

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЁВА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»

А.В. КИРИЛЛОВ

СТАТИСТИКА

Часть 1.

Общая теория статистики

*Утверждено Редакционно-издательским советом университета
в качестве учебного пособия*

САМАРА
Издательство СГАУ
2012

УДК СГАУ: 31
ББК 60.6
К 431

Рецензенты: д-р филол. наук, проф. А. Н. С о р о ч а й к и н,
д-р экон. наук, проф. М. И. Г е р а с ь к и н.

К 431 ***Кириллов А.В.***
Статистика. Часть 1. Общая теория статистики: учеб. пособие /
А.В. Кириллов. – Самара: Изд-во СГАУ, 2012. – 112 с.

ISBN 978-5-7883-0893-7

Учебное пособие составлено в соответствии с требованиями для высших учебных заведений на базе последних материалов, статей и документов. Книга содержит сведения, необходимые для изучения дисциплины «Статистика». Рассмотрены вопросы и основные понятия общей теории статистики. Приведены примеры, поясняющие некоторые приемы и способы используемых алгоритмов.

Предназначено для студентов экономических вузов. Подготовлено на кафедре менеджмента.

УДК СГАУ: 31
ББК 60.6

ISBN 978-5-7883-0893-7

© Самарский государственный
аэрокосмический университет, 2012

ОГЛАВЛЕНИЕ

Введение	6
Часть I. Общая теория статистики.....	7
Глава 1. Содержание, предмет и задачи курса статистики.....	7
1.1. Филологическое толкование статистики.....	7
1.2. Понятие о термине статистика. Задачи статистики.....	7
1.3. Краткие сведения из истории статистики.....	8
1.4. История зарождения статистической науки.....	10
Глава 2. Основные понятия статистики.....	11
2.1. Статистическая совокупность.....	11
2.2. Статистические признаки. Виды классификации.....	12
2.3. Закон больших чисел. Статистическая закономерность.....	13
Глава 3. Статистическое исследование.....	13
3.1. Задачи статистического исследования.....	13
3.2. Стадии статистического исследования.....	14
3.3. Статистическое наблюдение.....	14
3.4. Виды и способы статистического наблюдения.....	15
3.4.1. Выборочные наблюдения.....	17
3.4.2. Способы проведения статистического наблюдения.....	19
3.5. Программа статистического наблюдения.....	19
3.6. Организационные формы наблюдения.....	20
3.7. Ошибки наблюдения.....	22
3.8. Методы проверки и способы контроля достоверности данных...	24
Глава 4. Статистическая сводка. Группировка. Статистические таблицы... ..	26
4.1. Статистическая сводка и группировка.....	26
4.2. Группировка. Виды.....	27
4.3. Типологическая группировка.....	27
4.4. Структурная группировка.....	28
4.5. Метод аналитических группировок.....	29
4.6. Ряды распределения.....	30
4.7. Графическое отображение статистических данных.....	32
4.7.1. Линейные изображения.....	32
4.7.2. Диаграммы.....	33
4.8. Статистические таблицы.....	36
4.8.1. Элементы статистических таблиц.....	36
4.8.2. Правила заполнения статистических таблиц.....	37
Глава 5. Обобщающие показатели статистики.....	39
5.1. Сущность обобщающих показателей.....	39
5.2. Абсолютные величины статистики.....	40
5.3. Относительные величины.....	40
5.4. Виды относительных величин.....	41

Глава 6. Средние величины.....	43
6.1. Понятие о средних величинах.....	43
6.2. Виды средних величин.....	44
6.2.1. Виды простых средних.....	44
6.2.2. Виды средних взвешенных.....	45
6.3. Техника вычисления среднеарифметической и ее свойства.....	46
6.4. Свойства средней арифметической.....	48
6.5. Расчет средней арифметической величины по правилу «моментов».....	49
6.6. Вычисление среднего уровня вариационного ряда, составленного из относительных величин.....	50
6.7. Основное правило выбора вида средних величин.....	51
6.8. Мода и медиана.....	51
Глава 7. Показатели вариации.....	54
7.1. Система показателей вариации.....	54
7.2. Свойства дисперсии.....	56
7.3. Дисперсия альтернативного признака.....	57
7.4. Виды дисперсии. Правила сложений дисперсий.....	58
Глава 8. Динамические ряды.....	61
8.1. Понятие о рядах динамики. Виды динамических рядов.....	61
8.2. Правила составления рядов динамики.....	62
8.3. Аналитические показатели ряда динамики.....	63
8.4. Аналитические показатели ускорения (замедления).....	64
8.5. Средние показатели динамического ряда.....	66
Глава 9. Способы обработки динамических рядов.....	68
9.1. Способ укрупнения интервалов.....	68
9.2. Способ характеристики ряда средними величинами.....	69
9.3. Метод скользящей средней.....	69
9.4. Приведение динамических рядов к общему основанию.....	71
9.5. Смыкание динамических рядов.....	71
Глава 10. Приемы обработки статистической информации.....	72
10.1. Аналитическое выравнивание.....	72
10.2. Интерполяция и экстраполяция.....	75
10.3. Приемы коррелирования динамических рядов.....	76
10.4. Статистическое изучение сезонных колебаний.....	77
Глава 11. Индексы. Общие понятия об индексах.....	80
11.1. Сущность экономических индексов.....	80
11.2. Виды индексов.....	81
11.3. Основные методологические положения теории индексов.....	81
11.4. Индивидуальные индексы.....	82

11.4.1. Индивидуальный индекс физического объема отдельных видов продукции i_q	82
11.4.2. Индивидуальный индекс цен на отдельные виды продукции i_p	82
11.5. Общие индексы.....	83
11.5.1.Общий индекс стоимостного объема продукции I_{pq}	84
11.5.2.Общий индекс физического объема продукции I_q	84
11.5.3.Общий индекс оптовых цен I_p	85
11.5.4. Правило выбора весов в общих индексах.....	87
11.5.5. Способы построения общих индексов	88
11.6. Отдельные виды индексов, употребляемые в статистике России.....	89
11.6.1. Индекс стоимостного объема товарооборота I_{qp}	89
11.6.2. Индекс физического объема товарооборота I_q	89
11.6.3. Индекс оптовых цен I_p	89
11.6.4. Индекс розничных цен I_r	90
11.6.5. Индекс себестоимости продукции.....	90
11.6.6. Индекс заработной платы.....	90
11.6.7. Индекс урожайности.....	91
11.6.8. Индекс производительности труда.....	91
11.7. Индексный анализ динамики среднего уровня.....	93
11.7.1. Индексы постоянного состава.....	94
11.7.2. Индексы переменного состава.....	94
11.8. Особенности индексов выполнения плана.....	96
11.9. Территориальные индексы.....	98
11.10. Взаимосвязанные индексы.....	99
11.11. Индексы с постоянными и переменными весами.....	101
Глава 12. Статистическое изучение связи показателей коммерческой деятельности.....	103
12.1. Взаимосвязи показателей коммерческой деятельности и задачи статистики.....	103
12.2. Методы корреляционно-регрессионного анализа показателей коммерческой деятельности.....	105
12.3. Множественная регрессия.....	107
СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ.....	110

ВВЕДЕНИЕ

В настоящее время в эпоху глобализации возрастает интерес к информации, к статистическим данным, к статистике. Статистические данные, отражающие изменения социально-экономических процессов, протекающих в обществе, становятся необходимым атрибутом, условием объективного восприятия развития общества.

Знание статистических законов, методологии проведения статистических исследований, методов обработки статистических данных, приемов проведения обоснованных прогнозов позволит наиболее полно сформировать представление об окружающем мире, о протекающих социально-экономических процессах.

Перед статистической наукой стоят сложные и ответственные задачи дальнейшего развития и совершенствования системы методологии обработки сбора, обработки и анализа статистических данных в целях повышения степени достоверности статистической информации и объективного представления картины окружающего мира.

Настоящее учебное пособие описывает первую часть дисциплины статистики: общую теорию статистики. Также приведены отдельные примеры решения задачи расчетных работ. Особенностью данного учебника является системное изложение вопросов общей теории статистики.

Общая теория статистики изложена в соответствии с программой изучения данной дисциплины и содержит рассмотрение комплекса вопросов по организации и проведению статистического исследования: сбор статистической информации (статистическое наблюдение), обработку статистической информации, анализ сводных показателей, методы прогнозирования исследуемых процессов. В пособии учтены результаты вопросов и тестов Министерства образования.

Предлагаемое учебное пособие по своему содержанию соответствует требованиям образовательного стандарта Министерства образования и науки РФ.

Часть I

ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Глава 1. СОДЕРЖАНИЕ, ПРЕДМЕТ И ЗАДАЧИ КУРСА СТАТИСТИКИ

1.1. Филологическое толкование статистики

Термин «Статистика» зародился от латинского слова «status» – положение или состояние дел и вещей, близкое к данному корню «stato» – государство, отсюда появился и термин «статистика» – состояние дел в государстве.

В науку этот термин введен в XVIII в. немецким ученым Готфридом Ахенвалем. Он представлял статистику как совокупность сведений о государствах и народах.

1.2. Понятие о термине статистика. Задачи статистики

В настоящее время данный термин имеет многогранное значение. Он употребляется в трех значениях:

1) Статистика как отрасль практической деятельности в органах государственной статистики, которая имеет своей целью сбор, обработку, анализ и публикацию массовых данных о самых различных явлениях общественной жизни (в этом смысле «статистика» выступает как синоним словосочетания «статистический учет»);

2) Статистика как отрасль экономической науки. Это базис всей экономической науки в целом. Это особая научная дисциплина и соответственно учебный предмет в высших и средних специальных учебных заведениях.

3) Под статистикой понимается система показателей, характеризующая ту или иную область общественных явлений, статистика в «узком» смысле слова.

Все три определения вместе представляют статистику в «широком» смысле слова.

Статистика – отрасль общественных наук, имеющая целью сбор, упорядочение, анализ и сопоставление фактов, относящихся к самым разнообразным массовым явлениям.

Под предметом статистики понимается количественная сторона массовых общественных явлений в постоянной связи с их содержанием или количественной стороной, а также количественное выражение закономерностей общественного развития в конкретных условиях места и времени. Каждая отрасль имеет свою статистику. Статистика развивается как отдельная наука. Отраслевая статистика дополняет теорию статистики.

Теория статистики является основополагающей дисциплиной и служит фундаментом для применения статистического метода анализа для хозяйственных субъектов. На любом уровне и в любой сфере эффективность использования статистики во многом определяется качеством исходной информации.

Статистика как наука подразделяется:

- на общую теорию статистики,
- макроэкономическую статистику,
- экономическую статистику,
- отраслевую статистику.

Статистика изучает массовые общественные явления. Массовые общественные явления – это явления, которые встречаются в больших количествах, но отличаются друг от друга величиной определенного признака.

Статистика изучает закономерности развития с помощью количественных показателей, поэтому она определяет размеры, уровни и величины различных явлений, изучает структуру, динамику, взаимодействие явлений.

Задачи статистики:

1. Переход от отраслевого принципа сбора информации к статистике предприятия. Статистика предприятия дает достаточную информацию для взаимосвязанного анализа функционирования рынков труда, капитала, товаров и услуг.

2. Переход на качественно новые международные стандарты в области статистики цен, занятости, стоимости рабочей силы и уровня жизни населения.

3. Создана основа для широкого применения разнообразных математических и статистических методов для расчетов и контроля надежности статистических данных.

4. Создана система статистических показателей для трех уровней управления: федерального (макроэкономические показатели), территориального (отрасли и сектора экономики), предприятий (статистика предприятий).

1.3. Краткие сведения из истории статистики

Местом зарождения статистической практики считают Древний Китай. Так, известно, что еще 23 века до н.э. проводились переписи населения в Китае, причем проводились регулярно, т.к. в Китае всегда существовала проблема перенаселения. В Древней Персии велась административная статистика, осуществлялось сравнение военного потенциала разных стран, наличие сведений о материальных ресурсах, дорогах, военных поселениях.

В Древнем Египте зародилась земельная (кадастровая) статистика, развитие ее было обусловлено тем, что 90% населения проживало у русла р. Нил.

Наиболее высокого уровня достигла статистика в Древнем Риме. Этого уровня Западная Европа достигла в XVII веке. Статистика развивалась в виде censов (учет имущества граждан для начисления). Учет осуществлялся относительно свободных граждан Рима (учитывалось наличие земель, построек, имущества, рабов).

В дальнейшем многие правители использовали элементы статистики для решения собственных задач. Так, нормандский герцог Вильгельм, завоевав Англию, провел перепись населения («Книга страшного суда»), ввел непо-

мерные налоги. Татары, захватив Русь и Киев, провели перепись и на основании данной переписи обложили Русь данью.

В России элементами статистики можно считать деятельность боярина Морозова по созданию «Опытных книг» как результат проведения инвентаризации его имущества.

Монастырские книги явились своеобразным первоисточником к возникновению основ бухгалтерского учета. Монастыри представляли собой самостоятельное «государство в государстве». Именно в монастырях зародились элементы организации принципов самофинансирования и самоокупаемости, т.к. монахи вели учет поступлений и расходов и добивались согласования данных статей в бюджете монастыря.

Особенно бурного расцвета достигло развитие статистики в России при царе Петре I. Петр I при проведении реформ государства испытывал острую недостатку в финансовых средствах, одним из источников средств было введение налогов. Петр вводил различные налоги, например налог на бороды, для этого проводилась перепись по данному вопросу, рассчитывалась величина налога.

В дореволюционной России развитие статистики осуществлялось в двух формах: правительственная и земская статистика. Земская статистика образовалась после аграрной или земской реформы 1861 г. (19 февраля 1861 г. было подписано Александром II «Положение о крестьянах, вышедших из крепостной зависимости», в тот же день последовал манифест «О всемилостивейшем даровании крепостным людям прав состояния свободных сельских обывателей».) Земская статистика велась в земствах (выборных органах местного самоуправления). Вопросами земской статистики являлись исследования промыслов и сельского хозяйства.

Первая перепись населения России была проведена в 1897 году.

После 1917 года при советской власти было создано ЦСУ (Центральное Статистическое Управление), которое находилось на правах наркомата (народных комитетов – аналог министерств). В 20-е годы советская статистика была одной из самых передовых в мире. В 1924 году был сформирован первый народнохозяйственный баланс государства. В его разработке непосредственное участие принимал В.В. Леонтьев (впоследствии лауреат Нобелевской премии). Серьезный урон статистике принесли годы руководства Сталина. В это время против статистиков, которые несли справедливую, но горькую правду, применялись репрессии.

Только в 1946 году вновь было создано ЦСУ, но отставание России было существенно, в это время за рубежом широко развивались международные методологии статистики – СНС (система национальных счетов). В России же использовалась концепция материального производства, согласно которой национальный доход формировался в сфере производства (по К.Марксу и А.Смиту).

В 1988 году ЦСУ преобразовано в Госкомстат на правах Госкомитета и начался переход на систему национальных счетов, медленное ее внедрение связано с тем, что внедрение системы начали «сверху», а не с предприятий.

1.4. История зарождения статистической науки

Статистика как наука зародилась в XVII веке. У истоков статистической науки стояли две школы – немецкая описательная и английская школа политических арифметиков. Это был **первый этап** развития статистики.

Представители описательной школы считали, что задачей статистики является описание достопримечательностей государства: территории, населения, климата, вероисповедания, ведения хозяйства и т.п. – только в словесной форме, без цифр и вне динамики, т.е. без отражения особенностей развития государств в те или иные периоды, а только лишь на момент наблюдения. Видными представителями описательной школы были Г. Конринг (1606–1661 гг.), Г. Ахенваль (1719–1772 гг.) и др.

Политические арифметики ставили целью изучать общественные явления с помощью числовых характеристик – меры веса и числа. Это был принципиально новый этап развития статистической науки по сравнению со школой государственоведения, так как от описания явлений и процессов статистика перешла к их измерению и исследованию. Политические арифметики видели основное назначение статистики в изучении массовых общественных явлений, осознавали необходимость учета в статистическом исследовании требований закона больших чисел, поскольку закономерность может проявиться лишь при достаточно большом объеме анализируемой совокупности. Основателем этого направления был Вильгельм Петти (1623–1687 гг.). Наиболее известные его работы «Трактат о налогах и сборах» (1662 г.) и «Политическая арифметика» (1676 г.) История показала, что последнее слово в статистической науке осталось именно за школой политических арифметиков.

Вторым этапом развития статистики можно считать развитие учения бельгийского статистика Адольфа Кетле, основоположника учения о средних величинах (первая треть XIX в.). Он был сторонником и создателем теории устойчивости статистических показателей. Для выявления закономерностей общественной жизни А. Кетле широко использовал аппарат средних величин. Однако в его работах просматривались две ошибки. Первая: он абсолютизировал значение средних величин, приписывал свойства средних каждому. Вторая: Адольф Кетле смешивал законы общественной жизни с законами естественных процессов. Например, он строил модель общества исходя из принципов закона всемирного тяготения.

Третий этап статистики – математическое направление. Развивалось в работах исследователей-математиков: К. Пирсон (1857–1936 гг.) развивал теорию корреляции, разрабатывал количественные оценки связи между явлениями, В. Госсет (1876–1937 гг.), более известный под псевдонимом Стьюдента, разрабатывал теорию малой выборки, Р. Фишер (1890–1962 гг.) – методы количественного анализа, методы нахождения оценок неизвестных параметров, и др.

Прогрессу статистической методологии способствовали труды российских статистиков – А.А. Чупрова (1874–1926 гг.), В.С. Немчинова (1894–

1964 гг.), С.Г. Струмилина (1877–1974 гг.) и др. Развитие статистической науки, расширение сферы практической статистической работы привели к изменению содержания самого понятия «статистика».

Как и всякая наука, статистика имеет свой предмет изучения. Статистика изучает количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной, исследует количественное выражение закономерностей общественного развития в конкретных условиях места и времени.

Глава 2. ОСНОВНЫЕ ПОНЯТИЯ СТАТИСТИКИ

2.1. Статистическая совокупность

Статистика изучает массовые явления, совокупности. Совокупности можно считать *статистическими*, если они обладают следующими характерными чертами:

- объективность существования;
- качественная однородность образующих их явлений, единиц;
- варьирование изучаемых признаков в пространстве или во времени.

Иными словами, *статистическая совокупность* – это множество единиц изучаемого явления, объединенных одной качественной основой, но отличающихся друг от друга отдельными свойствами, особенностями.

Качественная однородность – принадлежность единиц (явлений) к одному определенному типу, что обеспечивается в равной степени у всех единиц совокупности наличием основных, характерных признаков данного типа.

Например: студенты очной формы обучения – качественная группа, по ряду признаков (не работают, имеют право получать пособие (стипендию) от государства) она отличается от других качественных групп (например, пенсионеры – они тоже не работают и тоже получают пособие (пенсию) от государства).

Количественная однородность – близость количественных значений основных признаков единиц совокупности друг к другу. Например, рост обычного человека (мужчины) находится в интервале от 165 до 195 см.

При подготовке статистического наблюдения необходимо определить такие понятия, как объект наблюдения, единицу наблюдения, единицу совокупности.

Объект наблюдения – совокупность общественных процессов или явлений, подвергающихся статистическому исследованию. Пример, объект наблюдения – промышленность – совокупность предприятий. Объект наблюдения при переписи населения – само население.

Единица наблюдения – составная часть объекта наблюдения, представляющая источник информации. Пример, предприятие, артель – это единицы наблюдения.

Единица совокупности – составной элемент объекта наблюдения, основа счета, который обладает признаками, подлежащими регистрации в процессе наблюдения.

Единицы совокупности обладают свойствами и особенностями, выражающими их характер и связи с окружающим миром. Эти свойства, особенности единиц совокупности называются *признаки*. *Признаки* – непосредственный объект статистического измерения.

2.2. Статистические признаки. Виды классификации

Статистические признаки имеют несколько видов классификации:

А) Классификация по изменчивости. Признаки разделяются на варьирующие и статические:

Варьирующиеся – признаки, значения которых меняются в пространстве и во времени.

Статические – признаки, которые не меняют своих значений.

Статистика изучает только варьирующиеся признаки. Значения признака у отдельных единиц называются *вариантами*. *Вариант* – конкретное значение изучаемого признака. Он принимает любые значения в пределах вариации.

Б) Классификация по характеру отображения свойств единиц изучаемой совокупности. Признаки разделяются на количественные и атрибутивные:

Количественные – такие признаки, варианты, которые выражаются числовой характеристикой или мерой. Пример: рост, вес, длина и т.п.

Атрибутивный (качественный) признак – такой признак, варианты которого, характеризуя особенности отдельных единиц, не имеют количественного выражения.

Примеры: национальность, профессия, пол, социальная принадлежность.

В) Классификация по степени важности. Выделяют существенные и несущественные признаки:

Существенные – такие признаки, которые выражают сущность совокупности, определяют однокачественность и этим обособляют исследуемые единицы совокупности от окружающего мира.

Несущественные – признаки, которые не характерны для качества совокупности, не показательны для определяющих ее закономерностей. Они носят добавочные, второстепенные сведения.

Г) Классификация по доступности. Признаки подразделяются на первичные и вторичные:

Первичные признаки непосредственно характеризуют единицу совокупности в целом, как нечто целое неделимое.

Пример: число рабочих в производстве.

Вторичные признаки определяют структуру единиц совокупности, отражают различные стороны их существования и выступают как относительные характеристики (средние) того или иного сложного явления.

Пример: стоимость основных фондов на одного рабочего.

2.3. Закон больших чисел. Статистическая закономерность

Главным обобщением опыта исследований любых массовых явлений служит закон больших чисел. Данный закон отражает смысл проведения исследований. Суть его в следующем. Каждое отдельное единичное явление, рассматриваемое как одно из явлений данного рода, содержит элемент случайного – оно могло быть или не быть, быть таким или иным. Однако при объединении большого числа таких явлений в общих характеристиках всей их массы случайность исчезает тем в большей мере, чем больше соединено единичных явлений, чем больше учитывается и исследуется единиц совокупности.

Пример: Средний рост мужчин в России около 178 см. Однако, если мы будем опрашивать конкретных людей, то встретить человека именно с этим ростом достаточно сложно (в каждом конкретном случае будут присутствовать отклонения от среднего роста в ту или иную сторону). При нахождении среднего роста и обследовании группы мужчин за счет погашения взаимоклонуений значение среднего роста будет приближаться к среднему по России при увеличении выборки.

Статистическая закономерность

Закономерность, которая проявляется в большой массе явлений через преодоление случайных отклонений, свойственных единичным явлениям, называется *статистической закономерностью*.

В одних ситуациях нужно только проверить предположение, в других случаях необходимо найти данную закономерность.

Пример – спрос на товар – явление массовое. Можно сказать, что уменьшение цены товара увеличивает его спрос.

Глава 3. СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ

3.1. Задачи статистического исследования

Для получения количественных характеристик социально-экономических процессов, явлений, их места в общей системе общественного производства необходимо проведение всестороннего и глубокого исследования, а именно статистического исследования.

Статистическое исследование – это комплекс действий, заключающийся в выяснении тех закономерностей, которые могут быть установлены на основе данных массового наблюдения.

Прежде чем приступить к использованию различных способов и методов статистической методологии, необходимо располагать исчерпывающими и достоверными данными об объекте исследования, знать цель и задачи исследования, какого рода можно сделать выводы на основе анализа, какова ценность и практическая значимость.

Организация статистического исследования предполагает всестороннюю проработку вопросов его подготовки и проведения, а также анализа полученных результатов.

3.2. Стадии статистического исследования

Выделяют несколько стадий статистического исследования (рис. 2.1):

- сбор статистической информации (статистическое наблюдение) и первичная обработка;
- систематизация и дальнейшая обработка данных, полученных в результате статистического наблюдения, на основе их сводки и группировки;
- обобщение и анализ результатов обработки статистических материалов, формулировка выводов и рекомендаций по итогам всего статистического исследования.



Рис. 3.1. Стадии статистического исследования

3.3. Статистическое наблюдение

Первый и исходный этап статистического исследования – *статистическое наблюдение*. На этом этапе идет формирование статистических данных или исходной статистической, формируется прочный фундамент будущего статистического здания.

Статистическое наблюдение – основа статистического исследования.

Однако не всякий сбор данных является статистическим наблюдением. О статистическом наблюдении можно говорить лишь тогда, когда присутствуют три важных признака: плановость, массовость и систематичность.

Плановость статистического наблюдения представляет собой подготовку и осуществление по заранее разработанному плану, который охватывает все вопросы методологии, организации и техники сбора статистической информации, контроля ее качества и достоверности, оформления итоговых документов для дальнейшей их сводки и обработки. План статистического наблюдения должен составлять неразрывную часть плана статистического исследования.

Массовый характер статистического наблюдения – охват возможно большего, массового числа случаев проявления данного явления или процесса для получения правдивых статистических данных для характеристики массы единиц совокупности.

Систематичность статистического наблюдения предполагает проведение исследования не стихийно, от случая к случаю, а систематически, либо непрерывно, либо регулярно, по возможности в равные промежутки времени – интервалы.

3.4. Виды и способы статистического наблюдения

На рис. 3.2 показаны основные виды статистических наблюдений. В зависимости от цели исследования существует несколько видов классификации видов наблюдения.

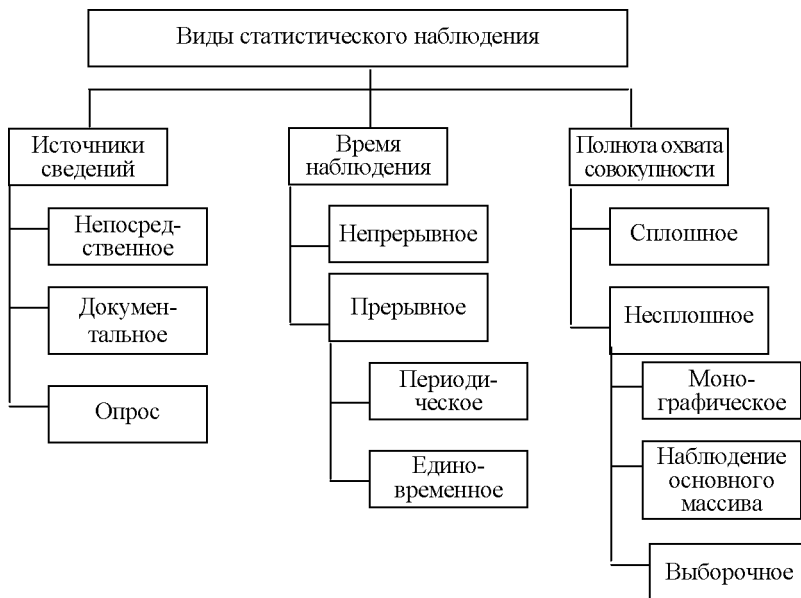


Рис. 3.2. Виды статистических наблюдений

Классификация видов наблюдения по способу сбора информации (источники сбора данных) включает следующие виды: *непосредственное наблюдение, документальное наблюдение и опрос.*

Непосредственное наблюдение представляет собой наблюдение, при котором происходит фиксирование фактов, лично установленных регистратором (счетчиком) в результате осмотра или измерения. Пример: замер времени трудовых операций при изучении их трудоемкости.

Документальное наблюдение предполагает, что в качестве источника информации выступают данные соответствующих документов. Пример: при составлении отчетности предприятия используются документы бухгалтерского отчета.

Опрос – это получение данных в форме ответов опрашиваемых лиц на поставленные вопросы по исследуемой теме. Пример: проведение опроса жителей города по животрепещущей проблеме.

Классификация по времени проведения наблюдения разделяет все наблюдения на *непрерывные* и *прерывные*.

Непрерывное (текущее) наблюдение – непрерывная регистрация фактов по мере их возникновения. Прослеживаются все изменения изучаемого процесса или явления. Пример: учет выработки продукции, учет отпуска материалов в производстве, отчетность.

