

ОБЩАЯ  
ТЕОРИЯ  
**СТАТИСТИКИ**



СТАТИСТИКА КАК НАУКА  
ОРГАНИЗАЦИЯ СТАТИСТИКИ

СБОР СТАТИСТИЧЕСКОЙ  
ИНФОРМАЦИИ  
И ОБЕСПЕЧЕНИЕ КАЧЕСТВА  
ЕГО МАТЕРИАЛОВ

СТАТИСТИЧЕСКАЯ СОВОКУПНОСТЬ  
И ЕЕ ХАРАКТЕРИСТИКА

СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ  
ВЗАИМОСВЯЗЕЙ

АНАЛИЗ ДИНАМИКИ

# ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

ПОД РЕДАКЦИЕЙ  
А.М. ГОЛЬДБЕРГА,  
В.С. КОЗЛОВА

Допущено Министерством  
высшего и среднего специ-  
ального образования СССР  
в качестве учебника для сту-  
дентов экономических спе-  
циальностей высших учебных  
заведений



ПРОВЕРКА  
200 9



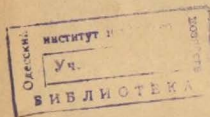
МОСКВА  
„ФИНАНСЫ И СТАТИСТИКА“  
1985



А. Я. БОЯРСКИЙ, Л. Л. ВИКТОРОВА,  
А. М. ГОЛЬДБЕРГ, В. Р. КАМЕНЕЦКИЙ, В. С. КОЗЛОВ,  
Г. Н. КОЗЛОВА, А. С. КРАВЕЦ, Л. П. ОЛЕСЕВИЧ,  
Я. М. ЭРЛИХ