Прерывные наблюдения проводятся либо регулярно через равные промежутки времени (периодические), либо одновременно – однократно (по мере надобности). К периодическим наблюдениям можно отнести перепись населения, ежегодную перепись скота, учет товарных остатков на момент переоценки на начало или конец месяца. К единовременным – это наблюдения (инвентаризация) на момент покупки или продажи объекта. Встречаются ситуации, когда единовременное наблюдение может перейти в регулярное.

Классификация по степени охвата единиц исследуемой совокупности разделяет все статистические наблюдения на *сплошные* и *несплошные*:

Сплошное наблюдение – это наблюдение, при котором происходит регистрация всех без исключения единиц исследуемой совокупности. Преимущества: полнота информации об изучаемых явлениях или фактах, наибольшая достоверность. Недостатки: необходимо привлечение больших финансовых, временных и человеческих ресурсов.

Несплошное наблюдение представляет статистическое наблюдение, при котором обследованию подвергаются не все единицы исследуемой совокупности, а только часть. Оно организуется как учет части единиц совокупности, на основе которой можно получить обобщающую характеристику для всей совокупности. Несплошное наблюдение охватывает определенную часть изучаемой совокупности. К видам несплошного наблюдения относятся: *монографические наблюдения, наблюдение основного массива и выборочные наблюдения*.

Монографическое наблюдение – это глубокое и всестороннее исследование отдельных единиц совокупности. монографическое наблюдение проводится для выявления зарождающихся тенденций, исследования новых форм и организации бизнес-процессов. Например – исследование передового опыта бригады (стахановское движение).

Наблюдение основного массива – обследование только тех единиц совокупности, у которых изучаемые признаки обладают наиболее существенными по значимости значениями. Пример – проводится обследование только 15-20% предприятий, выпускающих 90-95% всей продукции.

3.4.1. Выборочные наблюдения

Выборочное – наблюдение части единиц исследуемой совокупности, выделенных некоторым методом. В зависимости от метода выделяют следующие выборочные исследования: *случайная, механическая (систематический отбор), типическая (стратифицированная), серийная (гнездовая), многоступенчатая (компонентная) выборки.*

Случайная выборка

Главный принцип *случайной выборки* – равенство шансов каждой единицы совокупности попасть в выборочную совокупность методом случайного отбора (метод Монте-Карло, датчик случайных чисел, статистические таблицы, лотерея). Данный вид наблюдения широко применяется в промышленности для контроля качества продукции, в сельском хозяйстве для анализа цен на продукцию, в торговле – для характеристики спроса населения.

Использование различных типов случайного отбора позволяет несколько сгладить некоторые трудности, возникающие при проведении собственно случайного отбора. Например, некоторые типы случайного отбора позволяют упростить организацию опроса, но главное – это то, что они увеличивают эффективность выборки. Однако модификации случайного отбора не могут преодолеть всех трудностей, связанных со случайной выборкой. Это связано с тем, что все они являются *разновидностями именно случайного отбора* и в них используется принцип случайности. Из этого следует, что проводить любой случайный отбор невозможно без списка элементов исследуемой (генеральной) совокупности. Более того, большинство типов случайного отбора приводят к тем же трудностям при организации опроса, что и при собственно случайной выборке. Главное, чего достигают эти модификации случайного отбора, так это увеличения точности выборки.

Механическая выборка

Механическая выборка является упрощенным вариантом случайной выборки. Отбор единиц в выборочную совокупность производится на основе общего списка (генеральной совокупности) через равные интервалы (шаг выборки). При этом шаг выборки рассчитывается путем деления размера генеральной совокупности на объем выборки. Так, при 2%-й выборке выбирается каждая пятидесятая единица, при 10%-й выборке выбирается каждая десятая единица и т.д. В соответствии с данным подходом вся совокупность как бы механически разбивается на равновеликие группы, из каждой группы отбирается одна единица.

Процедура проведения механической выборки менее громоздка, чем проведение случайной выборки. Хотя применение компьютеров практически нивелирует это преимущество.

Механическая выборка может быть как более точной, так и менее точной по сравнению со случайной выборкой.

Типическая (стратифицированная) выборка

При *типической (стратифицированной) выборке* генеральная совокупность разделяется на однородные группы (страты) по какому-либо признаку.

Поскольку группы однородны, это означает, что единицы в группе похожи друг на друга, а сами группы сильно различаются между собой. Из выделенных страт производится случайный отбор единиц по принципам случайной или механической выборки. Типическая выборка обычно применяется при изучении сложной статистической совокупности, например, при выборочном обследовании труда работников, состоящих из отдельных групп по квалификации. Стратифицированная выборка в любом случае оказывается точнее собственно-случайной. Этот метод особенно хорош, когда генеральная совокупность неоднородна.

Однако стратифицированная выборка может быть применена лишь при наличии дополнительной информации о генеральной совокупности (например, нам необходимо процентное соотношение мужчин и женщин, в случае, если мы хотим стратифицировать выборку по полу). Отсутствие такой информации делает применение стратифицированной выборки невозможным. Еще один недостаток стратифицированного отбора – это возможность систематической ошибки.

Серийная (гнездовая) выборка

При проведении *серийной (гнездовой) выборки* принципы построения противоположны стратифицированной выборке. Из генеральной совокупности отбираются не отдельные единицы, а целые группы (гнезда). Группы отбираются случайно (равновероятностно). Внутри выделенной группы обследуются все без исключения единицы, проводится сплошное в рамках группы наблюдение.

Серийные выборки часто применяют в торговле, т. к. многие товары для их транспортировки, хранения и продажи упаковываются в пачки. Поэтому при контроле качества рациональна проверка качества отдельных упаковок товара.

При обследованиях городских жителей данный метод также удобен. В этом случае городские микрорайоны отбираются по равновероятностной (случайной) процедуре, а далее проводится сплошной опрос жителей микрорайонов.

Главное преимущество данного метода состоит в том, что он гораздо проще в организационном плане. Действительно, гораздо проще выбрать несколько групп и опросить их целиком, чем бегать за каждым респондентом. Это дает нам выигрыш в средствах и во времени. Но при этом необходимо следить, чтобы количество групп в генеральной совокупности было достаточно большим, иначе ни о каком принципе случайности не может быть и речи.

Многоступенчатая (компонентная) выборка

Построение *многоступенчатой (компонентной) выборки* основывается на описанных выше принципах. Этот тип выборки применяют, когда возникает необходимость в проведении поэтапного отбора единиц из разных совокупностей. При этом генеральная совокупность сначала разбивается на группы. Затем производится отбор группы. Внутри данной группы выделяются отдельные сегменты и т.д. Этапов может быть большое количество исходя из задачи исследования.

3.4.2. Способы проведения статистического наблюдения (способы сбора статистических данных)

Выделяют четыре способа проведения наблюдений: *отчетный, экспедиционный, саморегистрация, анкетирование.*

Отчетный: представление информации в определенные сроки по определенной форме (например, ежеквартальное предоставление бухгалтерской отчетности, которая подготавливается по определенной форме к определенному сроку).

Экспедиционный: командировка к единице наблюдения специальных лиц (регистраторов, счетчиков) к месту событий, которые непосредственно фиксируют исследуемый факт или явление.

Саморегистрация подразумевает, что специальные работники снабжают исследуемых специальными бланками и дают необходимые инструкции, однако заполнение бланков делает сам обследуемый. Данный способ широко распространен за рубежом, когда обследуемый сам, своей рукой « в трезвом уме и светлой памяти» записывает необходимую информацию. Рядом с данным испытуемым или обследуемым могут присутствовать советники или помощники для пояснения отдельных вопросов, однако ответы пишет сам обследуемый.

Анкетное наблюдение – наблюдению подвергается специально выделенный круг лиц. Анкетирование чаще всего применяется как способ опроса населения в социальных исследованиях для нужд PR-кампаний, рекламы, маркетинга, политических кампаний. Анкеты могут быть открытыми и закрытыми. В открытых анкетах даются развернутые и подробные ответы, в закрытых – предлагается перечень возможных ответов, из которых опрашиваемый выбирает один, соответствующий его точке зрения. Заполнение анкет носит добровольный и анонимный характер. Учитывая тот факт, что опрашиваемый не несет ответственности за предоставляемую информацию, в анкете должны быть разработаны защиты по распознаванию искажения информации, например метод «перекрестных вопросов». Смысл данного способа заключается в том, что один и тот же вопрос, но в разных интерпретациях, может задаваться несколько раз в вопроснике. Если у человека нет твердого мнения или он скупавил в каком-либо ответе, то спустя некоторое время в подобном вопросе он может дать иной ответ. Эти разногласия будут впоследствии выявлены.

3.5. Программа статистического наблюдения

Программа статистического наблюдения должна включать перечень вопросов, на которые должны получить ответы. Содержание требований и вопросы должны включать исследуемые признаки и вопросы должны быть четкие, предельно ясные.

Для проведения статистического наблюдения – инструменты: формуляры и инструкции.

Формуляр – специальный документ, в котором фиксируются ответы на вопросы программы. Существуют формуляры-карточки и формуляры-список. Формуляр-карточка используется для регистрации одной единицы статистической совокупности. Формуляр-список применяется, когда собираются сведения о нескольких единицах совокупности.

Для заполнения формуляров используются коды. Код – условное цифровое или буквенное обозначение какого-либо признака. Например, коды единой государственной классификации.

Бланк статистического формуляра содержит части:

- заголовочную (наименование, почтовый адрес, коды, подчиненность, название и т. д.);
- содержательную (наименование, коды и числовые характеристики показателей);
- оформительную (должностные лица и даты подписания документа).

В программе должны быть указаны:

- цель и задачи статистического исследования,
- сроки проведения,
- время проведения,
- сведения об объекте и единицах наблюдения,
- место наблюдения.

Время наблюдения представляет собой период, в течение которого ведется наблюдение.

Критический момент статистического наблюдения – это такой момент (дата), к которому приурочивается статистическое наблюдение. Например, Всероссийская перепись населения проводилась несколько дней с 14 по 25 октября 2010 года, но все данные собирались по состоянию на 0 часов 14 октября 2010 года.

3.6. Организационные формы наблюдения

Выделяют две организационные формы наблюдения: это статистическая *отчетность* и *специальное статистическое наблюдение*.

Статистическая отчетность – такая организационная форма статистического наблюдения, при которой единицы наблюдения (предприятие, организация, стройка и т. д.) представляют сведения о своей деятельности в виде формуляров, жестко регламентированных по содержанию, форме и срокам.

Преимущества:

- обязательность представления;
- документообоснованность (необходимо наличие аналитической, пояснительной записки);
- юридически подтвержденная достоверность, которая гарантирует ответственность руководителей отчитывающихся предприятий и организаций (наличие подписи ответственного лица).

В зависимости от вида статистическая отчетность подразделяется:

- 1) по охвату объектов наблюдения;
- 2) по периодичности представления;
- 3) по способу представления;
- 4) по порядку прохождения.

1. По охвату объектов наблюдения отчетность делится на *типовую отчетность* и *специализированную*. *Типовая отчетность* – имеет одинаковое содержание и форму для всех предприятий (организаций) данной отрасли или всех отраслей (например, бухгалтерский баланс предоставляют все предприятия, независимо от их величины, количества сотрудников, вида деятельности, организационно-правовой формы). *Специализированная отчетность* – отражает специфику той или иной отрасли. Пример: предприятие "Норильский никель" в Красноярском крае является монополистом, поэтому подает дополнительные сведения о своей деятельности.

2. По периодичности представления отчетность может быть: еженедельная, декадная, 2-недельная, ежемесячная, квартальная, годовая.

3. По способу представления выделяют: почтовую, телеграфную (срочная), телетайпную, факс-модемную с подтверждением почтой.

Почтовая – обычная отсылка документации заказными письмами или бандеролями (возможна курьерская почта).

Срочная (другая схема) – формируются 2 потока:

- поток плановых данных (передача почтой);
- поток фактических данных (передача по телеграфу, телетайпу, электронной почте).

4. По порядку прохождения статистическая отчетность делится на *централизованную* и *нецентрализованную (децентрализованная)*.

Централизованная отчетность подразумевает передачу информации от предприятий и организаций в систему государственной статистики, обработку и рассылку министерствам, ведомствам, органам управления на местах.

Преимущества – объективность передачи информации, недостаток – низкая оперативность, длительное время обработки данных.

Нецентрализованная (децентрализованная) представляет собой первоначальную передачу информации в профильные министерства и ведомства, а затем в сжатом виде информация передается в органы государственной статистики. Преимущества – быстрота и оперативность, недостаток – возможна некоторая корректировка информации в связи с тем, что министерства и ведомства имеют отраслевые взаимоотношения с объектами наблюдений.

Вторая организационная форма наблюдения – специальное статистическое наблюдение. Специальные статистические обследования освещают моменты, которые не охвачены статистической отчетностью и служат средством для проверки и анализа материалов этой отчетности. Специальные статистические обследования дают дополнительный материал как для национально-

хозяйственного прогнозирования и оперативных мероприятий, так и для познания закономерностей развития экономики. В зависимости от задач исследования и конкретных условий статистическое наблюдение может быть единовременным или текущим. Единовременное наблюдение – запись признаков единиц наблюдения, приуроченная к данному "критическому моменту" времени. Единовременное наблюдение или учет состояния проводится через некоторые периоды времени, охватывает длительно существующую совокупность. Такое наблюдение проводится для определения численности, состава и качественных особенностей совокупности. Примерами специальных статистических обследований могут служить проводимые Всероссийские (всесоюзные) переписи населения, обследования бюджетов семей.

3.7. Ошибки наблюдения

Основная задача статистических органов – это достоверность и своевременность получения статистических данных. Под достоверностью понимается степень объективного отображения статистическими данными сущности явлений и процессов. Своевременность характеризует поступление данных в специально оговоренные сроки. В случаях нарушения данных принципов возникают *ошибки наблюдения*. Под *ошибкой наблюдения* понимается расхождение между результатами наблюдения и истинным значением величины. Учитывая многообразие встречаемых ошибок, используется следующая классификация ошибок наблюдения (рис. 3.3).

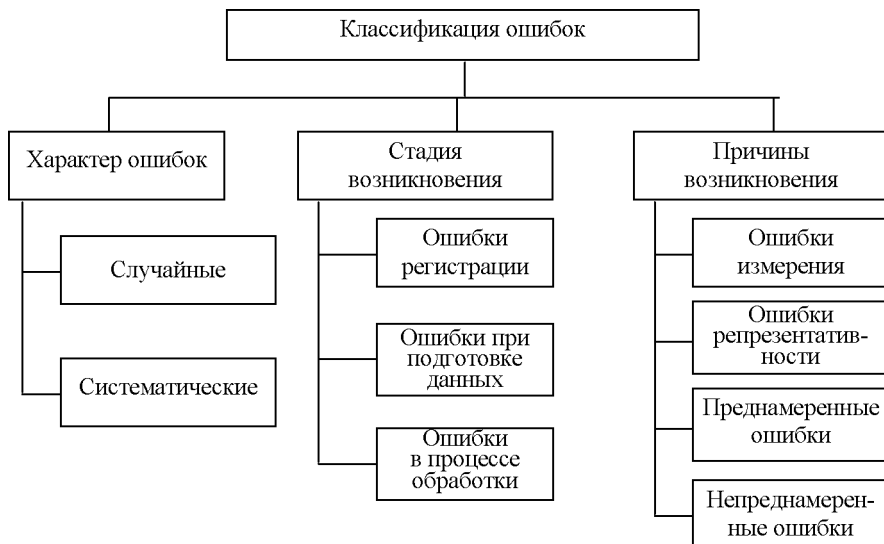


Рис. 3.3. Виды ошибок

По характеру ошибки разделяются на *случайные* и *систематические*.

Возникновение *случайных ошибок* обусловлено действием случайных факторов. Существует альтернативное определение *случайной ошибки* – ошибка измерения, возникающая из-за изменений респондента или ситуации, в которой производится измерение. Оба понятия в сущности похожи. Выявление случайных ошибок производится методами математической статистики.

Систематические ошибки представляют собой направленное искажение конечного результата наблюдения. Выявление этих ошибок значительно сложнее, используются методы системного анализа.

По стадии возникновения выделяют ошибки: *регистрации, при подготовке к машинной обработке, в процессе машинной обработки*.

Ошибки регистрации – ошибки, связанные с неточностями при записи данных в статистический формуляр (первичный документ, статистический отчет, бланк переписи и т. д.) или при вводе данных в ЭВМ через терминал (телетайп, дисплей). В настоящий момент существует термин – ошибка оператора, эта ошибка связана с вводом данных человеком.

Ошибки при подготовке к машинной обработке – некачественное состояние первичных документов (нечеткое заполнение) или машинных носителей дискет, флеш-карт, дисков.

Ошибки в процессе машинной обработки – это ошибки, связанные с тремя факторами: некорректные алгоритмы программ, перебой электропитания, наличие программ-вирусов. Первый тип ошибок обусловлен непроработанным алгоритмом программы. Второй тип ошибок обусловлен перебоями в электроснабжении, в настоящее время в качестве защиты необходимо использовать источники бесперебойного питания. Третий тип ошибок вызван наличием различных программ-вирусов, которые либо портят исходную информацию, либо разрушают аппаратные средства. В любом случае наносят непоправимый урон.

По признаку классификации причины возникновения ошибки разделяются на *ошибки измерения, репрезентативности, преднамеренные, непреднамеренные*.

Ошибки измерения – данные ошибки связаны с погрешностями при однократных статистических наблюдениях социально-экономических явлений. Для повышения достоверности данных необходимо как минимум трехкратное повторение наблюдений.

Ошибки репрезентативности – ошибки, связанные с несплошным (выборочным) наблюдением. Характерны для ситуаций, когда величина изучаемого признака в выборке или обобщающая характеристика выборки отличается от величины признака во всей совокупности или от обобщающей характеристики совокупности. Ошибки репрезентативности характерны только для несплошного наблюдения. Они возникают потому, что отобранная совокупность недостаточно точно воспроизводит (репрезентирует) исходную совокупность.

Ошибки преднамеренные – по причине сознательного искажения реальных данных с целью «приукрасить» действительное положение дел. Чаше всего встречаются два типа таких ошибок: завышение фактических показателей (приписки) в отчетных документах и занижение плановых показателей, сокрытие продукции.

Ошибки непреднамеренные – носят, как правило, случайный характер и могут рассматриваться как промахи в организации данной работы, ошибки менеджмента. Например, не была учтена низкая квалификация исполнителя или установлен завышенный норматив по обследованию респондентов и т.п.

Для устранения возможности появления ошибок используют методы проверки достоверности данных наблюдения.

3.8. Методы проверки и способы контроля достоверности данных

Целью проверки и контроля является установление достоверных данных, выявление причин их искажения или несвоевременного предоставления государственной отчетности.

Проверки бывают *тематические, комплексные, повторные*.

Тематическая проверка представляет собой проверку достоверности отчетных данных по одной или нескольким формам одной отрасли статистики.

Комплексная проверка – это одновременная проверка достоверности отчетных данных по двум или нескольким отраслям статистики.

Повторная проверка – это проверка или контроль исправления допущенных нарушений и ликвидация причин их возникновения. Обычно повторные проверки проводятся через год.

Способы контроля достоверности данных достаточно многочисленны (рис. 3.4). Среди них можно выделить три группы: *синтаксические (ручные), логические (с помощью ЭВМ) и арифметические (с помощью ЭВМ)*.

При синтаксических методах контроля наиболее часто используются:

- проверка *структуры документа* – проверяется состав и наличие всех составных элементов документа. Проверка носит достаточно формальный характер;

- проверка *полноты документа* – проверяются отдельные части документа по критериям достаточности и необходимости объема предоставления информации;

- *полнота заполнения строк* – проверяются отдельные строки документа на заполнение важными характеристиками.

Логические (с помощью ЭВМ) способы позволяют проводить различного рода сопоставления количественных и атрибутивных показателей с некоторыми нормативными. К логическим способам можно отнести следующие: *способ проверки соответствия кодов и исследуемых признаков, проверку отклонений исследуемых показателей от нормативных значений, проверку присутствия логических связей между показателями.*

Способ проверки соответствия кодов и исследуемых признаков заключается в проверке соответствия представленной информации соответствующему рубриктору (коду) и правильности отнесения данной информации к соответствующему разделу.



Рис. 3.4. Классификация способов контроля

Проверка отклонений исследуемых показателей от нормативных значений в настоящий момент становится распространенным способом. Это связано с тем, что подвергать полному контролю все показатели достаточно трудоемко, поэтому используют принцип достаточности. Если отчетный показатель укладывается в установленный норматив или диапазон, полная проверка с расчетами не производится, если показатель выпадает из установленного диапазона, то в этом случае данная составляющая подвергается полной проверке.

Проверка присутствия логических связей между показателями – широко распространенный прием подготовки бухгалтерской отчетности. Смысл данной проверки состоит в том, что расчет более сложного показателя формируется исходя из расчета (строк) представления других показателей.

Арифметические способы построены на принципе математических проверок отдельных составляющих. К ним можно отнести: проверку соответствия построчных контрольных сумм, контроль соответствия пографных контрольных сумм, балансовый метод (балансовая клетка).

Глава 4. СТАТИСТИЧЕСКАЯ СВОДКА. ГРУППИРОВКА. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

4.1. Статистическая сводка и группировка

В результате проведения наблюдения получаем материал, характеризующий отдельные единицы совокупности. У каждой единицы наблюдения много признаков и общие черты скрыты, спрятаны в индивидуальных показателях. Необходимо выделить общее, закономерное.

Возникает необходимость в специальной обработке – *сводке* материалов наблюдения.

Сводка – комплекс последовательных действий по обобщению конкретных единичных данных, образующих совокупность, в целях выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Все несущественное – отсеивается, а основное – выделено и зафиксировано.

Задача сводки – охарактеризовать исследуемый объект с помощью статистических показателей, выявить и измерить его существенные черты и особенности.

Четыре этапа сводки:

1. Систематизация данных, группировка материалов, собранных при наблюдении.

2. Уточняется система показателей, с помощью которых количественно характеризуются свойства и особенности предмета.

3. Вычисление скорректированной системы показателей и их обобщение по группам и по всей совокупности.

4. Представление результатов сводки в виде таблиц.

В сочетании всех этих этапов – понятие сводки в широком смысле слова.

Сводка в узком смысле слова – это подсчет итогов, данных.

Виды статистической сводки

Статистическая сводка имеет несколько разновидностей. Различают сводку *централизованную, децентрализованную и смешанную*.

Порядок осуществления *централизованной сводки* состоит в том, что весь первичный материал сосредотачивается и обрабатывается в едином центре, возглавляющем данные исследования.

Децентрализованная сводка заключается в том, что собранный материал обрабатывается на местах.

Каждая из этих двух видов сводки имеет свои преимущества и недостатки

В целях устранения недостатков и использования достоинств различных видов сводки стали широко применять *смешанную сводку*. Ее содержание состоит в том, что по важнейшим показателям сводка осуществляется децентрализованным способом, а по прочим показателям и их взаимосвязям сводка осуществляется централизованно.

4.2. Группировка. Виды

Для разделения совокупности единиц на однотипные группы статистика использует метод *группировки*.

Группировка – первый и наиболее важнейший этап сводки, позволяющий выделить из массы первичного статистического материала совокупности единиц, однородные группы единиц, обладающих общим сходством в количественном и качественном смысле.

Группировки решают 3 основные задачи:

- а) изучение социально-экономических типов явлений;
- б) изучение структуры однотипных совокупностей;
- в) установление связей и зависимостей между явлениями (причинно-следственные связи).

Группировки имеют несколько видов классификаций. В зависимости от степени сложности выделяют:

- простая группировка – выделение по одному признаку;
- комбинационная группировка – группа выделяется по одному признаку, подгруппа по другому признаку и т. д.;
- многомерная (n -мерная) группировка – используя теорию распознавания объектов, единицы совокупностей представляются в виде точек в n -мерном пространстве (n – число признаков, по которым характеризуется совокупность явлений).

В зависимости от вида решаемой задачи группировки разделяют на типологические, структурные и аналитические.

Отличие группировки от классификации

Классификация – общепринятый методологический стандарт расчленения совокупности на однородные группы, устанавливаемый на определенный промежуток времени.

Группировка проводится для конкретного исследования.

4.3. Типологическая группировка

Типологическая группировка позволяет выявлять качественно однородные типы явлений. Типологические группировки чаще всего используются первыми, т.к. они обрабатывают неоднородный статистический материал (неоднородные совокупности), который встречается чаще всего.

Характерным свойством типологических группировок является то, что они осуществляются по ограниченному числу наиболее существенных признаков. Типологические группировки широко применяются в социально-экономических исследованиях. Именно в выделении социально-экономических типов явлений, позволяющих проследить зарождение, развитие и их отмирание, состоит основная задача типологических группировок. Для данных группировок большое значение имеет выбор группировочного признака, чаще всего он атрибутивный. При атрибутивном признаке число групп определяется свойствами изучаемого явления: группировка населения

по половому признаку, предприятий и организаций – по формам собственности и т.п. Пример группировки по атрибутивному признаку: группировки по полу, профессии, национальности.

4.4. Структурная группировка

Группировка, с помощью которой производится расчленение однородной совокупности единиц на группы, характеризующие ее структуру по определенным признакам. Структурные группировки могут быть дискретные и интервальные. Дискретные группировки применяются в тех случаях, когда количество вариантов ограничено (обычно не более семи). Ниже приведен пример дискретной группировки.

Таблица 4.1. *Дискретная группировка студентов группы № 7113 СГАУ по возрасту*

№ п/п	Группы студентов по возрасту, лет	Количество студентов в группе	То же, в процентах (%)
1	2	3	4
1	17	1	3.2
2	18	7	22.6
3	19	21	67.8
4	20	2	6.4
5	Итого	31	100

Количество студентов в группе (3-й столбец табл. 4.1) называется *частотой*. *Частота* – это характеристика того, как часто встречается тот или иной вариант. В ряде случаев удобнее использовать понятие *частоты*. *Частость* – это частота, выраженная в процентах (столбец 4 в табл. 4.1).

В тех случаях, когда количество вариантов велико, используют интервальную группировку. Интервалом называется количественное значение признака, на основании которого исследуемые явления разбиваются на группы.

Границы интервалов устанавливают так, чтобы достигалась конечная цель группировок, внутри группы единицы совокупности должны быть однородны в количественном и качественном отношении.

Интервалы могут применяться равные и неравные.

В случае применения равных интервалов процедура выполнения структурной интервальной группировки заключается в следующем:

1. Выбираются экстремальные значения: $X_{\max}$, $X_{\min}$ ($X_{\max}$ – максимальное значение группировочного признака, $X_{\min}$ – минимальное значение признака).

2. Определяется количество интервалов. Для этого используется формула Стерджесса:

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

где n – число групп, N – количество исследуемых единиц совокупности.

3. Расчет величины интервала ведется по следующей формуле:

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

4. Прибавлением к минимальному значению величины интервала рассчитываются границы всех интервалов.

5. Единицы совокупности разносятся по группам.

Применение неравных интервалов обусловлено тем, что абсолютное изменение группировочного признака на одну и ту же величину имеет не одинаковое значение для групп с большим и малым значением признака.

Пример использования неравных (прогрессивно увеличивающихся) интервалов.

Численность работающих на промышленных предприятиях группируется:
до 100 человек

101 – 200

201 – 500

501 – 1000

1001 – 3000

3001 и выше.

При построении интервальной группировки могут использоваться открытые и замкнутые (закрытые) интервалы. Замкнутые интервалы имеют обе границы (нижнюю и верхнюю). Открытые интервалы имеют лишь одну границу, при этом употребляются термины «до», «менее», «более» и т.п.

Итак, отличительными признаками структурной группировки являются:

- 1) они осуществляются в однородной совокупности;
- 2) различия между группами в структурной группировке несущественны.
- 3) в структурной группировке применяется количественный группировочный признак.

4.5. Метод аналитических группировок

Это традиционный прием изучения причинно-следственных связей, которые обнаруживаются при параллельном сопоставлении значений признаков по группам.

Порядок построения аналитической группировки:

1. Из двух взаимосвязанных признаков выделяется факторный признак (причина) и результативный (следствие).
2. Группировка осуществляется по факторному признаку.
3. По каждой из групп находятся средние значения результативного признака.
4. Сопоставление параллельных рядов значений двух признаков, устанавливается характер зависимости.

Т а б л и ц а 4.2. *Зависимость успеваемости студентов от возраста*

№ п/п	Группы студентов по возрасту, лет	Количество	Средний балл
1	17	1	3.8
2	18	7	3.9
3	19	21	4.1
4	20	2	3.9
5	Итого	31	

4.6. Ряды распределения

На языке математической статистики группировка называется рядом распределения, т. е. ряд распределения – это упорядоченная группировка совокупности по одному признаку.

Ряды распределения бывают следующих видов: вариационные; атрибутивные.

Вариационные ряды – это ряды, построенные по количественному признаку. *Атрибутивные* – ряды, выполненные по качественному признаку.

Вариационные ряды, в свою очередь, подразделяются на дискретные и интервальные.

Дискретными рядами называются ряды распределения, в которых варианты способны принимать строго ограниченное число целых значений (табл. 4.3).

Т а б л и ц а 4.3. *Распределение рабочих производственной бригады по квалификации*

Группы по разряду	Численность группы	То же, в % к итогу	Нарастающий итог в %
2	6	15,0	15,0
3	8	20,0	30,0
4	14	35,0	70,0
5	9	22,5	92,5
6	3	7,5	100,0
Итого	40	100,0	100,0

Интервальные ряды – это такие ряды распределения, в которых интервалы указываются не конкретными числами, а в виде интервала значений с указанием наименьшего и наибольшего числа.

Интервальные ряды могут быть с прерывной и непрерывной вариацией.

Смысл непрерывной вариации выражается в том, что нижняя граница последующего интервала совпадает с верхней границей предшествующего интервала.

Смысл прерывной вариации выражается в том, что нижняя граница последующего интервала на единицу (на 0,1) больше верхней границы предшествующего интервала (см. табл. 4.5). Интервальные ряды с прерывной вариацией удобно использовать для простоты сбора информации, т.к. в этом случае

не возникает вопросов, в какой интервал включать единицы совокупности, варианты которых совпадают с границами интервалов.

При обработке данных в случае построения таблиц с интервалами с непрерывной вариацией существует правило, что если какое-либо обрабатываемое значение в точности совпадает с границами интервалов, то оно включается в тот интервал, для которого служит нижней границей, если не оговорено иное (например, не введен термин «включительно»).

Ниже приведен пример построения таблицы с непрерывной вариацией.

В табл. 4.4. приведено распределение промышленных предприятий по размеру производственных основных фондов. В данном конкретном случае таблица построена непрерывной вариацией границ, т.е. нижние границы последующих интервалов совпадают с верхними границами предшествующих интервалов.

Т а б л и ц а 4.4. *Распределение предприятий по величине основных фондов*

№ п/п	Группы предприятий по размерам основных фондов, млн. руб.	Число предприятий
1	До 0,5	3120
2	0,5 – 3,0	284
3	3,0 – 10,0	182
4	10,0 – 50,0	69
5	50,0 и более	51
	Итого	3706

В ряде случаев нижняя граница первого интервала и верхняя граница последнего являются открытыми, т. е. не указываются конкретным числом и обозначаются: "до", "менее", "свыше", "более". При обработке данных таблиц не всегда удобно использовать открытые интервалы, возникает неопределенность, какие значения включать в открытые интервалы. Существует правило закрытия интервалов: величина открытого интервала равна величине соседнего закрытого интервала.

Примером ряда распределения с прерывной вариацией служит распределение промышленных предприятий по численности работников.

Т а б л и ц а 4.5. *Распределение предприятий по численности персонала*

№ п/п	Группы предприятий по численности промышленного производственного персонала (человек)	Число предприятий
1	До 500	10500
2	501 – 3000	4500
3	3001 – 10000	7000
4	10001 и более	3200
	Итого	25200

Если исследуется однородная совокупность, то, как правило, осуществляется равноинтервальная группировка.

4.7. Графическое отображение статистических данных

Обычно проведение группировок и табличное представление статистических данных дополняют их графическим отображением. Схематически все множество графических представлений статистических данных разделяют на два класса: линейные изображения и диаграммы.

4.7.1. Линейные изображения

Линейный график – это график основополагающего типа. Он представляет собой ломаную линию, соединяющую отдельные значения исследуемого признака.

Важным достоинством линейных графиков является то, что на одном и том же поле графика можно изобразить несколько показателей, что позволяет сравнивать и выявлять специфику их развития во времени и характер изменения одного показателя по разным объектам в пространстве или территории.

К линейным графикам относятся *полигон распределения*, *кумулятивная кривая*, *огива*, *кривая концентрации*.

Полигоном распределения (полигоном частот) называют ломаную, отрезки которой соединяют точки. На оси абсцисс обычно откладывают варианты, на оси ординат – частоты.

При проведении структурной дискретной группировки графиком является полигон распределения. Так, для данных табл. 4.3 выполнена группировка рабочих по квалификации и на рис. 4.1 приведен полигон распределения рабочих. Графическое представление наглядно подчеркивает структуру изучаемой совокупности.



Рис. 4.1. Полигон распределения

Кумулятивная кривая – это ломаная, составленная по последовательно суммированным, т.е. накопленным частотам. При построении кумулятивной кривой дискретного признака на ось абсцисс наносятся значения признака, а ординатами служат нарастающие итоги частот. При построении кумуляты интервального признака на ось абсцисс откладываются границы интервалов и верхним значениям присваивают накопленные частоты. Кумулятивную кривую называют полигоном накопленных частот.

Для данных табл. 4.3 группировки рабочих по квалификации, в случае представления данных по накопленным частотам (столбец 4 табл.4.3), получим следующую кумулятивную кривую (рис. 4.2).



Рис. 4.2. Кумулятивная кривая нарастания квалификации рабочих

Если на ось ординат нанести значение признака, а накопленные частоты – на ось абсцисс, то получим кривую, называемую *огивой*.

4.7.2. Диаграммы

Диаграмма (от греческого *diagramma* – изображение, чертеж, рисунок) – это чертеж, на котором статистическая информация изображается посредством геометрических фигур, это графическое изображение, наглядно показывающее соотношение между сравниваемыми величинами. Диаграммы бывают различных видов: *полосовые (ленточные), столбиковые, квадратные, круговые, секторные, фигурные, радиальные, знаки Варзара*.

1. *Полосовые* – особенно наглядны при сравнении величин, связанных между собой в единое целое. Ширина полос должна быть одинаковой. По длине полосы разбиваются на части, пропорциональные изображаемым величинам.

Пример полосовой диаграммы показан на рис. 4.3.

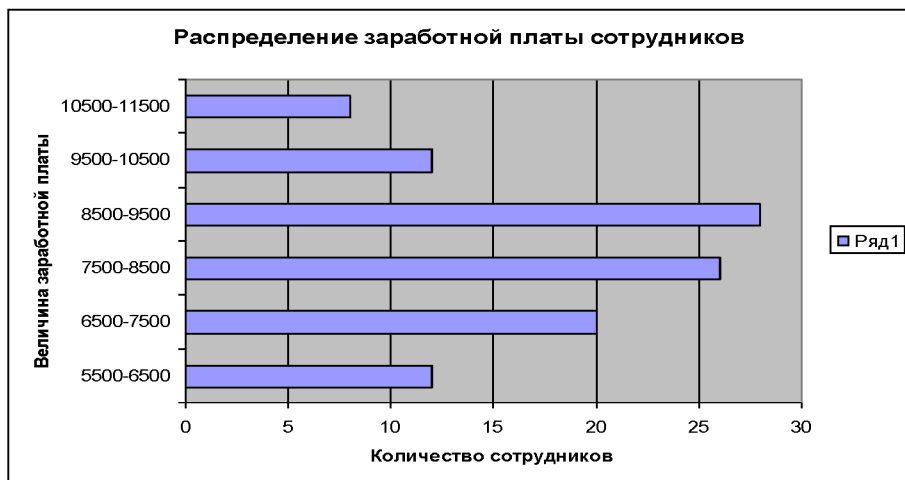


Рис. 4.3. Полосовая диаграмма распределения заработной платы сотрудников

2. *Столбиковые диаграммы* (рис. 4.4) предполагают вертикальное расположение полос (столбиков). Каждое значение изучаемого показателя изображается в виде вертикального столбика. По оси абсцисс размещается основание столбиков. Высота столбиков (в соответствии с принятым по оси ординат масштабom) должна строго соответствовать изображаемым данным. Расстояние между столбиками должно быть одинаковым. В качестве примера приведем графическое изображение использованных выше данных (см. данные табл.4.5).

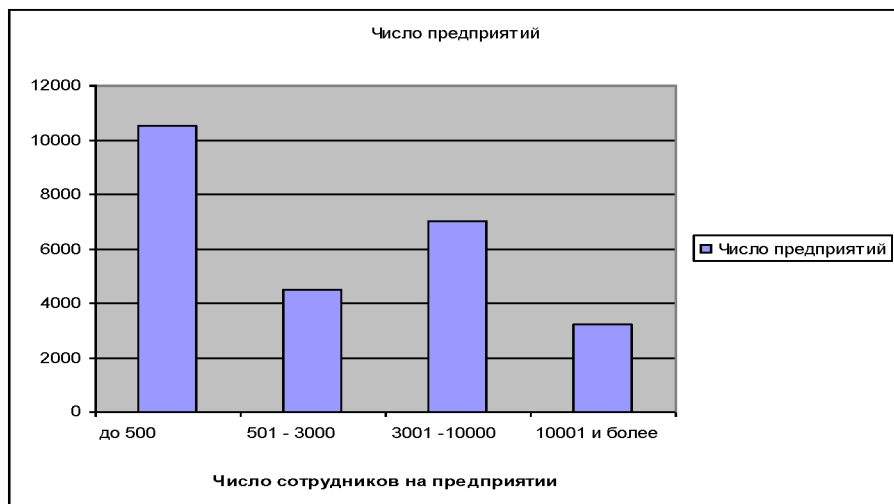


Рис. 4.4. Столбиковая диаграмма распределения предприятий

Основным видом столбиковых диаграмм являются *гистограммы*.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основанием которых служат частотные интервалы длины h , а высоты равны частоте m_i . Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними на расстоянии m_i проводят отрезки параллельные основанию. Пример построения гистограммы для распределения заработной платы среди сотрудников приведен на рис. 4.5.

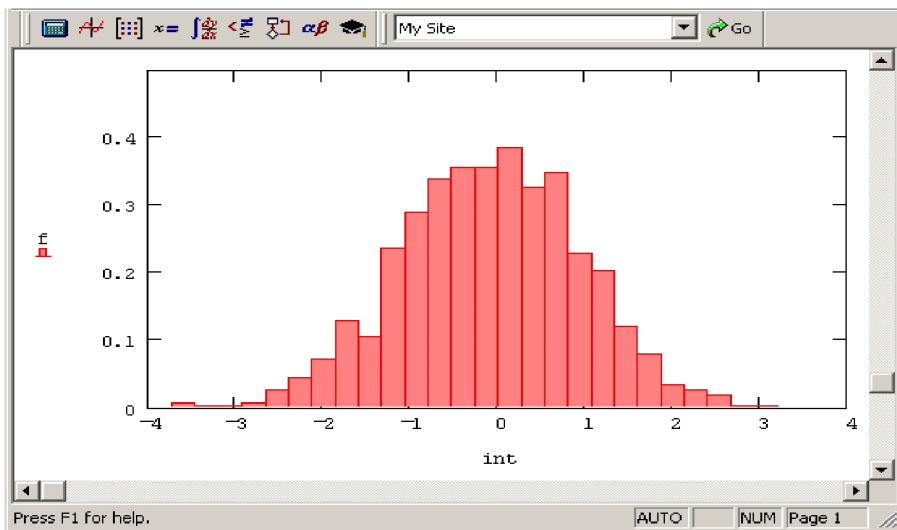


Рис. 4.5. Представление данных в виде гистограммы

Гистограммы чаще всего применяются при проведении структурной интервальной группировки.

3. При построении *квадратных и круговых диаграмм* площади квадратов или кругов выражают изображаемые величины.

4. *Круговые секторные* диаграммы применяют для графического изображения составных частей целого. Для их построения необходимо изображаемые данные выразить в градусах, т.к. 1% составляет 3,6 градусов, то соответствующие показатели для определения центральных углов надо умножить на 3.6. Чтобы легче различать сектора, используют различную расцветку или штриховку.

5. *Радиальные* – они строятся в полярной системе координат и используются для изображения признаков, периодически изменяющихся во времени (в большинстве своем сезонных колебаний). Вычисляется среднее арифметическое, затем строится окружность радиуса равного среднему арифметическому. Данная окружность делится на нужное число секторов (обычно 12) и на каждом радиальном направлении откладываются точки в соответствии со значениями $\bar{X}_i$.

6. *Фигурные* диаграммы строятся 2 основными способами: данные изображаются либо фигурами различных размеров, либо разной численностью фигур одинакового размера. Вторым способом чаще используется, каждая фигура содержит определенное число единиц признака и сравнение осуществляется по числу фигурок. При этом допускается дробление знака до половины.

7. Диаграмма "*знак Варзара*" названа по имени русского статистика В. Е. Варзара (1851-1940 гг.). С помощью данной диаграммы можно изображать многомерные признаки на плоскости посредством прямоугольников с разным соотношением между основанием и высотой. Одна из компонент признака изображается основанием прямоугольника, вторая – его высотой, третья – равна произведению двух других размеров получившейся площади.

4.8. Статистические таблицы

Результаты группировки и сводки принято оформлять в виде статистических таблиц. Таблица представляет способ наглядного и рационального изложения статистических данных.

Перед описательным изложением таблицы имеют следующие преимущества:

- 1) в наглядности и аналитичности, что обеспечивается путем параллельного изложения граф;
- 2) в экономии книжной площади, обеспечиваемой путем "сжатия" информации;
- 3) в доступности понимания материала, обеспечиваемой за счет сжатия информации в сказуемом таблицы.

4.8.1. Элементы статистических таблиц

Всякая таблица имеет следующие элементы:

- 1) строки;
- 2) графы;
- 3) клетки;
- 4) подлежащее таблицы, располагающееся в крайней левой стороне. Это наиболее общая и важная характеристика объекта;
- 5) сказуемое – показатели, которыми характеризуется объект;
- 6) заголовок таблицы, указание места и времени действия и характеристика задачи исследования;
- 7) нумерация строк и граф (графы нумеруются арабскими цифрами; строки нумеруются буквами славянского алфавита, римскими и арабскими цифрами);
- 8) итоговая графа и строка и балансирующая клетка;
- 9) макет таблицы- таблица без цифровых данных;
- 10) скелет таблицы – таблица без указания подлежащего и сказуемого.

4.8.2. Правила заполнения статистических таблиц

1. В таблице не должно быть исправлений, в случае совершения исправления необходимо сделать приписку «исправленному верить» с росписью руководителя.

2. Если по тем или иным причинам в таблице отсутствуют сведения, надо поставить три точки, т. е.

3. Если явление не имело места, то делается прочерк (–).

4. Если клетка не имеет смысла, то в ней ставится крест (на всю клетку).

5. Все округления в таблице делаются с одинаковой степенью точности. Чаще всего цифры указываются с точностью до десятых, при этом расчет ведется до сотых.

6. Если таблица или даже отдельные ее элементы заимствованы, то следует делать сноску с указанием первоисточника.

7. Если порядок расчета показателя в таблице отличается от используемого в таблице, то к таблице следует делать примечание, в котором разъясняется сущность применяемой методики.

Таблицы бывают следующих видов (рис. 4.6): *простые*; *группировочные*.

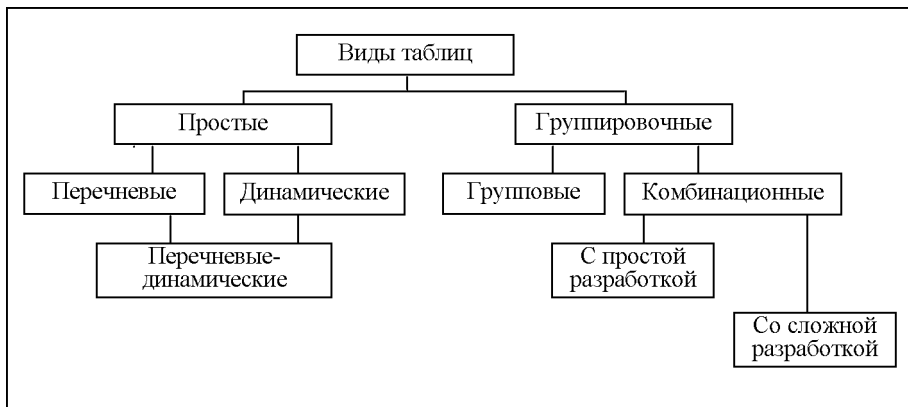


Рис. 4.6. Виды статистических таблиц

Простые таблицы – это таблицы, которые не являются результатом группировки, они подразделяются:

- на перечневые (содержащие перечень каких-либо элементов объекта),
- динамические (характеризующие динамику или процесс развития явления во времени).

В некоторых случаях таблицы содержат и перечень, и динамику явления и называются перечневыми-динамическими.

Группировочные таблицы – таблицы, которые являются результатом группировки.

Подразделяются:

- на групповые (выражают результат группировки по одному признаку),
- комбинационные (являются результатом группировки по двум и более признакам).

Комбинационные таблицы бывают с простой и сложной разработкой сказуемого.

При простой разработке сказуемого признаки рассматриваемого сказуемого располагаются параллельно, вне сочетаний друг с другом. Каждая клетка такой таблицы характеризуется двумя признаками: одним признаком из подлежащего, одним из сказуемого. В табл. 4.6 приведен пример построения такой группировки.

Т а б л и ц а 4.6. *Группировка населения по месту жительства, полу и трудоспособности*

Группы по месту жительства	Группы по полу		Группы по трудоспособности		
	мужской	женский	Моложе трудосп. возраста	Трудосп. возраст.	Старше трудосп. возраста
Городское население					
Сельское население			180 000		
Итого					

При сложной разработке сказуемого признаки сказуемого сочетаются не только с подлежащими, но и между собой, причем каждая клетка такой таблицы характеризуется тремя признаками:

одним – из подлежащего;

двумя – из сказуемого.

Т а б л и ц а 4.7. *Группировка населения по месту жительства, полу и трудоспособности*

Группы по месту жительства	Мужчины			Женщины		
	моложе 16 лет	16-59 лет	60 лет и старше	моложе 16 лет	16-54 лет	55 лет и старше
Городское население			340000			
Сельское население						
Итого						

Простую и сложную разработку может иметь не только сказуемое, но и подлежащее.

Глава 5. ОБОБЩАЮЩИЕ ПОКАЗАТЕЛИ СТАТИСТИКИ

5.1. Сущность обобщающих показателей

Непосредственным результатом статистической сводки является расчет обобщающих показателей. С их помощью осуществляется конечный этап исследований, а именно обобщение или анализ полученных результатов.

Процесс обобщения выражается в расчете соответствующих показателей, характерными чертами этих показателей являются:

- 1) их обобщающий характер, они характеризуют всю совокупность;
- 2) показатели имеют количественно-качественную форму, т.е. за количественными размерами скрывается определенная социально-экономическая сущность;
- 3) показатели должны быть адекватны социально-экономическим процессам.

Обобщающие показатели имеют несколько способов классификации:

По видам различают:

- 1) абсолютные величины;
- 2) относительные величины;
- 3) средние величины.
- 4) индексы.

По охвату совокупности: обобщающие показатели можно квалифицировать как индивидуальные и сводные.

Индивидуальные показатели – признаки, характеризующие отдельные единицы совокупности.

Сводные – признаки, которые характеризуют показатели в целом.

По времени исследования – обобщающие показатели можно подразделить на моментные и интервальные.

Моментные показатели характеризуют состояние явления на некоторый момент времени. Например: численность работающих на начало года, число единиц производственного оборудования на конец года и т. д.

Интервальными показателями называются признаки, характеризующие статистическую совокупность за некоторый период или интервал времени. Например: производство продукции в области за год; затраты материальных ресурсов и т. д.

По видам единиц измерения выделяют:

- 1) натуральные (тонны, км, м) – выражают физические свойства объектов;
- 2) условно-натуральные показатели, применяемые для расширения границ натурального метода (условный вагон, 40-процентное мыло, 7000 калорий топлива);
- 3) стоимостные показатели, получаемые в денежной оценке (обычно валюты собственного государства);

4) трудовые показатели, показатели продукции, получаемой путем умножения количества продукции на нормативную трудоемкость (трудоемкость – затраты труда, рабочего времени на производство единицы продукции):

$$T = \sum_i q_i \cdot t_{ni}, \quad (5.1)$$

где q_i – количество продукции или объем работы конкретного вида,

t_{ni} – нормативная трудоемкость (чел-час, чел-день) конкретного i -го вида продукции или работы;

5) безразмерные показатели. Обычно используются для сравнения одноименных показателей.

Особое место в статистике занимают расчетно-аналитические показатели, имеющие весьма сложную методологию исчисления (совокупный общественный продукт, валовой национальный продукт (ВНП)).

5.2. Абсолютные величины статистики

Абсолютные величины – это показатели, характеризующие количественные размеры изучаемого признака. Они могут быть представлены:

Во-первых: в виде *количественного показателя*, выражающего численность совокупности и определяемого методом прямого счета. Этот признак характеризуется числом единиц, входящих в совокупность. Например: численность населения страны, региона; количество единиц произведенной продукции и т. д.

Во-вторых: абсолютные величины могут быть представлены в виде *объемного показателя*, который характеризует суммарное свойство статистической совокупности и определяется путем суммирования значений признаков отдельных единиц. Пример: фонд зарплаты работников предприятия; материальные затраты на производство; валовой сбор овощей, зерна и т. д.

5.3. Относительные величины

С помощью абсолютных величин можно выполнить лишь начальные элементы анализа.

Целям анализа, сущность которых заключается в сопоставлении, сравнении знаний признаков, служат относительные величины, вычисляемые по схеме:

$$ОВ(\text{относительная величина}) = \frac{\text{показатель 1}}{\text{показатель 2}}. \quad (5.2)$$

Всякая относительная величина – есть соотношение: сравниваемая величина (показатель 1) / база сравнения (показатель 2).

Относительные величины имеют следующие единицы измерения: выражаются в коэффициентах (целые и десятые), $n/1$, $n/10$; в процентах (%), $n/100$;

в промилле (тысячные части числа), (‰), $n/1000$;
в процедимилле (десятитысячная часть числа), (‰‰), $n/10000$.

5.4. Виды относительных величин

Относительная величина динамики (ОВД) представляет собой процесс развития явления во времени. Относительные величины динамики разделяются на базисные и цепные.

Имеем значения динамического ряда: $Y_0, Y_1, Y_2, Y_3, \dots, Y_n$. Тогда базисные ОВД вычисляются путем отношения последовательно взятых уровней динамического ряда к одному и тому же начальному:

$$ОВД_{\text{баз}} = \frac{Y_i}{Y_0} 100\%, \quad (5.3)$$

где $i = 1, 2, 3, \dots, n$.

Цепные ОВД вычисляются отношением последующих уровней динамического ряда к предыдущему:

$$ОВД_{\text{цеп}} = \frac{Y_i}{Y_{i-1}} 100\%. \quad (5.4)$$

Относительная величина планового задания (ОВПЗ) представляет собой план, выраженный в процентах к фактическому уровню базисного периода:

$$ОВПЗ = \frac{Y_{\text{план}}}{Y_0} 100\%. \quad (5.5)$$

Пример: в 2009 году завод выпустил 400 станков; на 2010 год планировал 300 станков:

$$ОВПЗ = \frac{300}{400} 100\% = 75\%.$$

Относительная величина планового задания характеризует степень напряженности поставленного плана (намерений), а вот о выполнении или невыполнении его свидетельствует показатель: *относительная величина выполнения плана (ОВВП)*.

Относительная величина выполнения плана (ОВВП) отражает степень выполнения планового задания и представляет отношение фактического выполнения к установленному плановому заданию.

ОВВП вычисляется различными способами в зависимости от того, в какой форме представлено плановое задание.

Плановое задание может быть представлено в виде:

- 1) абсолютной величины;
- 2) % роста или снижения показателя.

Если плановое задание выражено абсолютной величиной, то ОВВП вычисляется:

$$ОВВП = \frac{Y_1}{Y_{\text{план}}} 100\%,$$

где Y_1 – фактическая величина отчетного периода,

$Y_{\text{план}}$ – плановая величина.

Пример, допустим, что плановый выпуск автомобилей Волжского автомобильного завода (АвтоВаз) составляет $Y_{\text{план}} = 800$ тыс. автомобилей в год. Коллектив фактически произвел 816 тыс. шт. Тогда относительная величина выполнения плана составит:

$$ОВВП = \frac{816}{800} 100\% = 102\% \quad (\text{план перевыполнен на } 2\%).$$

Если плановое задание выражено в виде % роста или снижения показателей, то

$$ОВВП = \frac{100\% + K_{\text{факт.роста}}}{100\% + K_{\text{план.задан}}} 100\% \quad (5.6)$$

Интерпретация результатов вычисления осуществляется различными способами в зависимости от того, планируется рост или снижение показателей.

Примеры:

1. По плану рост производительности труда должен был составить +3,4%. Фактически рост производительности труда составил +5,7%.

$$ОВВП = \frac{100 + 5,7}{100 + 3,4} 100\% = 102,4\%.$$

План по росту производительности труда перевыполнен на 2,4%.

2. Планом намечалось снизить материалоемкость изделия на -0,6%, а фактически снизилась на -0,2%.

$$ОВВП = \frac{100 - 0,2}{100 - 0,6} 100\% = \frac{99,8}{99,4} 100\% = 100,3\%.$$

План по снижению себестоимости недовыполнен на 0,3%.

3. Планом предполагалось снизить себестоимость на -2,1%, а фактически снизилась на -3,9%.

$$ОВВП = \frac{100 - 3,9}{100 - 2,1} 100\% = 98,1\%.$$

План по снижению себестоимости перевыполнен на 1,9%.

Нетрудно заметить, что три последние величины: относительная величина в динамике, относительная величина выполнения плана и относительная величина планового задания, связаны следующей зависимостью:

$$ОВД = ОВВП \cdot ОВПЗ. \quad (5.7)$$

Этой зависимостью можно пользоваться:

1) как одним из методов расчета относительных величин выполнения плана;

2) для контроля за расчетом всех трех величин.

Относительная величина структуры (ОВС)

Представляет собой показатель, служащий для измерения и анализа структуры статистических совокупностей.

ОВС вычисляются, как отношение части совокупности, обладающей определенным свойством, ко всей совокупности:

$$ОВС = \frac{V_i}{V_{\Sigma}} 100\%, \quad (5.8)$$

где V_i – исследуемая часть совокупности, V_{Σ} – вся совокупность.

Например: удельный вес городского населения:

удельный вес мужчин – 48,0%;

удельный вес женщин – 52,0%.

Относительная величина координации (ОВК)

ОВК представляет собой аналитический показатель, который обеспечивает сопоставление двух частей одного целого:

$$ОВК = V_1 / V_2. \quad (5.9)$$

Пример: соотношение численности рабочих и ИТР:

- в развивающихся странах на 1 менеджера 6-7 рабочих;

- в развитых странах на 1 менеджера 2-3 рабочих.

Пример: известна статистика рождаемости мальчиков: 106 мальчиков на 100 девочек.

Относительные величины структуры и координации во многом имеют сходное значение, однако ОВС применяется чаще, т.к. при ее использовании может быть несколько исследуемых частей, а в ОVK по определению только две.

Относительная величина сравнения (ОВСр)

ОВСр вычисляются путем сопоставления значений одноименных показателей, относящихся к различным объектам или территориям, взятым на один и тот же момент (или за один интервал) времени. Обычно исчисляются в %.

Т а б л и ц а 5.1. *Добыча нефти в % СССР и США)*

Годы	1940	1950	1960	1970	1982
% добычи	17	14	42	74	144

Глава 6. СРЕДНИЕ ВЕЛИЧИНЫ

6.1. Понятие о средних величинах

Средние величины – одна из разновидностей обобщающих показателей.

Средние величины – это обобщающие характеристики статистических совокупностей по количественному варьирующему признаку. Средние величины выражают значение определенного признака в расчете на абстрактную единицу совокупности.

Средние величины связаны прежде всего с двумя законами: с внутренним законом развития явления; с законом больших чисел.

1. Связь с внутренним законом состоит в том, что не средние величины регулируют уровни явлений, а некоторый внутренний закон развития. Средние величины способствуют выявлению, установлению этого закона.

2. Связь с законом больших чисел состоит в том, что по мере увеличения совокупности в средних величинах все в более полной мере происходит взаимопогашение случайных отклонений от закономерного уровня.

Важнейшим условием вычисления средних величин является однородность статистической совокупности.

В неоднородных совокупностях метод средних величин необходимо сочетать с методом статистических группировок и расчет общих средних надо сочетать с расчетом групповых средних.

6.2. Виды средних величин

Средние величины – это само по себе понятие математическое. Они представляют собой корень следующего выражения:

$$f(X_1, X_2, X_3, \dots, X_n) = f(\bar{X}, \bar{X}, \dots, \bar{X}). \quad (6.1)$$

n раз

Средней величиной данного ряда будет некоторое число $\bar{X}$, подстановка которого в уравнение n раз не изменит результата вычислений.

Все виды средних можно разделить на две группы: простые и взвешенные.

Простые применяются в том случае, когда в совокупности варианты встречаются только по одному разу. В тех случаях, когда варианты встречаются несколько раз и присутствуют частоты (m), используются формулы средние взвешенные.

6.2.1. Виды простых средних

Простые средние можно представить в виде формулы степенной средней:

$$\bar{X} = \sqrt[n]{\sum x_i^z}, \quad (6.2)$$

где x_i – некоторое конкретное значение признака,

z – показатель степени средней ($-1, 0, 1, 2, \dots$),

n – число единиц совокупности.

В зависимости от показателя степени z получим различные виды средних:

1. При $z = -1$ средняя гармоническая $\bar{X}_h$:

$$\bar{X}_h = \frac{n}{\sum \frac{1}{x_i}}. \quad (6.3)$$

Данная характеристика чаще всего применяется при хронологических исследованиях и оценках временных интервалов.

2. При $z = 0$ – *средняя геометрическая* $\bar{X}_{\text{геом}}$:

$$\bar{X}_{\text{геом}} = \sqrt[n]{\prod_i x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} , \quad (6.4)$$

где Π – произведение.

Данная характеристика является эквивалентной мерой единиц совокупности, имеющих несколько исследуемых параметров.

3. При $z = +1$ – *средняя арифметическая* $\bar{X}_a$:

$$\bar{X}_a = \frac{\sum_i x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n} . \quad (6.5)$$

Наиболее распространенная характеристика, применяется в расчетах показателей, характеризующих физические объекты.

4. При $z = +2$ – *средняя квадратическая* $\bar{X}_q$:

$$\bar{X}_q = \sqrt{\frac{\sum_i x_i^2}{n}} . \quad (6.6)$$

Данная величина представляет собой характеристику разброса данных. С помощью данной характеристики можно оценить степень рассеяния случайной величины относительно некоторого центра – средней арифметической.

5. При $z = +3$ получим *среднюю кубическую* $\bar{X}_k$:

$$\bar{X}_k = \sqrt[3]{\frac{\sum_i x_i^3}{n}} . \quad (6.7)$$

Данная характеристика используется как эквивалентная характеристика объемов фигур.

6. При $z = +4$ – *средняя биквадратическая* $\bar{X}_b$.

Средняя биквадратическая применяется в теории колебаний для расчета тонов и полутонов колебаний.

Между средними величинами существует следующее соотношение, которое описывает *правило мажорантности*: для одних и тех же статистических данных чем больше показатель степени Z , тем больше и сама средняя величина.

Данное правило можно представить в следующем виде:

$$\bar{X} : \bar{X}_h < \bar{X}_{\text{геом}} < \bar{X}_a < \bar{X}_q < \bar{X}_{\text{куб}} < \dots$$

$$Z : -1 < 0 < +1 < 2 < 3 < \dots$$

Все выше рассмотренные случаи относились к простым средним.

6.2.2. Виды средних взвешенных

Для описания *средних взвешенных* величин используется степенная зависимость следующего вида:

$$\bar{X} = \sqrt[z]{\frac{\sum x_i^z \cdot m_i}{\sum m_i}}, \quad (6.8)$$

где x_i – некоторое конкретное значение признака,

m_i – частота варианта x_i ,

z – показатель степени средней ($-1, 0, 1, 2, \dots$),

В зависимости от показателя степени z получим различные виды средних взвешенных.

Приведем наиболее часто употребляемые:

1. При $z = -1$ получим *среднюю взвешенную гармоническую* $\bar{X}_h$:

$$\bar{X}_h = \frac{\sum m_i}{\sum \frac{m_i}{x_i}}. \quad (6.9)$$

2. При $z = +1$ получим *среднюю взвешенную арифметическую* $\bar{X}_a$:

$$\bar{X}_a = \frac{\sum x_i \cdot m_i}{\sum m_i} = \frac{x_1 \cdot m_1 + x_2 \cdot m_2 + \dots + x_n \cdot m_n}{\sum m_i}. \quad (6.10)$$

3. При $z = +2$ – *среднюю взвешенную квадратическую* $\bar{X}_q$:

$$\bar{X}_q = \sqrt{\frac{\sum x_i^2 \cdot m_i}{\sum m_i}}. \quad (6.11)$$

6.3. Техника вычисления среднеарифметической величины и ее свойства

Из всех видов средних величин наиболее распространенной является среднеарифметическая величина. Вычисление средней арифметической величины различается в зависимости от того, в каком виде предоставлены исходные данные.

Так, в случае дискретного распределения величин расчет среднего взвешенного арифметического проводится непосредственно по формуле (6.10). Здесь в качестве вариантов (x_i) применяются значения дискретного распределения (см. данные второго столбца табл. 6.1), а в качестве частот, а точнее частостей (m_i), используются данные третьего столбца.

Т а б л и ц а 6.1. Данные о распределении работников фирмы «Тойота» по числу рационализаторских предложений

№ п/п	Число предложений	Численность в %
1	2	3
1	0	32
2	1	34
3	2	17
4	3	8
5	4	5
6	5 и более	4
	Всего	100%

Данные табл. 6.1 рассчитываются по формулам (6.10):

$$\bar{X}_a = \frac{\sum_i x_i \cdot m_i}{\sum_i m_i} = \frac{0 \cdot 32 + 1 \cdot 34 + 2 \cdot 17 + 3 \cdot 8 + 4 \cdot 5 + 5 \cdot 4}{100} = 1,32.$$

Ответ: 1,32 рациональных предложений приходится в среднем на одного работника

Аналогично осуществляется расчет средней величины в интервальных рядах, но только с той разницей, что в качестве вариантов принимаются значения середины или центра интервала, который вычисляется по следующей формуле:

$$\text{середина интервала} = \frac{\text{нижняя граница интервала} + \text{верхняя граница интервала}}{2}$$

До проведения расчетов нужно помнить следующие правила:

А) В тех случаях, когда первый и последний интервалы имеют открытые границы, при расчете середины этих интервалов предполагается, что величина первого интервала равна величине второго интервала, а величина последнего интервала равна величине предпоследнего.

Б) Если мы имеем дело с прерывным признаком, то при расчете средней интервала нужно в значении нижней границы каждого интервала вычесть 1 или 0,1 (в зависимости от вида прерывного признака).

Т а б л и ц а 6.2. Расчет среднеарифметической величины потерь рабочего времени

Потери времени, мин	Непрерывная вариация	Середина интервала	Численность работников
1	2	3	4
Менее 5	0-5	2.5	22
6-10	5-10	7.5	16
11-15	10-15	12.5	12
16-20	15-20	17.5	7
21 и более	20-25	22.5	3
Итого			60

Первоначально ряд с прерывной вариацией признака преобразовали в ряд с непрерывной вариацией, а затем рассчитали центры интервалов, с помощью которых по формуле (6.10) нашли среднюю арифметическую величину:

$$\bar{X}_a = \frac{\sum_i x_i \cdot m_i}{\sum_i m_i} = \frac{2,5 \cdot 22 + 7,5 \cdot 16 + 12,5 \cdot 12 + 17,5 \cdot 7 + 22,5 \cdot 3}{60} = 8,6 \text{ мин.}$$

Ответ: в среднем на одного рабочего приходится 8,6 мин потерь рабочего времени.

6.4. Свойства средней арифметической

1. Средняя арифметическая ряда, составленного из постоянной величины, равна этой постоянной:

$$\bar{X}_c = C,$$

где $C = \text{const}$.

Данное свойство вытекает из определения средней, если все значения равны некоторой постоянной, то средняя не может отличаться от этой постоянной.

2. Сумма отклонений вариантов от их средней арифметической равна нулю:

$$\sum_i (x_i - \bar{X}) = 0.$$

Данное свойство вытекает из закона «больших чисел», т.к. при достаточно большом количестве единиц совокупности будет происходить полное взаимопоглощение случайных отклонений от среднего значения, т.е. всегда на некоторое положительное отклонение от средней найдется такое же, но отрицательное отклонение.

3. Сумма квадратов отклонений вариантов от их среднеарифметической есть число наименьшее для данного ряда:

$$\sum_i (x_i - \bar{X})^2 = \min.$$

Данное свойство широко применяется в практике. Например в методе «наименьших квадратов», который используется для определения неизвестных параметров уравнений.

4. Если все варианты уменьшить на какое-нибудь число A , то среднеарифметическая уменьшится на это же число:

$$\frac{\sum_i (x - A)_i}{n} = \bar{X} - A.$$

5. Если все варианты разделить на какое-либо число d , то и среднеарифметическая уменьшится в d раз:

$$\frac{\sum_i \left(\frac{x_i}{d} \right)}{n} = \frac{\bar{X}}{d}.$$

6. Если все частоты разделим на какое-либо число, то средняя арифметическая не изменится.

Следствие 1. Абсолютные значения частот можно заменить их процентным выражением.

Следствие 2. Если все частоты равны между собой, то среднюю арифметическую можно вычислить по простой формуле.

Пользуясь свойствами средней арифметической, ее расчет можно осуществлять упрощенно по способу "моментов".

6.5. Расчет средней арифметической величины по правилу «моментов»

На основании вышеприведенных свойств средней арифметической величины сформирована формула вычисления средней арифметической:

$$\bar{X} = \frac{\sum_i \left(\frac{x_i - A}{d} \right) \cdot m_i}{\sum_i m_i} d + A = m' \cdot d + A, \quad (6.12)$$

где A – середина одного из центральных интервалов распределения;

x_i – конкретные значения, варианты;

m_i – частоты, соответствующие вариантам x_i ;

d – величина интервала;

m' – момент первого порядка, т.е. среднеарифметическая из значений $\left(\frac{x_i - A}{d} \right)$.

В табл. 6.3 приведены данные потерь рабочего времени в течение дня в механическом цеху одного из предприятий.

Т а б л и ц а 6.3. *Распределение потерь рабочего времени*

Потери времени, мин	Непрерывная вариация	Средина интервала	Численность работников	Центрированные интервалы	Нормированные интервалы	
		x_i	m_i			
Менее 5	0-5	2.5	22	-10	-2	-44
6-10	5-10	7.5	16	-5	-1	-16
11-15	10-15	12.5	12	0	0	0
16-20	15-20	17.5	7	5	1	7
21 и более	20-25	22.5	3	10	2	6
Итого			60			-47

$$\bar{X} = \frac{\sum_i \left(\frac{x_i - A}{d} \right) \cdot m_i}{\sum_i m_i} \cdot d + A = m^1 \cdot d + A = \frac{-47}{60} \cdot 5 + 12,5 = 8,6 \text{ мин.},$$

где A – середина одного из центральных интервалов, $A = 12,5$;
 d – наибольшая кратная, чаще всего величина интервала, $d = 5$.

Независимо от метода расчета, результат будет идентичным (см. результаты расчета табл. 6.2).

6.6. Вычисление среднего уровня вариационного ряда, составленного из относительных величин

В тех случаях, когда вариационный ряд из качественных или относительных показателей, среднюю величину этого ряда можно вычислять тремя тождественными способами в зависимости от исходных данных:

- 1) по логической формуле;
- 2) по среднеарифметической;
- 3) по средней гармонической

Рассмотрим использование этих способов на примере данных табл. 6.4.

Т а б л и ц а 6.4. *Данные по выработке трех рабочих, занятых изготовлением одной и той же продукции*

Рабочие	Затраты времени на одну деталь, мин	Количество деталей	Отработанное время, мин
	X_i	m_i	M_i
1	20	26	520
2	30	14	420
3	26	18	468
Итого		58	1408

Допустим, что мы располагаем данными об отработанном времени и количестве деталей.

1. Считаем средние затраты рабочего времени на 1 деталь:

$$\text{средние затраты на 1 деталь} = \frac{\text{отработанное время}}{\text{количество деталей}} = \frac{\sum M}{\sum m_i} = \frac{1408}{58} = 24,3 \text{ мин.}$$

2. Допустим, что мы располагаем данными о затратах времени на 1 деталь и об отработанном времени, расчет проводится по формуле средней гармонической:

$$\bar{X}_h = \frac{\sum M_i}{\sum \frac{M_i}{x_i}} = \frac{520 + 420 + 468}{\frac{520}{26} + \frac{420}{14} + \frac{468}{18}} = 24,3 \text{ мин.}$$

3. Если располагаем данными о затратах времени на 1 деталь и количестве деталей, расчет проводится по средней арифметической:

$$\bar{X} = \frac{\sum_i x_i \cdot m_i}{\sum_i m_i} = \frac{20 \cdot 26 + 30 \cdot 14 + 26 \cdot 18}{58} = 24,3 \text{ мин.}$$

6.7. Основное правило выбора вида средних величин

Средняя должна быть взята так, чтобы порядок вычисления соответствовал логической формуле.

Для того чтобы активировать процесс расчета (выбор средней), применяются два формальных правила о выборе вида средней:

1. а) Если в качестве весов (частот) вариационного ряда даны значения количественного показателя, то расчет средней следует осуществлять по средней арифметической.

б) Если же даны значения объемного признака, то расчет следует осуществлять по средней гармонической.

2. а) Если в качестве частот даны значения признака, находящегося в числителе логической формулы, то расчет выполняется по средней гармонической.

б) Если в качестве частот даны значения признака, находящегося в знаменателе логической формулы, то расчет следует выполнять по средней арифметической.

6.8. Мода и медиана

Мода – это вариант вариационного ряда, имеющий наибольшую частоту.

В дискретном ряду мода устанавливается по максимальной частоте.

В интервальном ряду первоначально устанавливается модальный интервал (интервал, имеющий максимальную частоту), а затем в рамках модального интервала находится конкретное значение моды по следующей формуле:

$$M_0 = X_0 + d \frac{(m_2 - m_1)}{(m_2 - m_1) + (m_2 - m_3)}, \quad (6.13)$$

где X_0 – нижняя граница модального интервала;

d – величина интервала;

m_1 – частота интервала, предшествующего модальному интервалу;

m_2 – частота модального интервала;

m_3 – частота интервала, следующего за модальным интервалом.

Расчет модального значения рассмотрим на примере, изложенным в табл. 6.5.

Т а б л и ц а 6.5. *Данные о возрасте вступающих в брак мужчин*

№ п/п	Группы по возрасту	Число вступивших в брак, тыс. чел.
1	Моложе 20	127,3
2	20-25	1232,6
3	25-30	423,1
4	30-35	248,7
...		
10	60 и старше	80,2
	Всего	2365,3

Первоначально найдем модальный интервал – это второй интервал (выделен цветом), у него наибольшая частота (частота равна 1232,6). Затем вычисляем значение моды:

$$M_0 = 20 + 5 \cdot \frac{1232,6 - 127,3}{(1232,6 - 127,3) + (1232,6 - 423,1)} = 22,9 \text{ лет.}$$

Вывод: на основании обработанных данных чаще всего вступают в брак мужчины в возрасте 22,9 лет.

Медиана – вариант вариационного ряда, расположенный в его середине, который делит вариационный ряд на две равные части по количеству единиц совокупности. Первая часть имеет значение меньше медианного значения, вторая часть – больше медианного значения.

В дискретном ряду медиана находится непосредственно путем пересчета количества единиц совокупности и выделения срединного значения. Например, если обследуется сто человек, нас интересует 50-й порядковый номер, и то значение, которое будет у пятидесятого номера, и будет медианным значением.

В интервальном ряду находится медианный интервал (интервал, в котором находится срединное значение по количеству единиц в исследуемой совокупности). Для нахождения медианного интервала желательно заполнить столбец с порядковыми номерами всех единиц совокупности, тогда легче определить медианный интервал. Затем вычисляется конкретное значение медианы по формуле

$$Me = X_0 + d \cdot \frac{\sum m_i - \int m_{Me-1}}{m_{Me}}, \quad (6.14)$$

где X_0 – нижняя граница медианного интервала;

d – величина интервала;

$$\frac{\sum m_i}{2} - \text{полусумма частот ряда};$$

$\int m_{Me-1}$ – сумма накопленных частот интервалов, предшествующих медианному интервалу;

m_{Me} – частота медианного интервала.

Для рассматриваемого примера оценки возраста вступления в брак мужчин (см. табл. 6.5) преобразуем таблицу к следующему виду (табл. 6.6) – добавим столбец, в котором будут записаны порядковые номера соответствующих обследуемых мужчин.

Т а б л и ц а 6.6. *Преобразованные данные о возрасте вступающих в брак мужчин*

№ п/п	Группы по возрасту	Число вступивших в брак, тыс. чел.	Порядковые номера
1	Моложе 20	127.3	1-127300
2	20-25	1232.6	127301-1359900
3	25-30	423.1	1359901-178300
4	30-35	248.7	178301-2031700
...			
10	60 и старше	80.2	2285101-2365300
	Всего	2365.3	

Для данного примера медианный интервал будет вторым (совпадение с модальным интервалом случайно), т.к. срединное значение числа мужчин $2365300/2 = 1182650$ находится в интервале порядковых номеров от 127301 до 1359900.

Для данного интервала находим по формуле (6.14) значение медианы:

$$M_e = 20 + 5 \cdot \frac{\frac{2365.2}{2} - 1127.3}{1232.6} = 24.3 \text{ лет.}$$

Важен вывод: половина мужчин вступает в брак моложе 24,3 лет, вторая половина – старше 24,3 лет.

В статистике используются также понятия: *процентиль* и *квартиль*.

Процентиль характеризует такое значение, ниже которого попадает заданное количество (в процентном выражении) единиц совокупности. Например, 10-й процентиль по заработной плате сотрудников характеризует такое значение заработной платы, ниже которого 10% сотрудников имеет зарплату.

Квартиль – подразумевает разделение вариационного ряда на четыре равных части по 25%. Термин был впервые использован Галтоном в 1882 году. Чаще всего с термином *квартиль* связывают понятия нижней, первой (25-я процентиль) квартили и верхней квартили (это 75-я процентиль). Вторая квартиль – это медиана или 50-й процентиль.

Глава 7. ПОКАЗАТЕЛИ ВАРИАЦИИ

7.1. Система показателей вариации

Характерным свойством экономических или общественных явлений является их способность варьировать, т. е. изменять свои количественные значения под влиянием множества факторов.

Вариация является выражением взаимосвязи философских категорий: случайности и необходимости.

Изучение вариации имеет практическое значение: во-первых, с общенаучных позиций мы определяем степень однородности совокупности; во-вторых, уровень вариации характеризует величину производственных резервов или использованных возможностей.

Для измерения уровня вариации применяется следующая система критериев или показателей:

1. *Размах вариации* P_v , т.е. разность между наибольшим и наименьшим значением признака

$$P_v = X_{\max} - X_{\min}. \quad (7.1)$$

Например, размах вариации по выставляемым оценкам студентам составляет

$$P_v = 5 - 2 = 3 \text{ балла.}$$

2. *Среднее линейное отклонение* представляет собой среднюю величину из абсолютных отклонений вариантов от их средней арифметической:

$$\bar{l} = \frac{\sum_i |x_i - \bar{X}| \cdot m_i}{\sum m_i}, \quad (7.2)$$

где x_i – конкретное значение признака, $\bar{X}$ – среднее арифметическое, m – частота.

Вычисление характеристик рассмотрим на примере (табл. 7.1).

Т а б л и ц а 7.1. *Данные распределения студентов по успеваемости*

Балл успеваемости	Число студентов	$ x_i - \bar{X} $	$ x_i - \bar{X} \cdot m_i$
x_i	m_i	$\bar{X} = 3,6$	
2	4	1.6	6.4
3	18	0.6	10.8
4	22	0.4	8.8
5	6	1.4	8.4
Итого	50		34.4

$$\bar{l} = \frac{\sum_i |x_i - \bar{X}| \cdot m_i}{\sum m_i} = \frac{34.4}{50} = 0,688 \text{ балла.}$$

Это означает, что успеваемость каждого отдельно взятого студента отличается от средней успеваемости в среднем на 0,6868 балла.

На практике среднее линейное отклонение применяется крайне редко, т.к. во многих случаях оно не отражает реальную степень рассеяния исследуемого параметра.

3). *Дисперсия* как характеристика разброса является более объективным показателем, который представляет собой среднюю величину из квадратов отклонений вариантов от их средней арифметической:

$$D = \frac{\sum_i (x_i - \bar{X})^2 \cdot m_i}{\sum_i m_i} . \quad (7.3)$$

Т а б л и ц а 7.2. *Результаты промежуточных вычислений для расчета дисперсии*

Балл успеваемости	Число студентов	$ x_i - \bar{X} $	$ x_i - \bar{X} \cdot m_i$	$(x_i - \bar{X})^2$	$(x_i - \bar{X})^2 \cdot m_i$
x_i	m_i	$\bar{X}=3,6$			
2	4	1,6	6,4	2,56	10,24
3	18	0,6	10,8	0,36	6,48
4	22	0,4	8,8	0,16	3,52
5	6	1,4	8,4	1,96	11,76
Итого	50		34,4		32

Для рассматриваемого примера найдем значение дисперсии:

$$D = \frac{32,00}{50} = 0,64 .$$

Расчет по приведенным формулам является достаточно трудоемким процессом. Часто используется упрощенный способ определения дисперсии:

$$D = \overline{X^2} - (\bar{X})^2 . \quad (7.4)$$

Данная формула представляет разность среднего значения квадрата вариантов и квадрата среднего значения вариантов.

Однако дисперсия имеет следующее неудобство для ее понимания, размерность получаемой характеристики представляется в квадрате, т.е. например, при расчете разброса заработной платы сотрудников получим значение в квадратных рублях и т.п. Поэтому для более простого восприятия используется характеристика – *среднее квадратическое отклонение*.

4. *Среднее квадратическое отклонение* представляет собой корень квадратный из дисперсии:

$$\sigma = \sqrt{D} . \quad (7.5)$$

Применение среднеквадратического отклонения на практике предпочтительнее, т.к. данный показатель выражается в тех же единицах, что и заданные значения.

Для нашего примера данный показатель характеризует, как успеваемость студентов отличается от среднего значения $\sigma = \sqrt{D} = \sqrt{0,64} = 0,8$ (балла).

В ряде случаев удобно использовать безразмерные характеристики вариации, например, когда сравниваются изменения различных признаков.

Для данных действий применяется показатель – *коэффициент вариации*, который представляет собой процентное отношение среднего квадратического отклонения к среднему арифметическому:

$$K_v = \frac{\sigma}{\bar{X}} 100\% . \quad (7.6)$$

Для нашего примера коэффициент вариации составляет 22,2%.

Коэффициент вариации необходим для решения следующих задач:

1. Для сравнения уровня вариации в различных вариационных рядах, имеющих различные единицы измерения.

2. Для оценки степени однородности совокупности. Критерием служит критический коэффициент вариации $K_{Vc} = 33\%$.

Если вариация меньше 33%, то исследуемая совокупность считается однородной, устойчивой, регулярной по данному признаку. Если вариация превышает 33%, то совокупность считается неоднородной, неустойчивой, в ней протекают какие-то процессы, присутствуют тенденции.

7.2. Свойства дисперсии

При вычислении дисперсии необходимо знать и использовать некоторые ее свойства:

1. Дисперсия постоянной величины равна 0:

$$Dc = 0,$$

где $c = \text{const}$.

Действительно, если все используемые значения признака равны между собой, то разброс равен нулю.

2. Если все варианты вариационного ряда уменьшить на одну и ту же величину, дисперсия не изменится:

$$D(x-A) = D(x).$$

3. Если все варианты ряда разделить на одно и то же число d , то дисперсия уменьшится в d^2 раз:

$$D\left(\frac{x}{d}\right) = \frac{1}{d^2} \cdot D(x) .$$

Использование вышеприведенных свойств позволяет существенно упростить вычисление дисперсии. Например, вместо вычисления дисперсии ряда значений: 1001, 1002, 1003, 1004, 1005, можно вычислить значение дисперсии для ряда 1, 2, 3, 4, 5.

7.3. Дисперсия альтернативного признака

Альтернативный признак – это качественный признак, способный принимать два противоположных значения (состояния).

Обозначим первое состояние (рабочее состояние) признака за 1, а соответствующую ему частоту свершения данного события (частость) через p .

Второе состояние (нерабочее состояние) обозначим через 0, а соответствующую частоту (частость) через q .

Возможны только два состояния объекта: полная вероятность (вероятность свершения одного или другого состояния) равна единице; в сумме частоты образуют единицу: $p + q = 1$.

Сформируем математическое представление альтернативного признака. Имеем атрибутивный ряд значений, сведенный в табл. 7.3.

Т а б л и ц а 7.3. *Атрибутивные значения признака*

x_i	m_i
1	p
0	q
Итого	$1 = p + q$

Для вычисления дисперсии первоначально найдем среднюю арифметическую:

$$\bar{X}_a = \frac{\sum_i x_i \cdot m_i}{\sum_i m_i} = \frac{1 \cdot p + 0 \cdot q}{p + q} = p.$$

Вывод: *среднеарифметическое значение альтернативного признака равно частости первого состояния (рабочего состояния) признака.*

Проведем расчет дисперсии, по определению (см. формулу (7.3)) получим:

$$\begin{aligned} D &= \frac{\sum_i (x_i - \bar{X})^2 \cdot m_i}{\sum_i m_i} = \frac{(1-p)^2 \cdot p + (0-p)^2 \cdot q}{p+q} = \frac{q^2 \cdot p + p^2 \cdot q}{1} = \\ &= p \cdot q \cdot (q + p) = p \cdot q. \end{aligned} \quad (7.7)$$

Вывод: *дисперсия значений альтернативного признака равна произведению частостей двух его состояний.*

Нетрудно заметить, что максимально возможное значение дисперсии альтернативного признака равно 0,25. Рассмотрим на примере: рабочее состояние – лампочка горит, нерабочее – не горит. В случае, когда 50% времени лампочка горит и 50% времени не горит, имеем $p = q = 0,5$. Дисперсия составит 0,25. Данное состояние называется равновероятностным состоянием. Во всех остальных случаях значение дисперсии будет меньше. Например: при $p = 0,9$

(рабочее состояние – 90% лампочка горит); нерабочее состояние будет наблюдаться в 10%, т.е. $q = 0,1$. Соответственно дисперсия обоих состояний:
 $p \cdot q = 0,09 < 0,25$.

7.4. Виды дисперсии. Правила сложений дисперсий

В общем случае любая исследуемая совокупность чаще всего является неоднородной и может быть представлена в виде нескольких однородных (типовых) групп. В отношении этой совокупности могут быть выделены и вычислены три вида дисперсии:

1. *Общая дисперсия*, которая может быть вычислена любым из ранее рассмотренных способов. Эта дисперсия рассматривается как результат влияния на изучаемое явление всех определяющих его факторов:

$$D = \frac{\sum_i (x_i - \bar{X})^2 \cdot m_i}{\sum_i m_i}, \quad (7.8)$$

где i – порядковый номер конкретного значения (варианта),

x_i – конкретное значение (вариант),

$\bar{X}$ – общая средняя (для всей совокупности),

m_i – частота i -го варианта во всей совокупности.

2. *Межгрупповая дисперсия*, которая представляет собой средний квадрат отклонений групповых средних от общей средней:

$$\delta^2 = \frac{\sum_j (\bar{x}_j - \bar{X})^2 \cdot m_j}{\sum_j m_j}, \quad (7.9)$$

где j – номер группы ($i = 1, 2, \dots$);

$\bar{x}_j$ – групповые средние в каждой из типических групп;

$\bar{X}$ – межгрупповая средняя;

m_j – численность единиц совокупности в j -й типической группе.

Эта дисперсия рассматривается как результат влияния на изучаемое явление признака, положенного в основу группировки.

3. *Внутригрупповая дисперсия* – дисперсия, определенная для отдельной группы:

$$\sigma_j^2 = \frac{\sum_i (x_i - \bar{x}_j)^2 \cdot m_{ij}}{\sum_i m_{ij}}, \quad (7.10)$$

где x_i – средняя вариантов j -й группы;

m_{ij} – частота повторения i -го варианта в j -й группе;

σ_j^2 – дисперсия i -й группы.

Можно найти среднюю из внутригрупповых дисперсий. Она представляет собой среднюю арифметическую взвешенную из групповых дисперсий, т. е. дисперсий в каждой из типических групп:

$$\bar{\sigma}_j^2 = \frac{\sum_j \sigma_j^2 \cdot m_j}{\sum_j m_j}, \quad (7.11)$$

где σ_j^2 – групповые дисперсии.

Эта дисперсия рассматривается как результат влияния на изучаемое явление всех прочих факторов.

Нетрудно заметить, что между указанными видами дисперсий существует зависимость

$$D = \delta^2 + \bar{\sigma}_j^2. \quad (7.12)$$

Данное уравнение носит название *правило сложения дисперсий*.

Исходя из данного правила видно, что чем больше межгрупповая дисперсия (компонента δ^2), тем в большей степени группировочный признак оказывает влияние на процесс формирования изучаемого явления.

Эту меру влияния можно выразить количественно с помощью показателя *эмпирического коэффициента детерминации*:

$$\eta^2 = \frac{\delta^2}{D}. \quad (7.13)$$

Корень квадратный из данного выражения есть эмпирическое корреляционное отношение:

$$\eta = \sqrt{\frac{\delta^2}{D}}. \quad (7.14)$$

Данная величина всегда положительна и заключена в интервале: $0 \leq \eta \leq 1$.

Причем чем ближе η к 1, тем более полным является влияние группировочного признака.

Рассмотрим вычисление этих дисперсий и покажем справедливость соотношений на примере, приведенном в табл. 7.4. Стояла задача, объективно ли распределение рабочих по присвоенным разрядам (квалификации). Факторным признаком является количество изготовленных деталей за смену.

Т а б л и ц а 7.4. *Производство деталей токарями двух бригад*

Бригада №1 (токари 5 разряда)			Бригада №2 (токари 6 разряда)		
№ п/п	Количество деталей, штук: X_1	X_1^2	№ п/п	Количество деталей, штук: X_2	X_2^2
1	8	64	1	9	81
2	8	64	2	9	81
3	9	81	3	10	100
4	11	121	4	10	100
			5	12	144
			6	13	169
Всего	36	330	Всего	63	675

Рассчитаем групповые средние (средние для бригад токарей 5 и 6-го разрядов).

Групповые средние:

$$\bar{x}_{5p} = \frac{36}{4} = 9(\text{дет.}), \quad \bar{x}_{6p} = \frac{63}{6} = 10.5(\text{дет.})$$

Общая средняя:

$$\bar{X} = \frac{36+63}{10} = 9.9(\text{дет.})$$

Групповые дисперсии:

$$\sigma_{5p}^2 = \overline{x_{5p}^2} - (\bar{x}_{5p})^2 = \frac{330}{4} - 9^2 = 1.5(\text{дет.}^2),$$

$$\sigma_{6p}^2 = \overline{x_{6p}^2} - (\bar{x}_{6p})^2 = \frac{675}{6} - 10.5^2 = 2.25(\text{дет.}^2).$$

Общая дисперсия:

$$D = \overline{X^2} - (\bar{X})^2 = \frac{330+675}{4+6} - 9.9^2 = 2.49(\text{дет.}^2).$$

Средняя арифметическая величина из групповых дисперсий:

$$\bar{\sigma}_j^2 = \frac{\sum_j \sigma_j^2 \cdot m_j}{\sum_j m_j} = \frac{1.5 \cdot 4 + 2.25 \cdot 6}{10} = 1.95(\text{дет.}^2).$$

Межгрупповая дисперсия составит:

$$\delta^2 = \frac{\sum_j (\bar{x}_j - \bar{X})^2 \cdot m_j}{\sum_j m_j} = \frac{(9.0 - 9.9)^2 + (10.5 - 9.9)^2}{10} = 0.54(\text{дет.}^2).$$

Проверка правила сложения дисперсий подтверждает его выполнение, итак:

$$D = \delta^2 + \bar{\sigma}_j^2 = 1.95 + 0.54 = 2.49 .$$

В данном примере: общая дисперсия (D) отражает суммарное влияние всех возможных факторов: квалификацию токарей, техническое состояние станков, обеспеченность станками и материалами и т. д.

Внутригрупповая дисперсия ($\bar{\sigma}_j^2$) показывает величину вариации иных признаков, кроме различий в квалификационном разряде. Внутри группы (бригады) токари имеют один разряд.

Межгрупповая дисперсия (δ^2) характеризует вариацию групповых средних, обусловленных различиями групп по квалификационному разряду.

Чем больше вклад межгрупповой дисперсии, тем сильнее влияние группового признака (квалификационного признака).

В рассматриваемом примере значение коэффициента детерминации составило:

$$\eta^2 = \frac{\delta^2}{D} = \frac{0.54}{2.49} = 0.217 .$$

В данном конкретном примере это тревожная ситуация, т.к. на назначение разрядов (квалификации) рабочим влияние выработки продукции (производительная программа) оказалось несущественным фактором (только 21,7%). В большей мере (78,3%) сказалось влияние прочих факторов (возможно, техническое состояние станков, благосклонность начальства и пр.).

Глава 8. ДИНАМИЧЕСКИЕ РЯДЫ

8.1. Понятие о рядах динамики. Виды динамических рядов

Явления общественной жизни обладают способностью варьировать не только в пространстве, но и во времени. Изучение изменения явления во времени составляет одну из наиболее важных задач статистики. Статистика решает эту задачу путем построения рядов динамики или динамических рядов.

Динамическим рядом называется ряд числовых значений, характеризующих процесс развития явлений во времени.

Любой динамический ряд включает два основных элемента:

- Уровни ряда динамики отображают количественную оценку (меру) развития во времени изучаемого явления. Они могут выражаться абсолютными, относительными или средними величинами.

- Показатель времени, соответствующий уровню ряда.

В зависимости от характера изучаемого явления уровни рядов динамики подразделяются на *моментные* и *интервальные*.

Моментным называется динамический ряд, уровни которого характеризуют величину явления по состоянию на определенные даты. Примером моментного ряда может служить информация о списочной численности по месяцам работников предприятия. Особенностью моментного ряда динамики является то, что в его уровни могут входить одни и те же единицы изучаемой совокупности, т. е. может возникнуть повторный счет. Посредством моментных рядов динамики изучают товарные запасы, количество оборудования и другие показатели, отражающие состояние изучаемых явлений на отдельные даты времени.

Интервальным называется ряд, уровни которого характеризуют величину изучаемого показателя, полученную за определенные периоды, интервалы времени. Особенностью интервального ряда динамики является то, что каждый его уровень складывается из данных за более короткие интервалы (субпериоды) времени. Например, суммируя по торговой фирме товарооборот за первые три месяца года, получают его объем за первый квартал, а сумма товарооборота четырех кварталов дает объем товарооборота за год и т.д.

Ряды динамики, как правило, представляют в виде таблиц, в подлежащем которых указываются уровни ряда, а в сказуемых – указатель времени, это могут быть годовые уровни, квартальные, месячные данные, декадные и т.д. (табл. 8.1).

Т а б л и ц а 8.1. *Динамика товарооборота региона*

Уровни ряда	2005	2006	2007	2008	2009	2010
Товарооборот, млрд. руб.	142,1	130,9	128,9	132,1	140,9	145,6

При графическом отображении динамического ряда на оси абсцисс строится шкала времени, а на оси ординат – шкала уровней ряда. При этом на оси абсцисс обязательно должен соблюдаться масштаб времени, в противном случае линейный график неправильно отразит характер изменения.

8.2. Правила составления рядов динамики

Динамические ряды в процессе их построения должны соответствовать ряду требований, содержание которых заключается в обеспечении сопоставимости уровней ряда:

- уровни динамических рядов, характеризующие динамику продукции и другие обобщающие показатели процесса, производства, должны быть вычислены в сопоставимых ценах и тождественных единицах измерения;

- уровни рядов динамики должны обладать сопоставимостью во времени, т.е. уровни интервальных рядов должны относиться к равным по величине интервалам времени, а уровни моментных рядов должны относиться к одним и тем же моментам времени. Всякие исключения из этого правила должны оговариваться;

- уровни динамических рядов должны быть сопоставимы по территории, а также обладать вещественной сопоставимостью.

8.3. Аналитические показатели ряда динамики

Ориентируясь непосредственно на уровни рядов динамики, можно сделать лишь самые общие выводы о характере зависимости. Для того чтобы более определенно характеризовать закономерности развития, необходимы надлежащие параметры, которые называются аналитическими показателями. К ним относятся следующие: абсолютный прирост, коэффициент роста или темп роста, коэффициент прироста или темп прироста.

В основе расчета показателей рядов динамики лежит сравнение его уровней. В зависимости от применяемого способа сопоставления показатели динамики могут вычисляться на постоянной и переменной базах сравнения.

При расчете показателей на постоянной базе каждый уровень сравнивается с одним и тем же базисным уровнем. Исчисляемые при этом показатели называются *базисными*.

При расчете показателей на переменной базе каждый последующий уровень ряда сравнивается с предыдущим. Исчисляемые при этом показатели динамики называются *цепными*.

Ниже приведены основные показатели динамики :

Абсолютный прирост является важнейшим статистическим показателем динамики. Он определяется в разностном сопоставлении двух уровней ряда динамики в единицах измерения исходной информации.

Базисный абсолютный прирост ΔV_i^B вычисляется как разность между сравниваемым уровнем V_i и уровнем, принятым за постоянную базу сравнения, V_0 :

$$\Delta V_i^B = V_i - V_0 . \quad (8.1)$$

Цепной абсолютный прирост ΔV_i^C представляет собой разность между сравниваемым уровнем V_i и уровнем, который ему предшествует, V_{i-1} :

$$\Delta V_i^C = V_i - V_{i-1} . \quad (8.2)$$

Коэффициент роста – относительный показатель, получающийся в результате деления двух уровней ряда, показывает, во сколько раз уровень сравниваемого периода V_i больше или меньше некоторого уровня, с которым производится сравнение.

Базисный коэффициент роста вычисляется делением сравниваемого уровня V_i на уровень, принятый за постоянную базу сравнения (базисный уровень), V_0 :

$$K_i^B = \frac{V_i}{V_0} . \quad (8.3)$$

Цепной коэффициент роста вычисляется делением сравниваемого уровня V_i на предыдущий уровень:

$$K_i^C = \frac{V_i}{V_{i-1}} . \quad (8.4)$$

Темп роста представляет собой коэффициент роста, переведенный в проценты:

$$T_p = K_p \cdot 100\%. \quad (8.5)$$

Коэффициент прироста характеризует относительную величину прироста, т.е. его величину по отношению к уровню, принятому за базу сравнения.

Этот показатель может быть вычислен двумя способами:

1-й способ. Как отношение абсолютного прироста и уровня ряда, с которым производится сравнение.

Базисный коэффициент прироста K_{np} вычисляется делением сравниваемого базисного абсолютного прироста U_i на уровень, принятый за постоянную базу сравнения (базисный):

$$K_{np}^B = \frac{\Delta V_i}{V_0}. \quad (8.6)$$

Цепной коэффициент прироста K_{np} – это отношение сравниваемого цепного абсолютного прироста U_i к предыдущему уровню V_i :

$$K_{np}^B = \frac{\Delta V_i}{V_{i-1}}. \quad (8.7)$$

2-й способ. С использованием коэффициента роста.

Коэффициент прироста (базисный или цепной) определяется как разность соответствующего коэффициента роста (базисного или цепного) и единицы:

$$K_{np} = K_p - 1. \quad (8.8)$$

Темп прироста – относительный показатель, показывающий, на сколько процентов один уровень больше (или меньше) другого уровня, принятого за базу сравнения. Темп прироста характеризует абсолютный прирост в относительных величинах.

Формула для вычисления темпа прироста имеет следующий вид:

$$T_{np_i} = T_{p_i} - 100\%. \quad (8.9)$$

Выраженный в процентах темп прироста показывает, на сколько процентов изменился уровень по сравнению с базисным, принятым за 100%.

Распространенным статистическим показателем динамики является абсолютное содержание одного процента прироста, который представляет соотношение величины абсолютного прироста и темпа прироста:

$$\text{Абсолютное содержание 1\% прироста} = \frac{\Delta U}{T_{np}(\%)} 1\%. \quad (8.10)$$

Этот показатель выражает материальное содержание (вес) 1% прироста. Данный статистический показатель наиболее часто применяется в сравнительном анализе.

8.4. Аналитические показатели ускорения (замедления)

Для получения аналитических показателей динамики социально экономических явлений, характеризующих внутренние тенденции развития ряда динамики (развитие с ускорением или замедлением), служат следующие показате-

тели: абсолютное ускорение, прирост темпов роста, относительное ускорение, темп ускорения, прирост темпов ускорения.

Абсолютное ускорение представляет собой разность величин абсолютных приростов:

$$\delta Y_i = \Delta Y_i - \Delta Y_{i-1} = (Y_i - Y_{i-1}) - (Y_{i-1} - Y_{i-2}). \quad (8.11)$$

Прирост темпов роста определяется как разность цепных темпов роста в процентных пунктах (пп.):

$$\Delta Tr_i = Tr_i - Tr_{i-1}. \quad (8.12)$$

Относительное ускорение представляет собой разность величин приростов темпов роста:

$$\delta Tr_i = \Delta Tr_i - \Delta Tr_{i-1} = (Tr_i - Tr_{i-1}) - (Tr_{i-1} - Tr_{i-2}). \quad (8.13)$$

Темп ускорения – определяется как отношение цепных темпов роста динамического ряда:

$$T_{уск_i} = \frac{Tr_i}{Tr_{i-1}} 100\%. \quad (8.14)$$

Прирост темпов ускорения представляет разность двух последовательно взятых темпов ускорения в процентных пунктах:

$$\Delta T_{уск_i} = T_{уск_i} - T_{уск_{i-1}}. \quad (8.15)$$

Для лучшего усвоения изложенного рассмотрим использование цепных аналитических показателей на конкретном примере.

Т а б л и ц а 8.2. *Расчет аналитических показателей динамики и ускорения по одному из экономических регионов*

Годы	Доход региона, млрд. руб.	Абсолютный прирост, млрд. руб.	Темпы роста, %	Прирост темпов роста, %	Абсолютное ускорение, %	Относительное ускорение, %	Темпы ускорения	Прирост темпов ускорения
t	Y_i	$\Delta Y_i^ц$	Tr_i	ΔTr_i	δY_i	δTr_i	$T_{уск_i}$	$\Delta T_{уск_i}$
2000	7,3	-	-	-	-	-	-	-
2001	7,7	0,4	105,5	-	-	-	-	-
2002	8,6	0,9	111,7	6,2	0,5	-0	105,9	-
2003	9,1	0,5	105,8	-5,9	-0,4	-12,1	94,7	-11,2
2004	9,2	0,1	101,1	-4,7	0,4	1,2	95,6	0,9

Если необходимо провести анализ нескольких динамических рядов одновременно, то употребляется динамический показатель – *темп опережения*, который представляет собой отношение темпов роста двух различных параметров за один и тот же промежуток времени:

$$T_{опреж}(X / Y) = \frac{Tr(X)}{Tr(Y)} 100\%. \quad (8.16)$$

8.5. Средние показатели динамического ряда

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

Средний уровень ряда динамики характеризует типическую величину абсолютных уровней. При определении средних уровней рядов динамики следует помнить, что моментные и интервальные ряды обладают различными методологическими свойствами.

Уровни интервальных рядов можно суммировать, получая показатели за более продолжительные периоды времени. Поэтому вычисление среднего уровня интервального ряда производим по формуле среднего арифметического из n уровней ряда:

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{Y_1 + Y_2 + \dots + Y_i + \dots + Y_n}{n}, \quad (8.17)$$

где Y_i – некоторое значение i -го интервала,
 n – количество значений.

Учитывая, что уровни моментных рядов суммировать нельзя, необходимо первоначально выполнить операцию перевода моментного ряда в интервальный.

Рассмотрим случай моментного ряда динамики с равновеликими промежутками времени (интервалами).

Перевод осуществляется путем нахождения полусумм этого ряда:

$$Y_I = \frac{Y_1 + Y_2}{2}, \quad Y_{II} = \frac{Y_2 + Y_3}{2}, \quad Y_j = \frac{Y_i + Y_{i+1}}{2}, \quad (8.18)$$

где Y_j – значение интервального ряда,

Y_i – значения моментного ряда.

Затем для полученных значений интервального ряда найдем среднее значение:

$$\bar{Y} = \frac{\frac{Y_1 + Y_2}{2} + \frac{Y_2 + Y_3}{2} + \dots + \frac{Y_{n-1} + Y_n}{2}}{n-1}.$$

Проведя некоторые преобразования формулы, окончательно получим:

$$\bar{Y} = \frac{\frac{Y_1}{2} + Y_2 + Y_3 + \dots + Y_i + \dots + \frac{Y_n}{2}}{n-1}. \quad (8.19)$$

Полученная характеристика называется *средней хронологической моментного ряда*.

В случае моментного ряда с неравноотстоящими датами *средний уровень* определяется по формуле

$$\bar{Y} = \frac{\sum_i t_i \cdot Y_i}{\sum_i t_i} = \frac{t_1 \cdot Y_1 + t_2 \cdot Y_2 + \dots + t_n \cdot Y_n}{t_1 + t_2 + \dots + t_n}. \quad (8.20)$$

Пример. Для данных табл. 8.3 необходимо найти среднюю численность сотрудников за I квартал.

Т а б л и ц а 8.3. *Численность сотрудников на предприятии*

Даты	01.01.10	01.02.10	01.03.10	01.04.10
Численность, чел	116	109	105	108

Решение проводим по формуле среднего уровня для моментного ряда с равновеликими промежутками времени (8.17):

$$\bar{Y} = \frac{0,5 \cdot 116 + 109 + 105 + 0,5 \cdot 108}{4 - 1} = 109 \text{ чел.}$$

Наряду с вычислениями средних уровней динамических рядов используются и другие средние значения аналитических показателей: *средний абсолютный прирост и средний темп роста*.

Средний абсолютный прирост представляет собой обобщенную характеристику индивидуальных абсолютных приростов ряда динамики и определяется как среднее арифметическое из цепных абсолютных приростов.

Учитывая, что сумма цепных абсолютных приростов равна базисному абсолютному приросту, т.е.

$$(Y_2 - Y_1) + (Y_3 - Y_2) + (Y_4 - Y_3) + \dots + (Y_n - Y_{n-1}) = Y_n - Y_1,$$

средний абсолютный прирост вычисляется по граничным уровням ряда динамики:

$$\Delta \bar{Y} = \frac{Y_n - Y_1}{n - 1}. \quad (8.21)$$

Возвращаясь к примеру, средний абсолютный прирост составляет:

$$\Delta \bar{Y} = \frac{Y_n - Y_1}{n - 1} = 0,475 \text{ млрд. руб.}$$

Данная величина свидетельствует о том, что в среднем национальный доход возрастал в год на сумму 0,475 млрд.руб.

Средний коэффициент (темп) роста – обобщающая характеристика индивидуальных темпов роста ряда динамики и для его определения применяется формула среднего геометрического из цепных коэффициентов роста:

$$\bar{Kp} = \sqrt[n]{Kp_1 \cdot Kp_2 \cdot Kp_3 \cdot \dots \cdot Kp_n}. \quad (8.22)$$

Аналогично находится *средний темп роста*:

$$\bar{Tp} = \sqrt[n]{Tp_1 \cdot Tp_2 \cdot Tp_3 \cdot \dots \cdot Tp_n}. \quad (8.23)$$

Для данных табл. 8.2 средний коэффициент роста составляет:

$$\bar{Kp} = 105,5 \cdot 111,7 \cdot 105,8 \cdot 101,1 = 107\%.$$

В среднем каждый год национальный доход возрастал на 7%.

Глава 9. СПОСОБЫ ОБРАБОТКИ ДИНАМИЧЕСКИХ РЯДОВ

В отдельных случаях вариация уровней динамических рядов может стать значительной и утрачивается возможность сделать какие-либо выводы о закономерностях развития, ориентируясь непосредственно на уровни ряда. В таких случаях для получения сколько-нибудь определенных выводов применяются искусственные приемы обработки рядов динамики:

1. Увеличение интервалов.
2. Способ характеристики ряда средними величинами.
3. Метод скользящей средней.
4. Приведение рядов динамики к общему основанию.
5. Смыкание динамических рядов.

Рассмотрим использование данных приемов на следующем примере (табл. 9.1).

Т а б л и ц а 9.1. *Данные о выпуске продукции по месяцам*

№ п/п	Месяцы	Месячный выпуск, млн. руб.	Кварталы	Квартальный выпуск, млн. руб.
1	Январь	235	I	700
2	Февраль	229		
3	Март	236		
4	Апрель	241	II	719
5	Май	238		
6	Июнь	240		
7	Июль	240	III	727
8	Август	242		
9	Сентябрь	245		
10	Октябрь	242	IV	744
11	Ноябрь	252		
12	Декабрь	250		

9.1. Способ укрупнения интервалов

Способ укрупнения интервалов заключается в том, что данные, относящиеся к каким-либо мелким интервалам времени, заменяются данными, относящимися к более крупным, продолжительным периодам времени. В нашем случае данные о выпуске продукции за месяц заменяем квартальными.

Получили: I кв. – 700 млн. руб.

II кв. – 719 млн. руб.

III кв. – 727 млн. руб.

IV кв. – 744 млн. руб.

Можно сделать вывод: устойчивый рост объема производства из квартала в квартал наглядно характеризует тенденцию увеличения объема производства в течение года.

У данного метода можно выделить два недостатка:

- существенное сокращение исходной статистической информации (в данном примере вместо двенадцати значений осталось четыре);
- принципиально меняются единицы измерения исходной информации (в нашем примере вместо данных за месяц используются данные за квартал).

9.2. Способ характеристики ряда средними величинами

Способ характеристики динамических рядов средними величинами состоит в том, что уровни ряда динамики заменяются их средними величинами, взятыми за несколько промежутков времени.

Заменим месячные данные о выпуске продукции среднемесячными значениями соответствующего квартала:

I кв. – $700 : 3 = 233$ млн. руб.

II кв. – $719 : 3 = 238$ млн. руб.

III кв. – $727 : 3 = 242$ млн. руб.

IV кв. – $744 : 3 = 248$ млн. руб.

Вывод: в течение года происходит увеличение объема производства.

Более отчетливое проявление тенденции здесь связано с тем обстоятельством, что в средних величинах происходит взаимное погашение случайных отклонений от закономерного уровня.

У данного метода существует один недостаток – существенное сокращение исходной информации (в данном примере вместо двенадцати значений осталось четыре).

9.3. Метод скользящей средней

Суть этого способа состоит в том, что последовательно, на протяжении всего динамического ряда фактические значения заменяются расчетными, полученными как средняя арифметическая величина из n уровней.

Здесь n – период скользящей средней ($n = 3, 4, 5, \dots, 12$).

Проведем все действия для случая $n = 3$:

1-е действие.

Берем первые три значения выпуска продукции (данные за январь, февраль, март месяцы), по этим значениям вычисляем среднее арифметическое и относим его ко второму периоду времени (табл. 9.2).

$$\text{Для нашего примера } \frac{235 + 229 + 236}{3} = 233.$$

Таблица 9.2. Фрагмент обработки значений по методу скользящей средней

№ п/п	Месяцы	Месячный выпуск, млн. руб.	Среднее значение выпуска
1	Январь	235	233
2	Февраль	229	
3	Март	236	

2-е действие.

Отбрасываем первый уровень ряда (значение выпуска продукции за январь месяц), добавляем четвертый уровень (за апрель месяц), находим среднее арифметическое и относим его к третьему уровню динамического ряда.

3-е действие.

Отбрасываем второй уровень, добавляем пятый, находим среднее арифметическое и т.д. для всех уровней динамического ряда. Окончательно получим следующий сглаженный ряд (табл. 9.3). На основании сглаженных данных (табл. 9.3) отчетливо проявилась тенденция к увеличению объема производства на протяжении всего года. Данный способ обработки данных не имеет недостатков, присущих первым двум способам. Сокращение расчетных или сглаженных уровней будет незначительно, всего на $n - 1$ меньше, чем в исходном ряду. В нашем примере период скользящей средней был $n = 3$, соответственно число расчетных значений составило:

число исходных данных: $N = 12$;

число сглаженных данных: $N - (n - 1) = 12 - (3 - 1) = 10$.

При данном методе наблюдается потеря информации, но она не так велика, как в двух предыдущих методах.

Есть ограничение при использовании метода скользящей средней: не рекомендуется брать четным период скользящей средней. При четном периоде скользящей средней расчетные значения будут относиться не к конкретным периодам времени, а к промежуткам между ними, придется дополнительно осуществлять операцию центрирования, которая заключается в нахождении центров путем расчета полусумм из каждой пары расчетных значений. В связи с этими сложностями рекомендуется брать нечетным период скользящей средней ($n = 3, 5, 7$).

Т а б л и ц а 9.3. *Сглаженные значения месячного выпуска продукции*

№ п/п	Месяцы	Среднее значение выпуска
1	Январь	-
2	Февраль	233
3	Март	235
4	Апрель	238
5	Май	239
6	Июнь	240
7	Июль	241
8	Август	242
9	Сентябрь	243
10	Октябрь	247
11	Ноябрь	248
12	Декабрь	-

9.4. Приведение динамических рядов к общему основанию

Данный способ является приемом изучения взаимосвязей динамических тенденций у нескольких рядов. Алгоритм данного метода заключается в следующем:

Первый этап. Первоначально в нескольких различных, но взаимосвязанных рядах начальные уровни динамических рядов принимаются за 1 или за 100 %.

Второй этап. Все последующие уровни каждого ряда выражают в процентах или коэффициентах по отношению к начальным уровням соответственно.

Третий этап. В результате параллельного сопоставления сравниваемых рядов определяется характер взаимосвязи.

Приведем пример выпуска средств производства и предметов потребления за период с 2003 по 2009 гг. и сопоставим приведенные данные.

Таблица 9.4. *Соотношение выпуска средств производства и предметов потребления*

Годы	2003	2004	...	2009
Производство средств производства (А)	100,0%	104,0%	...	111,6%
Производство предметов потребления (В)	100,0%	104,1%	...	113,5%
Коэффициент опережения (В) : (А)	1,000	1,010	...	1,017

Из табл. 9.4 видно, что опережающими темпами росло производство предметов потребления.

9.5. Смыкание динамических рядов

В социально-экономической жизни общества иногда возникают коллизии (необычные явления), которые разрывают динамический ряд, делают его 2-е части несопоставимыми.

Например:

- изменение географических границ региона,
- изменение цен и тарифов и др.

Сущность способа смыкания динамических рядов заключается в выполнении следующих процедур:

1. В переходный период (год, когда произошли коллизии) вычисляются два уровня ряда:

- первый уровень вычисляется исходя из прежних условий производства;
- второй уровень – исходя из новых условий.

2. Оба уровня принимаются за 100%.

3. Все уровни пересчитываются в процентные отношения, затем предшествующие уровни вычисляются в процентах к первому уровню переходного периода, а последующие в процентах ко второму уровню.

Данный прием рассмотрим на примере обработки данных по реализации продукции по годам производственного объединения (табл. 9.5).

Т а б л и ц а 9.5. *Выпуск продукции производственного объединения*

Годы	2005	2006	2007*	2008	2009
Реализация	4285	4521	4826	—	-
продукции, тыс. руб.	—	—	3689	3805	4017

Примечание: * в 2007 году из состава производственного объединения выделилось малое предприятие и продукция объединения сократилась, таким образом, данные за 2008 и 2009 годы несопоставимы с уровнями 2005 года.

Выполняем смыкание уровней (см. табл. 9.5), объединяя приведенные значения в единый динамический ряд, используя основное правило данного метода. Получим следующие данные (табл. 9.6), обладающие свойством сопоставимости (в нашем случае в процентах к уровню 2007 года).

Т а б л и ц а 9.6. *Динамический ряд после обработки методом смыкания динамических рядов*

Годы	2005	2006	2007*	2008	2009
Реализация	88,8	93,7	100,0	-	-
продукции, в %	-	-	100,0	103,1	108,9

Из обработанных данных табл. 9.6 наглядно просматривается тенденция: независимо от того, что от предприятия отделилась часть производственных цехов, рост выпуска продукции в период с 2007 по 2009 гг. сохранился.

Глава 10. ПРИЕМЫ ОБРАБОТКИ СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ

10.1. Аналитическое выравнивание

Для более точного описания характера исследуемого процесса или явления необходимо получение математического описания. Чаще всего для этого используется метод аналитического выравнивания. При реализации этого метода вводится предположение, что уровни ряда являются функцией времени:

$$Y = f(t) . \quad (10.1)$$

Простейшим выражением этой функции является уравнение прямой:

$$Y_t = a_0 + a_1 \cdot t , \quad (10.2)$$

где Y_t – выражает расчетное значение Y при некотором значении t .

Параметры этого выражения имеют реальный экономический смысл:

a_0 – некоторое центральное значение ряда,

a_1 – скорость изменения уровней ряда.

Неизвестные параметры a_0 и a_1 этого выражения вычисляются методом наименьших квадратов. Параметры вычисляются при условии, что сумма квадратов отклонений фактических значений Y_i от расчетных Y_t будет наименьшей:

$$\sum_i (Y_i - Y_t)^2 \Rightarrow \min. \quad (10.3)$$

Данный способ реализуется путем исследования этой функции на экстремум. Находятся производные данного выражения (10.3) и приравняются к нулю:

$$\sum_i [2(Y_i - a_0 - a_1 \cdot t_i)] = 0.$$

Преобразовав, получим

$$\sum_i Y_i = \sum_i a_0 + \sum_i a_1 \cdot t_i.$$

Учитывая, что $\sum_i a_0 = n \cdot a_0$, получим

$$\sum_i Y_i = n \cdot a_0 + \sum_i a_1 \cdot t_i = n \cdot a_0 + a_1 \sum_i t_i.$$

Взяв производную по a_1 , получим выражение

$$\sum_i Y_i \cdot t_i = a_0 \cdot \sum_i t_i + a_1 \cdot \sum_i t_i^2.$$

Оба уравнения представим в виде системы

$$\sum_i Y_i = n \cdot a_0 + a_1 \sum_i t_i,$$

$$\sum_i Y_i \cdot t_i = a_0 \sum_i t_i + a_1 \sum_i t_i^2. \quad (10.4)$$

Составляем данную систему по фактическим значениям ряда и решаем ее относительно a_0 и a_1 , в результате получаем искомые параметры:

$$a_0 = \frac{\sum_i Y_i \cdot \sum_i t_i^2 - a_1 \sum_i t_i \cdot \sum_i Y_i \cdot t_i}{n \sum_i t_i^2 - \sum_i t_i \cdot \sum_i t_i},$$

$$a_1 = \frac{n \sum_i Y_i \cdot t_i - \sum_i Y_i \cdot \sum_i t_i}{n \sum_i t_i^2 + \sum_i t_i \cdot \sum_i t_i}. \quad (10.5)$$

Решение системы можно упростить, используя следующую замену: введем такой параметр времени t_i^* , чтобы $\sum_i t_i^* = 0$.

Тогда выражение (10.5) получит следующий вид:

$$a_0 = \frac{\sum_i Y_i}{n},$$

$$a_1 = \frac{\sum_i Y_i \cdot t_i}{\sum_i t_i^2}. \quad (10.6)$$

Применение метода аналитического выравнивания рассмотрим на примере.

Т а б л и ц а 10.1. *Динамика основных производственных фондов предприятия*

№ п/п	Годы	Уровни, млрд.руб.	Скорректированное время	Рабочие параметры формулы (10.6)		Расчетные значения
	t_i	Y_i	t_i^*	$Y_i \cdot t_i^*$	t_i^2	Y_t
1	2002	11,8	-3,5	-41,30	12,25	11,9
2	2003	12,4	-2,5	-31,00	6,25	12,5
3	2004	13,5	-1,5	-20,25	2,25	13,1
4	2005	13,6	-0,5	-6,80	0,25	13,7
5	2006	14,3	0,5	7,15	0,25	14,3
6	2007	14,9	1,5	22,35	2,25	14,9
7	2008	15,6	2,5	39,00	6,25	15,5
8	2009	15,9	3,5	55,65	12,25	16,1
	Итого	112,0		24,80	42,00	

В данной таблице за середину (центр) выбрана середина ряда (2005 год).

Тогда временной параметр 2002г. равен $t_{2002}^* = -3,5$,

временной параметр 2003г. равен $t_{2003}^* = -2,5$ и т.д.

Решение системы (10.6) позволило найти следующие значения параметров a_0 и a_1 :

$$a_0 = 112/8 = 14,0 \text{ млрд. руб.}, \quad a_1 = 24,8/42 = 0,59 \text{ млрд. руб.}$$

Уравнение тренда будет иметь следующий вид:

$$Y_t = 14,0 + 0,59t.$$

Для того чтобы получить расчетные значения Y_t , в уравнение "тренда" необходимо подставить соответствующее значение t_i^* . Для примера, чтобы определить значение на уровне 2002 года, необходимо в уравнение тренда подставить значение t_i^* , равное $-3,5$:

$$Y_t = 14,0 + 0,59 \cdot (-3,5) = 11,9 \text{ (млрд. руб.)}$$

Расчетные значения несколько отличаются от фактических, для оценки степени соответствия выбранного уравнения исходным данным обычно определяется ошибка, которая находится по следующей формуле:

$$\delta = \frac{|Y_t - Y_i|}{Y_t} 100\% . \quad (10.7)$$

Если ошибка менее 3%, используемое уравнение удовлетворительно описывает фактические данные. Если ошибка будет больше, необходимо использовать более сложные зависимости. Для этого применяют следующие уравнения:

- уравнение параболы $Y_t = a_0 + a_1 \cdot t + a_2 \cdot t^2$, где a_2 – ускорение;
- уравнение в виде степенной функции: $Y_t = a_0 \cdot a_1^t$;
- уравнение в виде тригонометрических функций: $Y_t = a_0 \cdot \sin x^{a_1}$ и другие функции.

10.2. Интерполяция и экстраполяция

В некоторых случаях при анализе динамических рядов обнаруживаются пропуски или возникает необходимость найти значения, расположенные за пределами рядов динамики (в прошлом или в будущем). Для этого применяются приемы *интерполяции* и *экстраполяции*.

Интерполяция – нахождение неизвестных уровней внутри ряда. Возможны следующие варианты интерполяции:

1. В случае, если выявляется постоянный абсолютный прирост ($\Delta Y_i = \text{const} = \Delta \bar{Y}$), то неизвестное значение определяется прибавлением к предыдущему известному уровню среднего значения абсолютного прироста для данного ряда:

$$Y_n = Y_{n-1} + \Delta \bar{Y} . \quad (10.8)$$

2. В случае, если присутствует постоянный коэффициент роста ($Kp_i = \text{const} = \bar{K}p$), то неизвестное значение определяется путем умножения предыдущего уровня ряда на среднюю величину коэффициента роста:

$$Y_n = Y_{n-1} \cdot \bar{K}p . \quad (10.9)$$

3. Если заранее нельзя сделать предположения о соответствии аналитических показателей (абсолютного прироста или коэффициента роста) постоянным значениям, неизвестное значение определяется путем нахождения среднего по двум соседним уровням:

$$Y_n = (Y_{n-1} + Y_{n+1}) / 2 , \quad (10.10)$$

где Y_n – неизвестное значение динамического ряда,

Y_{n-1} – предыдущее фактическое значение,

Y_{n+1} – последующее фактическое значение.

Экстраполяция – нахождение неизвестных уровней ряда за его пределами. Основное положение экстраполяции – распространение выявленных в анализе закономерностей развития изучаемого явления на будущее или прошлое, т.е. за пределами фактических значений. Точность прогноза зависит от

того, насколько обоснованными окажутся эти предположения. Следует иметь в виду, что чем короче сроки прогноза, тем надежнее результаты экстраполяции.

Применение методов экстраполяции зависит от характера изменений в базисном ряду динамики и предопределяется постановкой задачи исследования. В основном используются три приема:

1. При экстраполяции уровней развития изучаемого явления на базе ряда динамики с постоянными абсолютными приростами ($\Delta Y = \text{const}$) применяется формула

$$Y_l = Y_n + \Delta Y \cdot l, \quad (10.11)$$

где Y_l – экстраполируемый уровень,

Y_n – конечный уровень базисного ряда динамики,

l – срок прогноза (период упреждения).

Если необходимо рассматривать значения в прошлом периоде, то вместо знака «плюс» в формуле (10.11) необходимо поставить знак «минус».

2. При экстраполяции уровня развития изучаемого явления на базе динамики со стабильными коэффициентами ($Kp = \text{const}$) применяется формула

$$Y_l = Y_n \cdot (Kp)^l. \quad (10.12)$$

Необходимо сделать замечание, что если исследуется прошлый период, в формуле (10.12) вместо знака «умножение» должен быть знак «деление».

На практике результат экстраполяции прогнозируемых уровней обычно выполняется не точечными (дискретными), а интервальными оценками.

Для определения границ интервалов используется формула

$$Y_l = \bar{Y}_l \pm t_\alpha \cdot \sigma_{\text{ост}}, \quad (10.13)$$

где t_α – коэффициент доверия по распределению Стьюдента,

$\sigma_{\text{ост}}$ – остаточное среднее квадратическое отклонение тренда, скорректированное по числу степеней свободы ($n-m$), данное значение вычисляется по известным формулам математической статистики,

n – число уровней базисного ряда динамики,

m – число параметров адекватной модели тренда.

В заключение данного раздела важно иметь в виду, что экстраполяция в рядах динамики носит не только приближенный, но и условный характер.

10.3. Приемы коррелирования динамических рядов

В рядах динамики нередко возникает необходимость измерить тесноту взаимосвязи этих процессов. Эта задача решается путем *коррелирования* уровней динамических рядов.

Корреляция – статистическая взаимосвязь двух или нескольких величин. Математической мерой корреляции двух случайных величин служит коэффициент корреляции. Простейшим приемом коррелирования является нахождение

ние коэффициента корреляции по принципу вариационных рядов. В этом случае расчет коэффициента корреляции при оценке степени взаимосвязи двух рядов динамики (X_t и Y_t) определяется по следующей формуле:

$$r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}, \quad (10.14)$$

где $\bar{x}, \bar{y}$ – средние значения признаков рядов динамики,

σ_x, σ_y – соответствующие среднеквадратические отклонения признаков.

Коэффициент корреляции изменяется в интервале от -1 до $+1$.

Возможны три предельные ситуации:

а) если коэффициент корреляции стремится к -1 , то существует обратно пропорциональная зависимость. Например, чем больше пропусков занятий, тем ниже успеваемость студентов;

б) если коэффициент корреляции стремится к $+1$, то существует прямо пропорциональная зависимость. Например, чем больше студентов присутствуют на занятии, тем выше будет успеваемость курса;

в) если коэффициент корреляции близок к нулю, то между двумя параметрами не существует взаимосвязи.

Следует также заметить, что коэффициенты корреляции могут стремиться к граничным значениям $-1, +1$, однако это крайне редкая ситуация. В случае равенства граничным значениям присутствует жесткая линейная связь. Например, в тригонометрии расчет длины окружности: $L = 2\pi R$. Между параметрами R и L существует жесткая линейная связь. В социально-экономической жизни общества таких связей практически не бывает.

Встречаются ситуации, когда при сопоставлении двух процессов, формула (10.14) может показать наличие взаимосвязи там, где ее на самом деле нет (например, динамика успеваемости студента и изменения вкладов в швейцарских банках). Причина заключается в том, что оба динамических ряда обладают явлением *автокорреляции*. *Автокорреляция* – статистическая взаимосвязь между процессами, обладающими однонаправленными трендами. Для того чтобы определить действительную меру тесноты связи, нужно устранить явление автокорреляции. Для этого используются специальные приемы математико-статистического анализа.

10.4. Статистическое изучение сезонных колебаний

При рассмотрении динамики общих явлений, подверженных влиянию природных и других естественных факторов, нередко наблюдается так называемая сезонность.

Сезонность проявляется в последовательной циклической смене снижения уровней явлений, их повышения, и наоборот. Сезонность – это внутригодовые колебания уровней динамических рядов под влиянием времен года. Например: производство электроэнергии, надой молока (рис. 10.1) и др.

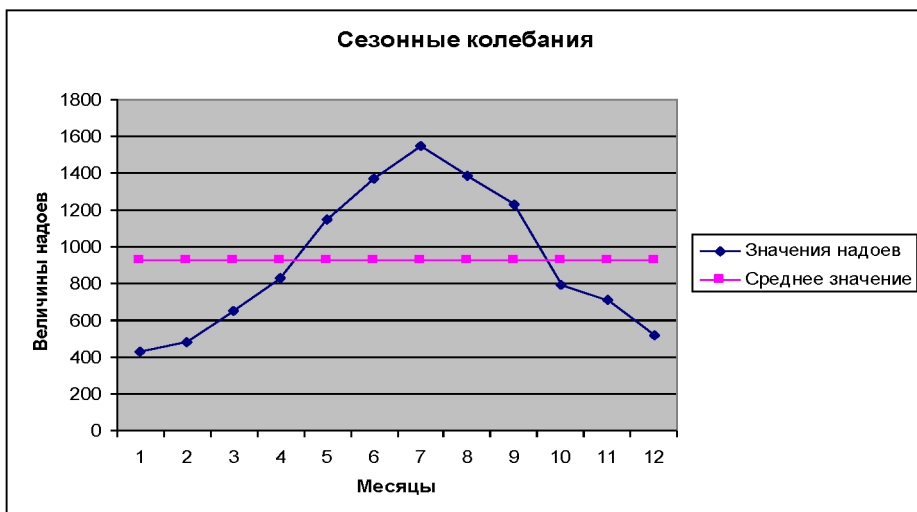


Рис.10.1. Сезонные колебания величины надоев

Задача статистики заключается в том, чтобы измерить сезонность и тем самым предусмотреть ее влияние на будущее, предотвратив нежелательные последствия.

Для измерения сезонных колебаний обычно используются *индивидуальный и сводный коэффициенты сезонности*.

Индивидуальные (ежемесячные) коэффициенты сезонности вычисляются методом процентных отношений:

$$k_i = \frac{Y_i}{\bar{Y}} \cdot 100\% \quad (i = 1, 2, \dots, 12), \quad (10.15)$$

где Y_i – месячные уровни показателей,

$\bar{Y}$ – среднемесячный уровень показателя за год.

Относительная величина индивидуального коэффициента сезонности непосредственно характеризует исследуемый процесс. Процентный показатель сразу дает представление о текущем значении относительно среднего уровня (100%).

Сводный коэффициент сезонности вычисляется, как среднее арифметическое из абсолютных значений отклонений индивидуальных коэффициентов сезонности от 100%:

$$K = \frac{\sum_{i=1}^{12} |k_i - 100\%|}{12}, \quad (10.16)$$

где k_i – индивидуальный коэффициент сезонности,

i – порядковый номер месяца.

Сводный коэффициент сезонности характеризует сезонные колебания за некоторый промежуток времени в целом. Формула (10.16) приведена для расчета сводного коэффициента сезонности в расчете на годовой период.

Рассмотрим использование коэффициентов на примере данных табл. 10.2.

Среднемесячная величина надоев для рассматриваемого примера составляет

$$\bar{y} = \frac{\sum y_i}{n} = \frac{11074}{12} = 922,8 \text{ центнера.}$$

Величины индивидуальных коэффициентов сезонности вычислялись по формуле (10.15). Так, значение индивидуального коэффициента за январь (первый месяц) составило

$$k_1 = \frac{427}{922,8} \cdot 100\% = 46,3\% .$$

Сводный коэффициент сезонности составляет

$$K = \frac{\sum |k_i - 100\%|}{12} = \frac{436,9}{12} = 36,4\% .$$

Таблица 10.2. *Данные о внутригодичных колебаниях надоев молока молочной фермы*

№ п/п	Месяцы	Уровни надоев, центнеры	Индивидуальные коэффициенты сезонности	Абсолютные отклонения
	t_i	y_i	k_i	$ k_i - 100\% $
1	Январь	427	46.3	53.7
2	Февраль	483	52.3	47.7
3	Март	653	70.8	29.2
4	Апрель	830	90.0	10.0
5	Май	1150	124.6	24.6
6	Июнь	1338	145.0	45.0
7	Июль	1550	168.0	68.0
8	Август	1387	150.3	50.3
9	Сентябрь	1233	133.6	33.6
10	Октябрь	793	85.9	14.1
11	Ноябрь	710	76.9	23.1
12	Декабрь	520	56.4	43.6
	Всего	11074		436.9

Был рассмотрен наиболее простой случай, когда средний уровень ряда является постоянной величиной. В реальных условиях возможно присутствие тенденции, т.е. средний уровень ряда либо убывает, либо возрастает. В этом случае вычисление коэффициентов сезонности производится с использованием уравнения тренда:

$$k_i = \frac{Y_i}{Y_t} \cdot 100\% \quad (i = 1, 2, \dots, 12), \quad (10.17)$$

где Y_t – выровненные значения уровней ряда по уравнению прямой либо по какому-либо иному аналитическому выражению.

Для анализа внутригодовой динамики могут применяться гармоники ряда Фурье. При аналитическом выражении изменений уровней ряда динамики используется формула

$$Y_t = a_0 + \sum_k (a_k \cdot \cos kt + b_k \cdot \sin kt), \quad (10.18)$$

где a_0, a_k, b_k – параметры,

k – характеризует номер гармоники, который используется с различной степенью точности (обычно k принимается от 1 до 4).

Параметры уравнения определяются методом наименьших квадратов:

$$\begin{aligned} a_0 &= \frac{\sum Y_i}{n}, \\ a_k &= \frac{2}{n} \sum Y_i \cdot \cos kt_i, \\ b_k &= \frac{2}{n} \sum Y_i \cdot \sin kt_i. \end{aligned} \quad (10.19)$$

При анализе ряда внутригодовой динамики по месяцам значение k принимается за 12. Представляя месячные периоды как части окружности, ряд внутригодовой динамики можно записать в таком виде (табл. 10.3):

Т а б л и ц а 10.3. *Радийная привязка координат для уравнения*

Периоды (t_i)	0	$\frac{1}{6}\pi$	$\frac{1}{3}\pi$	$\frac{1}{2}\pi$	$\frac{2}{3}\pi$	$\frac{5}{6}\pi$	π	$\frac{7}{6}\pi$	$\frac{4}{3}\pi$	$\frac{3}{2}\pi$	$\frac{5}{3}\pi$	$\frac{11}{6}\pi$
Уровни (Y_i)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12

Глава 11. ИНДЕКСЫ. ОБЩИЕ ПОНЯТИЯ ОБ ИНДЕКСАХ

11.1. Сущность экономических индексов

С помощью показателей динамических рядов осуществляется сравнение уровней отдельных явлений, непосредственно сопоставимых между собой.

Однако задача сопоставления уровней явлений может быть значительно сложнее, если мы сталкиваемся с проблемой несоизмеримости явлений. Под *несоизмеримостью* следует понимать невозможность непосредственного суммирования. Например, нельзя непосредственно просуммировать разнотипную продукцию предприятия.

Индексами называются относительные показатели, служащие для сопоставления уровней сложных явлений, отдельные элементы которых непосредственно несопоставимы.

Для определения индекса производится сопоставление не менее двух величин, в числителе располагается сравниваемая величина, обычно рассматривается за текущий (или отчетный) период, а в знаменателе – величина, с которой производится сравнение за базисный период. Явление несоизмеримости преодолевается в индексах путем введения весовых показателей или весов. Веса трактуются как соизмерители. Например, соизмерителями продукции являются: цена продукции, трудоемкость продукции и т.д.

Индексы обладают синтетическими и аналитическими свойствами.

Синтетические свойства индексов состоят в том, что посредством индексного метода производится соединение (агрегирование) в целое разнородных единиц совокупности.

Аналитические свойства состоят в том, что посредством индексного метода определяется влияние факторов на изменение изучаемого показателя. Использование индексов в аналитических целях – один из важных аспектов экономических разработок.

11.2. Виды индексов

Индексы могут быть различных видов:

- 1) индивидуальные i , которые вычисляются по отдельным элементам системы. Например, индекс цен на конкретный товар («городская» булка);
- 2) групповые или субиндексы I , которые вычисляются по однородным группам. Например, индекс цен на хлебобулочные изделия;
- 3) общие индексы I вычисляются по всей совокупности явлений в определенной системе. Например, индекс цен на продовольственные товары.

Второй и третий виды индексов имеют название сводных – общее название, т.к. они характеризуют соотношение уровней нескольких элементов совокупности.

11.3. Основные методологические положения теории индексов

Изучение методологических положений теории индексов наиболее удобно рассматривать на примере (табл. 11.1). Имеются данные о динамике производства продукции и цен.

Стоит задача: определить изменение в отчетном периоде в сравнении с базисным объемов отдельных видов продукции и всей продукции в целом, а также изменения цен.

Указанная задача по отдельным видам продукции решается с помощью индивидуальных индексов.

Т а б л и ц а 11.1. Цены и объемы продукции предприятия в отчетном и базисном периодах

Вид продукции	Количество продукции		Цена за единицу	
	Базисный период	Отчетный период	Базисный период	Отчетный период
А	q_0^A	q_1^A	p_0^A	p_1^A
Б	q_0^B	q_1^B	p_0^B	p_1^B
В	q_0^B	q_1^B	p_0^B	p_1^B

Примечание: q – quantitas (количество), индекс в нижней части означает временной период: 0 – базисный период, 1 – отчетный период. Пример: q_0^A – количество продукции вида А, произведенной в базисном периоде; p – pretium (цена), p_1^B – цена единицы продукции вида Б, произведенной в отчетном периоде.

11.4. Индивидуальные индексы

11.4.1. Индивидуальный индекс физического объема отдельных видов продукции i_q

Индивидуальные индексы динамики количества отдельных видов продукции определяются по следующим формулам для каждого в отдельности вида продукции (для данных табл. 11.1):

$$i_q^A = \frac{q_1^A}{q_0^A}; \quad i_q^B = \frac{q_1^B}{q_0^B}; \quad i_q^B = \frac{q_1^B}{q_0^B},$$

где i_q^A – индивидуальный индекс количества продукции вида А.

В общем виде формула индивидуального индекса количества отдельных видов продукции имеет вид

$$i_q = \frac{q_1}{q_0}, \tag{11.1}$$

где q_1 и q_0 – количество продукции конкретного вида, произведенное в отчетном и базисном периодах соответственно. В данной формуле (для простоты записи) обозначение конкретного вида продукции опускается.

Индивидуальный индекс количества продукции показывает, во сколько раз изменился выпуск количества продукции в отчетном периоде по сравнению с базисным.

11.4.2. Индивидуальный индекс цен на отдельные виды продукции i_p

Динамика цен на отдельные виды продукции отражается в следующем виде:

$$i_p^A = \frac{p_1^A}{p_0^A}; \quad i_p^B = \frac{p_1^B}{p_0^B}; \quad i_p^B = \frac{p_1^B}{p_0^B}.$$

В общем виде формула индивидуального индекса цен отдельных видов продукции имеет вид

$$i_p = \frac{p_1}{p_0}, \quad (11.2)$$

где p_1 и p_0 – цена продукции конкретного вида, произведенного в отчетном и базисном периодах соответственно.

Индивидуальный индекс цены отдельного вида продукции показывает, во сколько раз изменилась цена продукции в отчетном периоде по сравнению с базисным. В табл. 11.2 приведены расчеты индивидуальных индексов цен.

Т а б л и ц а 11.2. *Данные о ценах некоторого продукта*

Виды цен	Сентябрь	Ноябрь	Индивидуальный индекс цен $p_1 : p_0$
Цена на рынке за 1 кг, руб.	30	40	1.33 или 133 %

Вычисленные индивидуальные индексы показывают, что цена за 1 кг данного продукта на рынке была в ноябре на 33 % выше, чем в сентябре месяце.

11.5. Общие индексы

Всякий общий индекс состоит из двух частей:

- 1) индексированных величин;
- 2) весов.

Под индексированными величинами понимают показатели, общие значения которых в отчетном периоде сравниваются с общей их величиной в базисном периоде.

Весами называются показатели, с помощью которых удастся соизмерить индексированные величины.

В случае формирования индекса количества продукции I_q , где сопоставляются объемы продукции q_1 и q_0 , весами являются цены p . В случае сопоставления цен (индекс цен I_p) в отчетном и базисном периодах (соответственно p_1 и p_0) весами являются количества продукции q .

Примером индексного метода изучения динамики сложных статистических совокупностей являются данные табл. 11.3 о ценах и реализации товаров за два периода.

Т а б л и ц а 11.3. *Цены и объемы продаж товаров*

Товар	Ед. изм.	1-й период		2-й период		Индивидуальные индексы	
		Цена за единицу	Количество	Цена за единицу	Количество	цен	физического объема
		руб. (p_0)	(q_0)	руб. (p_1)	(q_1)	$i_p = p_1 : p_0$	$i_q = q_1 : q_0$
А	кг	40	7500	45	9500	1.125	1.27
Б	кг	50	2000	55	5500	1.100	2.75
В	м	75	1000	40	1500	0.53	1.50

Индивидуальные индексы показывают, что в текущем периоде по сравнению с базисным цена товара А повысилась на 12,5 %, на товар Б увеличилась на 10%, на товар В снизилась на 47 %. Количество реализации товара А возросло на 27 %, товара Б – на 175 %, а товара В – на 50 %.

Разновеликие по направлению и интенсивности изменения индивидуальных индексов обуславливают необходимость определения общего для данного ассортимента изменения цен и количества реализованных товаров. Для этого вычисляются соответствующие общие индексы.

11.5.1. Общий индекс стоимостного объема продукции I_{pq}

Для построения данного индекса найдем несколько составляющих. Первоначально найдем количество продукции базисного периода в ценах базисного периода:

$$q_0^A \cdot p_0^A + q_0^B \cdot p_0^B + q_0^C \cdot p_0^C = \sum q_0 \cdot p_0. \quad (11.3)$$

Продукция отчетного периода в ценах отчетного периода:

$$q_1^A \cdot p_1^A + q_1^B \cdot p_1^B + q_1^C \cdot p_1^C = \sum q_1 \cdot p_1. \quad (11.4)$$

Таким образом, получены стоимостные показатели продукции в ценах соответствующих периодов. Их соотношение дает общий индекс стоимостного объема продукции:

$$I_{pq} = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_0}. \quad (11.5)$$

Этот индекс показывает, во сколько раз стоимость продукции отчетного периода больше стоимости продукции базисного периода.

Величина индекса I_{pq} стоимостного объема продукции определяется двумя факторами: изменением количества продукции и изменением цен.

11.5.2. Общий индекс физического объема продукции I_q

Для того чтобы измерить динамику количества продукции, необходимо исключить динамику цен, т.е. необходимо продукцию отчетного периода выразить в неизменных ценах – ценах базисного периода. Данная процедура называется «замораживанием» цен. Стоимость продукции объема отчетного года в ценах базисного года определяется по следующей формуле:

$$q_1^A \cdot p_0^A + q_1^B \cdot p_0^B + q_1^C \cdot p_0^C = \sum q_1 \cdot p_0.$$

Индекс физического объема продукции определяется следующим образом:

$$I_q = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0}. \quad (11.6)$$

11.5.3.Общий индекс оптовых цен I_p

Динамика цен может выражаться следующими общими индексами цен:

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0} \quad \text{и} \quad I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}.$$

В первом случае при определении индекса цен в качестве соизмерителя индексируемых величин p_1 и p_0 применяются данные о количестве реализованной продукции в базисном периоде q_0 . При этом в числителе индексного отношения образуется значение $\sum p_1 \cdot q_0$, т.е. сумма стоимости продажи товаров базисного периода по ценам текущего периода. В знаменателе индексного отношения образуется значение $\sum p_0 \cdot q_0$, т.е. сумма стоимости продажи товаров в базисном периоде по ценам того же базисного периода.

Формула такого общего индекса имеет вид

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0}. \quad (11.7)$$

Расчет общего индекса цен по формуле (11.7) предложен немецким экономистом Э.Ласпейресом, поэтому его принято называть индексом Ласпейреса.

Применим формулу Ласпейреса для расчета индекса цен по данным табл. 11.3:

числитель индексного отношения

$$\sum p_1 \cdot q_0 = 45 \cdot 7500 + 55 \cdot 2000 + 40 \cdot 1000 = 487500 \text{ руб.},$$

знаменатель индексного отношения

$$\sum p_0 \cdot q_0 = 40 \cdot 7500 + 50 \cdot 2000 + 75 \cdot 1000 = 475000 \text{ руб.}$$

Полученные величины подставим в формулу (11.7):

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0} = \frac{487500}{475000} = 1.026 \text{ или } 102,6 \, \%.$$

Применение формулы (11.7) показывает, что по ассортименту в целом повышение цены составило в среднем 2.6 %.

Разность числителя и знаменателя формулы (11.7) представляет собой показатель прироста товарооборота ΔT при продаже товаров в базисном периоде по ценам текущего периода:

$$\Delta T = 487500 - 475000 = 12500 \text{ руб.}$$

Полученная сумма прироста товарооборота показывает, что повышение цен в текущем периоде в среднем на 2.6% соответствует увеличению объема товарооборота на 12.5 тыс.руб.

При определении общего индекса цен I_p в качестве соизмерителя индексируемых величин p_1 и p_0 могут применяться данные о количестве реализованной продукции отчетного периода q_1 . При этом в числителе индексного

отношения образуется значение $\sum p_1 \cdot q_1$, т.е. сумма стоимости продажи товаров текущего периода по ценам текущего периода. В знаменателе индексного отношения образуется значение $\sum p_0 \cdot q_1$, т.е. сумма стоимости товаров в текущем периоде по ценам базисного периода.

Формула такого общего индекса имеет следующий вид:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}. \quad (11.8)$$

Расчет индекса цен по формуле (11.8) был предложен немецким экономистом Г. Пааше. Поэтому данный индекс называют индексом Пааше.

Применим формулу (11.8) для расчета индекса цен по данным табл. 11.3: числитель индексного отношения

$$\sum p_1 \cdot q_1 = 45 \cdot 9500 + 55 \cdot 5500 + 40 \cdot 1500 = 790000 \text{ руб.},$$

знаменатель индексного отношения

$$\sum p_0 \cdot q_1 = 40 \cdot 9500 + 50 \cdot 5500 + 75 \cdot 1500 = 767500 \text{ руб.}$$

Тогда значение индекса Пааше получит следующий вид:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1} = \frac{790000}{767500} = 1.03 \text{ или } 103 \, \%.$$

Применение формулы (11.8) показывает, что по данному ассортименту товаров в целом цены повысились в среднем на 3%. Прирост товарооборота за счет фактора изменения цен в текущем периоде по сравнению с базисным периодом составил

$$\Delta T = 790000 - 767500 = 22500 \text{ руб.}$$

Полученная величина прироста говорит о том, что повышение цен на данный ассортимент товаров в среднем на 3 % привело к увеличению товарооборота в текущем периоде на 22,50 тыс. руб.

Таким образом, выполненные по формулам (11.7) и (11.8) расчеты имеют разные показания индексов цен. Это объясняется тем, что индексы Пааше и Ласпейреса характеризуют различные качественные особенности изменения цен.

Индекс Пааше характеризует влияние изменения цен на стоимость товаров, реализованных в отчетном периоде. Индекс Ласпейреса показывает влияние изменения цен на стоимость количества товаров, реализованных в базисном периоде.

Применение индексов Пааше и Ласпейреса зависит от цели исследования. Если анализ производится для определения экономического эффекта от изменения цен в отчетном периоде по сравнению с базисным, то применяется индекс Пааше. Данный индекс отображает разницу между фактической стоимостью продажи товаров в отчетном периоде $\sum p_1 \cdot q_1$ и некоторой условной стоимостью продажи этих же товаров по ценам базисного периода $\sum p_0 \cdot q_1$.

Если целью анализа является определение объема товарооборота при продаже в предстоящем периоде такого же количества товаров, что и в базисном периоде, но по новым ценам, то применяется индекс Ласпейреса. Этот индекс позволяет вычислять разность между суммой возможного объема товарооборота при объемах продаж базисного периода по ценам текущего периода $\sum p_1 \cdot q_0$ и суммы фактического товарооборота базисного периода $\sum p_0 \cdot q_0$. Эти особенности индекса Ласпейреса обуславливают его применение при прогнозировании объема товарооборота в связи с намечаемыми изменениями цен на товары в предстоящем периоде.

В ряде случаев, когда происходят существенные изменения объемов продаж или анализируются несколько периодов времени, в качестве соизмерителей выступают средние величины реализации товаров за два или большее число периодов. При таком способе расчета формула общего индекса записывается в следующем виде:

$$I_p = \frac{\sum p_1 \cdot \bar{q}}{\sum p_0 \cdot \bar{q}}, \quad (11.9)$$

где $\bar{q}$ – среднее количество товаров, реализованных за анализируемые периоды.

Данный индекс был предложен Лоу, поэтому его называют индексом Лоу. Если известны данные о количестве реализованной продукции в базисном и текущем периодах, то их средняя величина определяется методом средней простой:

$$\bar{q} = \frac{q_1 + q_2}{2},$$

где q_1 и q_2 – объемы продукции сравниваемых периодов времени.

Применительно к данным табл. 11.3 при средних величинах реализации товара А – 8500 кг, товара Б – 3750 м, товара В – 1250 м расчет общего индекса составит

$$I_p = \frac{\sum p_1 \cdot \bar{q}}{\sum p_0 \cdot \bar{q}} = \frac{45 \cdot 8500 + 55 \cdot 3750 + 40 \cdot 1250}{40 \cdot 8500 + 50 \cdot 3750 + 75 \cdot 1250} = 1,028 \text{ или } 102,8\%.$$

Индекс цен Лоу применяется в расчетах при закупках или реализации товара в течение продолжительных периодов времени (пятилетие, десятилетие).

11.5.4. Правило выбора весов в общих индексах

При построении общих индексов возникает проблема выбора периода времени, на уровне которого принимаются веса.

При этом исходят из допущения о том, что при изменении динамики количества продукции экономически обоснованно принимать цены базисного периода.

Тогда, учитывая зависимость $I_{qp} = I_q \cdot I_p$, определяем индекс цен I_p по весам отчетного периода (q_1), иначе указанная взаимосвязь будет неправомерной.

Существует *формальное правило выбора периода времени*: при построении общих индексов количественных показателей веса принимаются на уровне базисного периода, а при построении индексов условно-качественных показателей веса принимаются на уровне отчетного периода. Чтобы корректно использовать данное правило, необходимо определиться, какой параметр является количественным, а какой будет условно-качественным. Так, в рассмотренном индексе товарооборота количественным параметром будет количество продукции, соответственно цена будет являться условно-качественным параметром.

Существует единственное исключение из этого правила. Расчет индекса цен потребительской корзины (необходимый набор продуктов) проводится по весам базисного периода. Это сделано для того, чтобы учитывать реально существующую (сложившуюся) структуру потребления продукции.

11.5.5. Способы построения общих индексов

Всякий общий индекс может быть построен двумя тождественными способами:

- 1) способом агрегатных индексов;
- 2) способом средних индексов.

Агрегатным индексом называется общий индекс, в числителе и знаменателе которого берутся наборы (агрегаты) индексированных величин. Ниже приведена формула общего (агрегатного) индекса количества продукции:

$$I_q = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0}.$$

Второй способ заключается в вычислениях средней величины из индивидуальных индексов. Используя отношение индивидуального индекса количества продукции $i_q = \frac{q_1}{q_0}$, получим выражение для вычисления количества продукции в отчетном периоде:

$$q_1 = i_q \cdot q_0.$$

Запишем общий индекс количества продукции с учетом последнего выражения q_1 через индивидуальные индексы:

$$I_q = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0} = \frac{\sum i_q \cdot q_0 \cdot p_0}{\sum q_0 \cdot p_0}. \quad (11.10)$$

Проводя аналогии с формулой среднего арифметического взвешенного $\bar{X}_a = \frac{\sum x_i \cdot m_i}{\sum m_i}$, можно принять, что вариантами x_i является индивидуальные

индексы i_q , а частотами m_i – товарообороты базисного периода $q_0 \cdot p_0$ по конкретным i -м позициям ассортимента продукции.

Вывод: полученная величина I_q (формула 11.10) представляет собой среднее арифметическое из индивидуальных индексов, взвешенных по товарообороту продукции базисного периода.

11.6. Отдельные виды индексов, употребляемые в статистике России

11.6.1. Индекс стоимостного объема товарооборота I_{qp}

Расчет этого индекса аналогичен расчету индекса стоимостного объема продукции I_{qp} (формула 11.5), но только значение q берется как количество реализованной продукции, p – как розничная цена:

$$I_{pq} = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_0}. \quad (11.11)$$

11.6.2. Индекс физического объема товарооборота I_q

Индекс физического объема товарооборота характеризует собой измерение динамики количества реализованной продукции q и записывается аналогично индексу количества продукции I_q (см. формулу 11.6):

$$I_q = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0},$$

в данной формуле в качестве весов используются цены базисного периода p_0 .

11.6.3. Индекс оптовых цен I_p

Динамика оптовых цен выражается общим индексом цен. При этом возможны две записи, в зависимости от выбора периода времени у соизмерителей это либо отчетный период, либо базисный:

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0} \quad (11.12)$$

$$\text{или } I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}. \quad (11.13)$$

В данных формулах фигурируют оптовые цены отчетного и базисного периодов.

11.6.4. Индекс розничных цен I_p

Индекс розничных цен I_p выглядит как общий индекс оптовых цен, только индексируемые величинами являются розничные цены.

11.6.5. Индекс себестоимости продукции

Индивидуальный индекс себестоимости единицы продукции имеет следующий вид:

$$i_c = \frac{c_1}{c_0}, \quad (11.14)$$

где c_1 , c_0 – себестоимость единицы продукции в отчетном и базисном периодах.

Общий индекс себестоимости продукции:

$$I_c = \frac{\sum c_1 \cdot q_1}{\sum c_0 \cdot q_1}. \quad (11.15)$$

Суммирование ведется по отдельным видам продукции.

В связи с тем что проводится расчет условно-качественного показателя (себестоимости продукции), веса (количество произведенной продукции) берутся на уровне отчетного периода. Методика исчисления индекса себестоимости близка к методике исчисления индекса цен. Отличие в том, что вместо цен на отдельные виды продукции принимается себестоимость их производства.

Числитель данной формулы выражает фактический размер затрат на производство продукции в отчетном периоде, а знаменатель – условную величину затрат на производство того же объема продукции, исходя из себестоимости единицы продукции базисного периода.

Экономический эффект (экономия – знак «-», или перерасход – знак «+») от снижения себестоимости определяется по следующей формуле:

$$\mathcal{E}_c = \sum c_1 \cdot q_1 - \sum c_0 \cdot q_1. \quad (11.16)$$

11.6.6. Индекс заработной платы

Индивидуальный индекс заработной платы отдельной группы работников имеет следующий вид:

$$i_l = \frac{l_1}{l_0}, \quad (11.17)$$

где l_1 , l_0 – зарплата отдельных групп работников соответственно в отчетном и базисном периодах.

Общий индекс заработной платы:

$$I_l = \frac{\sum l_1 \cdot T_1}{\sum l_0 \cdot T_1}, \quad (11.18)$$

где T_1 – численность групп работников отчетного периода.

Данный индекс выражает относительную меру повышения зарплаты в текущем периоде по сравнению с каким-либо другим периодом, принятым за базу сравнения.

11.6.7. Индекс урожайности

Индивидуальный индекс урожайности отдельного вида сельскохозяйственной культуры имеет следующий вид:

$$i_U = \frac{U_1}{U_0}, \quad (11.19)$$

где U_0 , U_1 – урожайность отдельных видов сельскохозяйственных культур соответственно в базисном и отчетном периодах.

Общий индекс урожайности:

$$I_U = \frac{\sum U_1 \cdot S_1}{\sum U_0 \cdot S_1}, \quad (11.20)$$

где S_1 – весенняя продуктивная площадь полей под соответствующими культурами.

Числитель данной формулы выражает фактический валовой сбор культур, знаменатель – валовой сбор с тех же посевных площадей, исходя из урожайности базисного периода.

Разность числителя и знаменателя показывает абсолютную величину прироста валового сбора над весенним повышением урожайности:

$$\Delta \text{Вал.сбор} = \sum U_1 \cdot S_1 - \sum U_0 \cdot S_1.$$

11.6.8. Индекс производительности труда

В связи со сложностью исчисления данного показателя применяется целая система показателей. Их можно подразделить на две группы: *прямые и обратные*.

Прямые показатели непосредственно характеризуют выработку продукции в единицу времени, их называют *производительностью труда по выработке*:

$$W = \frac{q}{T}, \quad (11.21)$$

где q – произведенная продукция,

T – общий фонд затрат рабочего времени на выпуск продукции.

Обратные показатели выражают затраты времени на производство единицы продукции и называются показателями *трудоемкости продукции*:

$$t = \frac{T}{q} . \quad (11.22)$$

Прямые и обратные показатели связаны обратной зависимостью между собой:

$$W = \frac{1}{T} \quad \text{или} \quad t = \frac{1}{W} .$$

Рассмотрим индивидуальный индекс производительности труда по выработке:

$$i_W = \frac{W_1}{W_0} , \quad (11.23)$$

где W_1 , W_0 – выработка продукции в единицу времени в отчетном и базисном периодах соответственно.

Общий индекс производительности труда применительно к выработке будет иметь следующий вид:

$$I_W = \frac{\sum W_1 \cdot T_1}{\sum W_0 \cdot T_1} = \frac{\sum W_1 \cdot d_1}{\sum W_0 \cdot d_1} , \quad (11.24)$$

где d_1 – удельный вес затрат рабочего времени (берется в %).

Общий индекс производительности труда, определяемый по показателям трудоемкости, будет иметь следующий вид:

$$I_t = \frac{\sum t_1 \cdot q_1}{\sum t_0 \cdot q_1} \quad (11.25)$$

где t_1 , t_0 – показатели трудоемкости в отчетном и базисном периодах соответственно.

Следует отметить, что соотношение показателей трудоемкости взвешено по объему производства в отчетном периоде.

$$\text{Поскольку } W = \frac{1}{t}, \text{ то } I_W = \frac{1}{I_t} ,$$

$$I_W(t) = \frac{1}{I_t} = \frac{1}{\frac{\sum t_1 \cdot q_1}{\sum t_0 \cdot q_1}} = \frac{\sum t_0 \cdot q_1}{\sum t_1 \cdot q_1} . \quad (11.26)$$

Это формула определения индекса производительности труда через трудоемкость. Разность между числителем и знаменателем показывает экономию рабочего времени, достигнутую вследствие снижения трудоемкости продукции.

Применение вышеприведенных формул (11.21) – (11.26) зависит от наличия информации. В случае обработки данных по однородной продукции используют общий индекс производительности труда (11.24). В случае обработки данных по неоднородной продукции используется формула общего индекса производительности труда через трудоемкость (11.25).

Ниже приводится пример использования формул (11.21) – (11.26).

Т а б л и ц а 11.4. *Данные по производству разнородной продукции*

Вид продукции	Базисный период		Отчетный период	
	Количество единиц продукции q_0 , тонн,	Затрачено труда T_0 , чел.-дн.	Количество единиц продукции q_1 , тонн	Затрачено труда T_1 , чел.-дн.
А	12,6	27,0	13,4	28,1
Б	31,0	53,0	37,7	59,3
Итого		80,0		87,4

Так как продукция неоднородная, то расчет производительности труда проводим, используя показатели трудоемкости. Рассчитываются показатели трудоемкости в базисном периоде по обоим продуктам:

$$t_0^A = \frac{T_0}{q_0} = \frac{27,0}{12,6} = 2,14 \text{ чел.-дн./т.},$$

$$t_0^B = \frac{T_0}{q_0} = \frac{53,0}{31,0} = 1,71 \text{ чел.-дн./т.},$$

$$I_W(t) = \frac{\sum t_0 \cdot q_1}{\sum t_1 \cdot q_1} = \frac{2,14 \cdot 13,4 + 1,71 \cdot 37,7}{28,1 + 59,3} = 1,06.$$

Индексы прямых показателей можно исчислять методом среднего арифметического из индивидуальных индексов производительности труда: $i_W = W_1/W_0$. Взвешивание осуществляется по численности работников с соответствующими показателями роста производительности труда:

$$I_W = \frac{\sum i_W \cdot T_1}{\sum T_1}. \quad (11.27)$$

11.7. Индексный анализ динамики среднего уровня

При рассмотрении динамики качественных (атрибутивных) показателей, которые можно представить в виде средних величин, возможно построение нескольких различных видов индексов, имеющих между собой принципиальные различия.

11.7.1. Индексы постоянного состава

По своей форме они выражаются в виде известных уже агрегатных индексов, например:

$$I_U = \frac{\sum U_1 \cdot S_1}{\sum U_0 \cdot S_1} - \text{индекс урожайности.}$$

Такое название – агрегатный индекс – объясняется тем, что в числителе и знаменателе берутся наборы (агрегаты) индексируемых величин.

Величина этих индексов определяется лишь одним фактором: изменением значения показателей в отчетном периоде в сравнении с базисным периодом.

11.7.2. Индексы переменного состава

Индексы переменного состава представляют по своей сути индекс средней величины качественного показателя.

Индекс переменного состава (например индекс урожайности) обозначается: $I_{\bar{U}}$ (здесь $\bar{U}$ – среднее арифметическое значение урожайности):

$$I_{\bar{U}} = \frac{\sum U_1 \cdot S_1}{\sum S_1} : \frac{\sum U_0 \cdot S_0}{\sum S_0} = \bar{U}_1 : \bar{U}_0. \quad (11.28)$$

Свое название индекс (*индекс переменного состава*) получил потому, что в индексе принимаются различные веса: первая компонента индекса строится по весам отчетного периода, вторая – по весам базисного периода.

На величину данного индекса (например индекса урожайности (11.28)) оказывают влияние два фактора:

1. Изменение индексируемых величин в отчетном периоде по сравнению с базисным (U_1 к U_0). Определяется с помощью индекса постоянного состава.

2. Изменение в структуре совокупностей (например S_1, S_0). В этих условиях влияние второго фактора можно определить с помощью индекса структурного индекса:

$$I_{стр} = I_{\bar{U}} : I_U, \quad (11.29)$$

где $I_{стр}$ – структурный индекс урожайности;

$I_{\bar{U}}$ – индекс переменного состава;

I_U – индекс постоянного состава.

Приведенные формулы можно проиллюстрировать следующим примером (табл. 11.5).

**Т а б л и ц а 11.5. Данные об урожайности и валовом сборе
по одному из фермерских хозяйств**

Сорт с/х культуры	Базисный период			Отчетный период			Индек- сы
	Валовой сбор, тыс. ц	Посевн. пло- щадь, тыс. га	Урожай- ность, ц/га	Валовой сбор, тыс. ц	Посевн. пло- щадь, тыс. га	Урожай- ность, ц/га	
	$\sum U_0 \cdot S_0$	S_0	U_0	$\sum U_1 \cdot S_1$	S_1	U_1	$U_1 : U_0$
Пшеница озимая	81,0	4,5	18,0	129,9	6,8	19,1	1,06
Пшеница яровая	97,2	7,9	12,3	79,9	6,6	12,1	0,984
Рожь озимая	37,8	2,8	13,5	35,9	2,6	13,8	1,022
Итого (средняя)	216,0	15,2	14,2	245,7	16,0	15,4	

Анализ динамики урожайности проводится за счет исследования влияющих на нее факторов:

$$I_U = \frac{\sum U_1 \cdot S_1}{\sum U_0 \cdot S_1} = \frac{245,7}{18,0 \cdot 6,8 + 12,3 \cdot 6,6 + 13,5 \cdot 2,6} = 1,021 \text{ (102,1\%)}.$$

Это означает, что в среднем по трем культурам урожайность повысилась на 2,1%, за счет непосредственного изменения уровня урожайности.

Однако при расчете индекса изменения средней урожайности получим

$$I_{\bar{U}} = \bar{U}_1 : \bar{U}_0 = \frac{245,7}{16,0} : \frac{216,0}{15,2} = 1,085 \text{ (108,5\%)}.$$

Наблюдается парадокс: общее значение не принимает значения частных показателей.

Превышение общего индекса над индивидуальными объясняется тем, что кроме изменения урожайности отдельных культур на общий индекс повлияло еще и изменение в структуре посевных площадей.

Величина индекса $I_{\bar{U}}$ означает, что средняя урожайность повысилась на 8,5%, а урожайность в среднем по трем культурам повысилась на 2,1%.

Определим влияние второго фактора (структуры полей) на урожайность:

$$I_{cmp} = \frac{I_{\bar{U}}}{I_U} = \frac{1,085}{1,021} = 1,063 \text{ (106,3\%)}.$$

За счет изменений в структуре посевной площади урожайность дополнительно возросла на 6,3%.

Аналогично можно получить индексы переменного состава других наиболее важных показателей.

Индекс цен:

$$I_{\bar{p}} = \bar{p}_1 : \bar{p}_0 = \frac{\sum p_1 \cdot q_1}{\sum q_1} : \frac{\sum p_0 \cdot q_0}{\sum q_0} . \quad (11.30)$$

Индекс заработной платы:

$$I_{\bar{l}} = \bar{l}_1 : \bar{l}_0 = \frac{\sum l_1 \cdot T_1}{\sum T_1} : \frac{\sum l_0 \cdot T_0}{\sum T_0} . \quad (11.31)$$

Индекс переменного состава производительности труда *по выработке* рассчитывается по следующим формулам:

а) в случае, если продукция выражается в натуральном выражении (измеряется в штуках, тоннах, декалитрах и т.п.), расчетная формула выглядит следующим образом:

$$I_W = \bar{W}_1 : \bar{W}_0 = \frac{\sum q_1 \cdot T_1}{\sum T_1} : \frac{\sum q_0 \cdot T_0}{\sum T_0} ; \quad (11.32)$$

б) в случае, если используется стоимостной эквивалент объема продукции, расчет ведется по формуле

$$I_W = \frac{\sum q_1 \cdot p_{c1}}{\sum p_{c1}} : \frac{\sum q_0 \cdot p_{c0}}{\sum p_{c0}} ,$$

где P_c – сопоставимая цена периода (измеряется в рублях).

11.8. Особенности индексов выполнения плана

С помощью индексов можно анализировать выполнение плана. В отношении отдельных элементов совокупности задача решается с помощью индивидуальных индексов:

$$i_z = \frac{z_1}{z_{план}} ,$$

где $z_1, z_{план}$ – фактическое и плановое значения выпуска продукции соответственно.

При вычислении общего индекса выполнения плана возможна двойная система взвешивания:

1) взвешивание по весам, значения которых – показатели плана:

$$I_z = \frac{\sum z_1 \cdot d_{план}}{\sum z_{план} \cdot d_{план}} , \quad (11.33)$$

где $d_{план}$ – плановое значение удельного вида конкретного вида продукции;

2) взвешивание по весам – фактическим значениям:

$$I_z = \frac{\sum z_1 \cdot d_1}{\sum z_{\text{план}} \cdot d_1}, \quad (11.34)$$

где d_1 – фактическое значение удельного вида конкретного вида продукции.

Теоретически наиболее обоснованным является первый вариант, т.к. при этом устраняется влияние нарушения планового ассортимента выпуска продукции.

Однако на практике применяется второй вариант взвешивания, т.е. по фактическим весам, т.к. за счет перевыполнения плана по «выгодным» позициям перекрывается невыполнение плана по «невыгодным» позициям (У предприятий всегда есть "выгодные" и "невыгодные" виды продукции).

Для предотвращения данных попыток разработан способ учета продукции по методу "зачетов".

Согласно данному методу:

- если $q_1 > q_{\text{план}}$, то $q_{\text{зачет}} = q_{\text{план}}$,

- если $q_1 < q_{\text{план}}$, то $q_{\text{зачет}} = q_1$.

Сущность этого метода состоит в том, что за $q_{\text{зачет}}$ всегда принимается значение не свыше $q_{\text{план}}$, причем минимальное из двух значений $q_{\text{план}}$, q_1 .

Использование данного способа рассмотрим на следующем примере (табл.11.6). Предприятие выпускает два вида продукции. Продукция вида А имеет низкую цену, прибыль от реализации данного вида малая – позиция «невыгодная». Продукция вида Б пользуется спросом, на нее установлена высокая наценка, продукция «выгодная».

Т а б л и ц а 11.6. *Данные о выпуске продукции на предприятии*

Продукция	Плановый объем выпуска, шт.	Цена за единицу, руб.	Плановый товарооборот, руб.	Фактический объем выпуска, шт.	Фактический товарооборот, руб.	Зачетный объем продукции, шт.	Зачетный товарооборот, руб.
А	1000	50,0	50000	800	40000	800	40000
Б	1000	500,0	500000	1400	700000	1000	500000
Всего	2000		550000	2200	740000	1800	540000

Предприятие фактически по валовому продукту перевыполнило план как по количеству (фактически 2200 шт. при плане 2000 шт.), так и в денежном выражении (фактически 740 000,0 руб. при плане 550 000,0 руб.). Однако перевыполнение плана связано с увеличенным выпуском продукции вида Б, а по виду продукции А план не выполнен. Метод «зачетов» выявляет данный факт. В «зачет» по виду А будет зачислено 800 шт., по виду Б – только 1000 шт. и всего 1800 при плане 2000 шт. (План не выполнен). Аналогичная ситуация

будет наблюдаться и при расчете в денежном выражении. В «зачет» будет зачислено лишь 540 000,0 руб. при плане 550 000,0 руб. Откуда следует, что план в денежном выражении тоже невыполнен.

11.9. Территориальные индексы

Наряду с ранее рассмотренными задачами с помощью индексной методологии можно осуществить межхозяйственные и другие пространственные сравнения по принципу *территориальных индексов*.

Длительное время *территориальные индексы* вычислялись по схеме динамических индексов, т.е. индексы качественных показателей брались по структуре исходного объекта, а индексы количественных показателей вычислялись по структуре базисного объекта.

Например, индекс урожайности Самарской области в сравнении с урожайностью Саратовской вычислялся следующим образом:

$$I_{U_{Cam/Cap}} = \frac{\sum U_{Cam} \cdot S_{Cam}}{\sum U_{Cap} \cdot S_{Cam}}. \quad (11.35)$$

При таком построении территориального индекса в качестве соизмерителей или весов использовалась структура полей Самарской области (для сельскохозяйственных культур, которые имеют лучшую урожайность, выделялись и большие площади). В результате при расчете по данной формуле преимущество получали показатели урожайности Самарской области, соответственно индекс урожайности для Самарской области получался больше единицы, т.е. $I_{U_{Cam/Cap}} > 1$.

Аналогичная ситуация наблюдалась и в статистических отчетах Саратовской области при сравнении урожайности Саратовской области по отношению к Самарской области:

$$I_{U_{Cap/Cam}} = \frac{\sum U_{Cap} \cdot S_{Cap}}{\sum U_{Cam} \cdot S_{Cap}}. \quad (11.36)$$

Учитывая, что в качестве весов выступала структура полей Саратовской области, то на величину общего индекса урожайности влияла данная компонента, и соответственно получался индекс урожайности для Саратовской области больше единицы, т.е. $I_{U_{Cap/Cam}} > 1$.

Происходила парадоксальная ситуация, когда по данным расчета урожайности двух областей получались противоположные результаты. Чтобы избежать таких ситуаций, в течение ряда последних лет территориальные индексы вычисляются по схеме «стандартных» весов.

В качестве стандарта возможно применение структуры какого-либо более сложного объекта, составными частями которого являются сравниваемые единицы.

В результате индекс урожайности необходимо вычислять по следующей формуле:

$$I_{U_{Сам/Сар}} = \frac{\sum U_{Сам} \cdot S_{Поволжье}}{\sum U_{Сар} \cdot S_{Поволжье}}. \quad (11.37)$$

Формула (11.40) расчета индекса урожайности теоретически абсолютна правильная, однако ее реализация практически очень сложна, т.к. помимо структуры засева полей сравниваемых территорий необходимо знать все составляющие по структурам полей всех областей, входящих в данный территориальный регион.

Для упрощения расчетов территориальных индексов принято решение, чтобы в качестве весов использовать суммарные показатели сравниваемых объектов:

$$I_{U_{Сам/Сар}} = \frac{\sum U_{Сам} \cdot (S_{Сам} + S_{Сар})}{\sum U_{Сар} \cdot (S_{Сам} + S_{Сар})}. \quad (11.38)$$

11.10. Взаимосвязанные индексы

Взаимосвязанные индексы – это индексы, в которых признаки связаны между собой в виде мультипликативной модели, т.е. в виде произведения признаков.

С помощью взаимосвязанных индексов может быть решен целый класс аналитических задач:

- анализ взаимосвязей количественных и качественных показателей;
- изучение влияния изменения индексированных величин и структуры совокупностей;
- анализ взаимосвязей базисных и цепных индексов и т.д.

Наиболее важным качеством является взаимосвязь индексов, вытекающая из природы причинно-следственных связей явлений. Примерами таких взаимосвязей являются:

- *Товарооборот* = *Цена продукции* × *Количество*,
- *Валовый сбор* = *Посевная площадь* × *Урожайность*.

Этим взаимосвязям будет соответствовать следующая связь их индексов.

Так, например, индекс товарооборота можно представить как произведение индексов цен и количества продукции:

$$I_{pq} = I_p \cdot I_q, \quad (11.39)$$

индекс валового сбора зерна можно представить как произведение индексов урожайности и посевных площадей:

$$I_{вал.сбор} = I_{урожай} \cdot I_{площ.} \quad (11.40)$$

Изменение этих взаимосвязей имеет значение в двух смыслах:

- 1) по двум известным значениям взаимосвязанных индексов всегда можно найти третий индекс в этой системе;

2) с помощью взаимосвязанных индексов можно осуществить разложение приростов сложных явлений на составляющие их компоненты.

Рассмотрим влияние каждой компоненты сложного явления.

Введем следующие обозначения:

a – значение количественного показателя (признака);

b – значение условно-качественного показателя;

Q – значение объемного показателя.

Взаимосвязанный индекс будет иметь вид

$$Q = a \cdot b,$$

тогда объемные показатели базисного и отчетного периодов запишем в следующем виде:

$$Q_0 = a_0 \cdot b_0,$$

$$Q_1 = a_1 \cdot b_1.$$

Объемный показатель отчетного периода Q_1 можно представить как:

$$Q_1 = Q_0 \cdot I_a \cdot I_b,$$

где $I_a = \frac{a_1}{a_0}, \quad I_b = \frac{b_1}{b_0}.$

Задача состоит в том, чтобы разложить прирост сложного явления на отдельные составляющие, на прирост его под воздействием количественного признака и под влиянием качественного признака:

$$\Delta Q = \Delta Q(a) + \Delta Q(b),$$

$$\Delta Q(a) = ? \quad \Delta Q(b) = ?$$

Учитывая, что $Q_1 = Q_0 \cdot I_a \cdot I_b$, общий прирост объемного показателя ΔQ запишем в следующем виде:

$$\Delta Q = Q_1 - Q_0 = Q_0 \cdot I_a \cdot I_b - Q_0 = Q_0 \cdot (I_a \cdot I_b - 1).$$

В процессе анализа необходимы допущения, в данном случае предполагаем, что на сложное явление действует количественный (экстенсивный) фактор, а затем – условно-качественный (интенсивный) фактор.

Предположим, что сначала действует количественный фактор – параметр a , который меняет свое состояние из a_0 в a_1 , второй параметр b остается постоянным (рис.11.1). Объемный параметр принимает промежуточное состояние $Q^1 = a_1 \cdot b_0$, тогда

$$\Delta Q(a) = Q^1 - Q_0 = Q_0 \cdot I_a - Q_0 = Q_0 \cdot (I_a - 1). \quad (11.41)$$

После перехода в промежуточное состояние происходит воздействие условно-качественного фактора:

$$\Delta Q(b) = Q_1 - Q^1 = Q_0 \cdot I_a \cdot I_b - Q_0 \cdot I_a = Q_0 \cdot I_a \cdot (I_b - 1). \quad (11.42)$$

При достаточно точных вычислениях обеспечивается условие

$$\Delta Q(a) + \Delta Q(b) = \Delta Q.$$

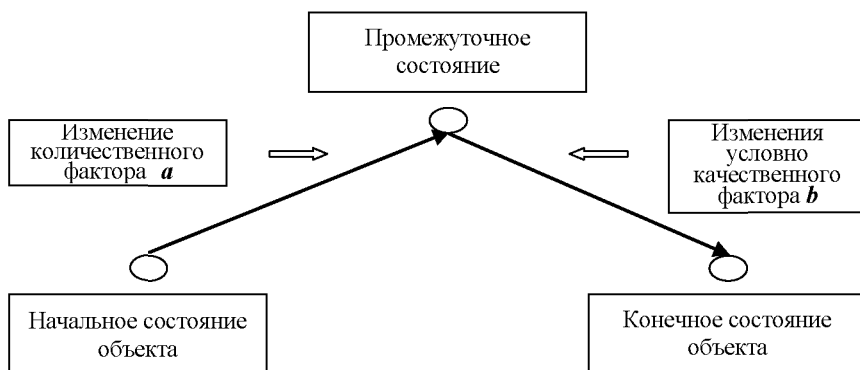


Рис. 11.1. Схема расчета прироста сложного показателя

Пользуясь этим выражением, можно определить относительную меру влияния количественных и качественных признаков на прирост сложного явления:

$$\text{а) } \% \left\langle \begin{array}{c} \text{влияния} \\ \text{количественного} \\ \text{признака (a)} \end{array} \right\rangle = \frac{\Delta Q(a)}{\Delta Q} 100\%,$$

$$\text{б) } \% \left\langle \begin{array}{c} \text{влияния} \\ \text{качественного} \\ \text{признака (b)} \end{array} \right\rangle = \frac{\Delta Q(b)}{\Delta Q} 100\%,$$

В сумме оба компонента будут равны 100%.

11.11. Индексы с постоянными и переменными весами

При классическом рассмотрении теории индексов учитываются два периода времени: базисный и отчетный. Но на практике имеют место задачи, когда встречается несколько периодов времени (пример: квартальные сводные данные – статистическая отчетность РФ).

В этом случае возникает проблема определения базисных и цепных индексов.

Для исследования данных, полученных за несколько периодов времени (в том числе и квартальных), используются следующие определения:

- *базисными* называются индексы, получаемые отношением уровней всех периодов времени к одному и тому же начальному (данные первого квартала принимаются за начальные, см. табл. 11.7);

- *цепными* называются индексы, получаемые отношением уровней последующих периодов времени к предыдущему.

Т а б л и ц а 11.7. *Квартальное распределение цен и объемов продукции*

Периоды (кварталы)	Цена единицы продукции	Количество продукции
I кв.	p_0	q_0
II кв.	p_1	q_1
III кв.	p_2	q_2
IV кв.	p_3	q_3

Рассмотрим построение индивидуальных индексов.

Индивидуальные индексы цен:

$$\text{Базисные: } i_{p_{II/I}} = \frac{p_1}{p_0}; \quad i_{p_{III/I}} = \frac{p_2}{p_0}; \quad i_{p_{IV/I}} = \frac{p_3}{p_0}. \quad (11.43)$$

В нижней части индекса римскими цифрами записаны отношения сравниваемых кварталов.

$$\text{Цепные: } i_{p_{II/I}} = \frac{p_1}{p_0}; \quad i_{p_{III/II}} = \frac{p_2}{p_1}; \quad i_{p_{IV/III}} = \frac{p_3}{p_2}. \quad (11.44)$$

Аналогичным образом рассчитываются индивидуальные индексы количества продукции:

$$\text{Базисные: } i_{q_{II/I}} = \frac{q_1}{q_0}; \quad i_{q_{III/I}} = \frac{q_2}{q_0}; \quad i_{q_{IV/I}} = \frac{q_3}{q_0}. \quad (11.45)$$

$$\text{Цепные: } i_{q_{II/I}} = \frac{q_1}{q_0}; \quad i_{q_{III/II}} = \frac{q_2}{q_1}; \quad i_{q_{IV/III}} = \frac{q_3}{q_2}. \quad (11.46)$$

Общие индексы физического объема продукции q :

$$\text{Базисные: } I_{q_{II/I}} = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0}; \quad I_{q_{III/I}} = \frac{\sum q_2 \cdot p_0}{\sum q_0 \cdot p_0}; \quad I_{q_{IV/I}} = \frac{\sum q_3 \cdot p_0}{\sum q_0 \cdot p_0}. \quad (11.47)$$

$$\text{Цепные: } I_{q_{II/I}} = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0}; \quad I_{q_{III/II}} = \frac{\sum q_2 \cdot p_0}{\sum q_1 \cdot p_0}; \quad I_{q_{IV/III}} = \frac{\sum q_3 \cdot p_0}{\sum q_2 \cdot p_0}. \quad (11.48)$$

Рассмотренные выше общие индексы имеют в качестве весов одну и ту же величину (p_0) и поэтому они получили название индексов с *постоянными весами*.

При рассмотрении общих индексов цен можно заметить отличие

$$\text{как в базисных } I_{p_{II/I}} = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}; \quad I_{p_{III/I}} = \frac{\sum p_2 \cdot q_2}{\sum p_0 \cdot q_2}; \quad I_{p_{IV/I}} = \frac{\sum p_3 \cdot q_3}{\sum p_0 \cdot q_3}, \quad (11.49)$$

$$\text{так и в цепных } I_{p_{II/I}} = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1}; \quad I_{p_{III/II}} = \frac{\sum p_2 \cdot q_2}{\sum p_1 \cdot q_2}; \quad I_{p_{IV/III}} = \frac{\sum p_3 \cdot q_3}{\sum p_2 \cdot q_3}. \quad (11.50)$$

При рассмотрении каждого нового периода времени применяются новые веса и поэтому эти индексы называются *индексами с переменными весами*.

Индексы с постоянными и переменными весами обладают различными методологическими свойствами, например, цепные индексы с постоянными весами можно перемножать, получим базисный индекс конечного уровня к начальному. Индексы с переменными весами перемножать нельзя, это не даст никакого результата.

Проблема выбора постоянных либо переменных весов ассоциирует с проблемой выбора периода времени, на уровне которого принимаются веса.

Индексы с постоянными весами рассматриваются при построении индексов количественных показателей, а при рассмотрении индексов качественных показателей принимаются переменные веса.

Глава 12. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СВЯЗИ ПОКАЗАТЕЛЕЙ КОММЕРЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

12.1. Взаимосвязи показателей коммерческой деятельности и задачи статистики

Изучение взаимосвязей на рынке товаров и услуг – важнейшая функция работников коммерческих служб: менеджеров, коммерсантов и т.д. Особую актуальность это приобретает в условиях развивающейся рыночной экономики.

Изучение взаимодействия спроса и предложения, прибыли и состава предложения товара и других качественных показателей имеет первостепенное значение для прогнозирования конъюнктуры рынка, решения вопросов успешного ведения бизнеса.

Задачей статистики является изучение коммерческой деятельности с количественной стороны, статистика придает выявленным связям количественные характеристики.

Статистические показатели коммерческой деятельности, отображающие взаимосвязь, могут состоять в следующих видах связи:

- балансовой,
- компонентной,
- факторной.

Балансовая связь показателей характеризует зависимость между источниками формирования ресурсов и их использованием.

Она проявляется в формуле товарного баланса

$$O_n - П = В - O_k, \quad (12.1)$$

где O_n , O_k – остаток товаров на начало и конец исследуемого периода,

$П$ – поступление товаров за период,

$В$ – выбытие товаров в изучаемом периоде.

Левая часть характеризует предложение товаров ($O_n + II$), а правая часть – использование товарных ресурсов ($B + O_k$).

Компонентные связи показателей коммерческой деятельности характеризуются тем, что изменения статистических показателей определяются изменением компонентов, входящих в этот показатель как множители:

$$a = b \cdot c. \quad (12.2)$$

Компонентные связи используются в индексном методе для выявления роли отдельных факторов в совокупном изменении сложного показателя. Например, индекс товарооборота в фактических ценах J_{qp} представляет произведение двух компонентов – индекса физического объема товарооборота J_q и индекса цен J_p , т.е.

$$J_{qp} = J_p \cdot J_q. \quad (12.3)$$

Важное значение данных показателей, состоящих в компонентных связях, состоит в том, что можно определить величину одного из неизвестных компонентов:

$$J_q = \frac{J_{pq}}{J_p} \text{ и наоборот } J_p = \frac{J_{qp}}{J_q}. \quad (12.4)$$

Факторные связи в коммерческой деятельности характеризуются тем, что они проявляются в согласованной вариации изучаемых показателей. При этом одни показатели выступают как факторные, а другие – как результативные. По своему характеру этот вид связи является причинно-следственной зависимостью.

В свою очередь *факторные связи* могут рассматриваться как *функциональные и корреляционные*.

При *функциональной связи* изменение результативного признака Y обусловлено действием факторного признака X :

$$Y = f(X). \quad (12.5)$$

Пример: длина окружности l от радиуса (r): $l = 2\pi r$.

Особенностью функциональной связи является то, что она проявляется с одинаковой силой во всех единицах совокупности. Поэтому, установив зависимость на части совокупности, можно ее распространить на всю совокупность. Знание функциональной зависимости позволяет абсолютно точно спрогнозировать событие (например, солнечное затмение).

При *корреляционной связи* изменение результативного признака Y обусловлено изменением факторного признака не всецело, а лишь частично, т.к. возможно влияние других факторов ε :

$$Y = \psi(X) + \varepsilon. \quad (12.6)$$

По своему характеру корреляционные связи – это связи соотносительные. Примером такой связи является зависимость сумм издержек обращения от

объема товарооборота. В этой связи помимо факторного признака (объема товарооборота) на результативный признак (сумму издержек обращения Y) влияют и другие факторы, в том числе и неучтенные ε . Поэтому корреляционные связи не являются полными (жесткими) зависимостями.

Характерной особенностью корреляционных связей является то, что они проявляются не в единичных случаях, а в массе. Поэтому изучаются корреляционные связи по так называемым эмпирическим данным, полученным в статистических наблюдениях.

При статистическом изучении корреляционной связи определяется влияние учтенных факторов при отвлечении от прочих аргументов. Применяемый способ научной абстракции хотя и ведет к некоторому упрощению (аппроксимации) реального механизма связи, но делает возможным установление закономерностей, что позволяет получать количественные характеристики корреляционной связи.

12.2. Методы корреляционно-регрессионного анализа показателей коммерческой деятельности

Использование возможностей вычислительной техники делает практически осуществимым оперативное решение задач изучения корреляционной связи показателей коммерческой деятельности методами корреляционно-регрессионного анализа.

Наиболее разработанной в теории статистики является методология так называемой парной корреляции, рассматривающая влияние вариации факторного признака X на результативный Y .

При изучении связи показателей коммерческой деятельности применяются различного вида уравнения прямолинейной и криволинейной связи.

При анализе прямолинейной зависимости применяется уравнение

$$Y_x = a_0 + a_1 X. \quad (12.7)$$

При криволинейной зависимости применяется ряд математических функций:

- полулогарифмическая $Y_x = a_0 + a_1 \cdot \ln X$,
- показательная $Y_x = a_0 + a_1^X$,
- степенная $Y_x = a_0 \cdot X^{a_1}$ и другие.

Решение уравнений связи предполагает вычисление параметров уравнений. Это осуществляется методом наименьших квадратов. В основу этого метода положено требование поиска параметров при выполнении критерия минимальности сумм квадратов отклонений эмпирических данных Y_i от выровненных Yx_i :

$$\sum_i (Y_i - Yx_i)^2 \Rightarrow \min. \quad (12.8)$$

При изучении корреляционной связи показателей коммерческой деятельности в условиях малого и среднего бизнеса анализу подвергаются сравнительно небольшие по составу единиц совокупности. При численности объектов анализа до 30 единиц возникает необходимость испытания параметров уравнения на их типичность. При этом осуществляется проверка, насколько вычисленные параметры характерны для отображаемого комплекса условий.

Применительно к совокупностям, у которых $n < 30$, для проверки типичности параметров уравнения используется критерий Стьюдента. Для этого определяются фактические значения функционального t -критерия:

$$t_{a_0} = a_0 \frac{\sqrt{n-2}}{\sigma_{\Sigma}}, \quad (12.9)$$

$$t_{a_1} = a_1 \frac{\sqrt{n-2}}{\sigma_{\Sigma}} \sigma_X,$$

где $\sigma_{\Sigma} = \sqrt{\frac{\sum (y_i - y_{x_i})^2}{n}}$ – среднеквадратическое отклонение (СКО) результативного признака y_i от выровненных значений y_{x_i} ;

$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$ – среднеквадратическое отклонение (СКО) факторного признака x_i от общей средней $\bar{x}$.

Полученные t_{a_0} и t_{a_1} сравнивают с критическим значением t_k , который получают по таблице Стьюдента с учетом принятого уровня значимости α ($\alpha = 0.05$) и числа степеней свободы ($n - 2$). Полученные в анализе корреляционные связи параметров уравнения регрессии признаются типичными, если t_{a_0} и t_{a_1} удовлетворяют следующему условию:

$$t_{a_0} > t_k < t_{a_1}. \quad (12.10)$$

По проверенным на типичность параметрам уравнения регрессии производится построение модели связи.

Смысл полученных данных состоит в том, что они характеризуют среднюю величину результативного признака Y_x в зависимости от вариации факторного признака (X).

Важным этапом корреляционного анализа связи является оценка практической значимости. Проверка практической значимости математических моделей осуществляется посредством показателей тесноты связи между признаками X и Y .

Для статической оценки тесноты связи применяются следующие показатели вариации:

1. Общая дисперсия результативного признака σ_y^2 , отражает влияние всех факторов:

$$\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n}. \quad (12.11)$$

2. Факторная дисперсия результативного признака $\sigma_{y_x}^2$, отображающая вариацию Y только от воздействия фактора X :

$$\sigma_{y_x}^2 = \frac{\sum (y_{x_i} - \bar{y})^2}{n}. \quad (12.12)$$

3. Остаточная дисперсия σ_ε^2 , отображающая вариацию результативного признака Y от всех прочих, кроме X , факторов:

$$\sigma_\varepsilon^2 = \frac{\sum (y_i - y_{x_i})^2}{n}. \quad (12.13)$$

4. Соотношение между факторной $\sigma_{y_x}^2$ и общей σ_y^2 дисперсиями характеризует меру тесноты связи между признаками X и Y , которая определяется индексом детерминации:

$$R^2 = \sigma_{y_x}^2 / \sigma_y^2. \quad (12.13)$$

5. При прямолинейной форме связи показатель тесноты связи определяется по формуле линейного коэффициента корреляции r :

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \left[\sum y^2 - \frac{(\sum y)^2}{n} \right]}}. \quad (12.15)$$

Заметим, что линейный коэффициент корреляции r равен индексу корреляции R только при линейной связи.

12.3. Множественная регрессия

На практике часто приходится исследовать зависимость результативного признака от нескольких факторных признаков. В этом случае статическая модель может быть представлена уравнением регрессии с несколькими переменными величинами. Такая регрессия называется *множественной*.

Например, уравнение линейной регрессии с m независимыми переменными:

$$y_i = a_0x_0 + a_1x_1 + a_2x_2 + \dots + a_jx_j + \dots + a_mx_m + \varepsilon_i. \quad (12.16)$$

Уравнение линейной регрессии имеет вид

$$y_{xi} = a_0x_0 + a_1x_1 + a_2x_2 + \dots + a_jx_j + \dots + a_mx_m. \quad (12.17)$$

При оценке параметров этого уравнения в каждом i -м наблюдении фиксируют значение результативного признака y и факторных признаков $x_{j0} \dots x_{jm}$. Слагаемое ε_i является случайным возмущением, имеющим математическое ожидание, равное 0, и дисперсию σ^2 , x_0 – фиктивная переменная, равная 1.

Модель (12.15) в векторном виде выглядит так:

$$Y = Xa + e. \quad (12.18)$$

Линейная модель записывается в виде

$$Y_x = Xa, \quad (12.19)$$

где $e = (e_i)$ – вектор ошибок в уравнении,

$a = (a_j)$ – вектор оценок параметров,

$j = 1, \dots, m$ – число неизвестных параметров,

$X = (X_{ij})$ – матрица значений независимых переменных размерностью $n(m+1)$.

$Y = (Y_i) - i = 1, 2, \dots, n$ – вектор значений зависимой переменной,

n – число наблюдений.

Обозначим сумму квадратов отклонений через $Q = \sum (Y_i - Y_{xi})^2$. Тогда ее значение определяется следующим образом:

$$\begin{aligned} Q &= \sum (Y_i - Y_{xi})^2 = \sum e_i^2 = e^T e = (Y - Xa)^T (Y - Xa) = \\ &= Y^T Y - a^T X^T Y - Y^T Xa + a^T X^T Xa = Y^T Y - 2a^T X^T Y + a^T X^T Xa. \end{aligned} \quad (12.20)$$

Для нахождения вектора неизвестных параметров a возьмем производную от функции Q по a и приравняем к нулю:

$$\frac{\partial Q}{\partial a} = -2X^T Y + 2(X^T X)a = 0,$$

$$X^T Y = X^T Xa,$$

$$a = (X^T X)^{-1} \cdot (X^T Y),$$

$$X = \begin{vmatrix} 1 & x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{im} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nm} \end{vmatrix} \quad (12.21)$$

Таким образом, рассчитываются все параметры уравнения линейной регрессии (12.17). После вычисления параметров уравнения рекомендуется оценить значимость коэффициентов регрессии a_j . Коэффициент регрессии считается статистически значимым, если его значение превосходит средне-квадратическую ошибку не менее чем в 1,95 раза, т.е. фактическое значение функционального t -критерия (12.9) больше значения 1,95.

Вычисление уравнений (12.15)–(12.21) в случае трех и более переменных в ручном режиме (с помощью калькулятора) становится затруднительным, поэтому рекомендуется проводить расчеты либо с помощью универсального программного пакета Excel, либо использовать специализированные пакеты Status, Statistika или иные.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Балинова, В.С. Статистика в вопросах и ответах [Текст]: учеб. пособие / В.С. Балинова. – М.: ТК Велби, Изд-во Проспект, 2004. – 344 с.
2. Гусаров, В.М. Статистика [Текст]: учеб. пособие для вузов / В.М. Гусаров. – М.: ЮНИТИ-ДАНА, 2003. – 463 с.
3. Громыко, Г.Л. Теория статистики [Текст]: учеб. для вузов / Г.Л. Громыко. – М.: Инфра-М, 2003. – 480с.
4. Дуброва, Т.А. Статистические методы прогнозирования [Текст]: учеб. пособие для вузов / Т.А. Дуброва. – М.: ЮНИТИ-ДАНА, 2003 – 206 с.
5. Елисеева, И.И. Общая теория статистики [Текст] / И.И. Елисеева, М.М. Юзбашев. – М.: Финансы и статистика, 2006. – 656 с.
6. Ефимова, М.Р. Общая теория статистики [Текст]: учеб. / М.Р. Ефимова, Е.В. Петрова, В.Н. Румянцев. – М.: ИНФРА-М, 2006. – 416 с.
7. Общая теория статистики [Текст]: учеб. / под ред. И.И. Елисеевой – М.: Финансы и статистика, 2001. – 480 с.
8. Общая теория статистики [Текст]: учеб. / под ред. О.Э. Башиной, А.А Спирина. – М.: Финансы и статистика, 2007. – 440 с.
9. Общая теория статистики: Статистическая методология в изучении коммерческой деятельности [Текст]: учеб. / А.И. Харламов, О.Э. Башина, В.Т. Бабурин [и др.]; под ред. А. А. Спирина, О.Э. Башиной. – М.: Финансы и статистика, 1994. – 279 с.
10. Практикум по теории статистики [Текст] / под ред. Р.А. Шмойловой. – М.: Финансы и статистика, 2001. – 416 с.
11. Статистика [Текст]: учеб. / [И.И. Елисеева и др.]; под ред. И.И. Елисеевой. – М.: Высш. образование, 2009. – 566 с.
12. Статистика [Текст]: учеб. / под ред. В.Г. Ионина. – М.: ИНФРА-М, 2008. – 445 с.
13. Статистика [Текст]: учеб. / под ред. В.Г. Минашкина. – М.: Велби; Проспект, 2007. – 172 с.
14. Теория статистики [Текст]: учеб. / под ред. Г.Л. Громыко. – М.: ИНФРА-М, 2009. – 476 с.

Учебное издание

Кириллов Александр Владимирович

СТАТИСТИКА

Часть 1. Общая теория статистики

Учебное пособие

Редактор Т. К. К р е т и н и н а

Верстка Т. Е. П о л о в н е в а

Подписано в печать 31.10.2012 г. Формат 60х84 1/16.

Бумага офсетная. Печать офсетная. Печ. л. 7,0.

Тираж 100 экз. Заказ . Арт. 19/2012.

Самарский государственный
аэрокосмический университет.
443086 Самара, Московское шоссе, 34.

Изд-во Самарского государственного
аэрокосмического университета.
443086 Самара, Московское шоссе, 34.

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЁВА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»**

А.В. КИРИЛЛОВ

СТАТИСТИКА

Часть 1.

Общая теория статистики

САМАРА 2012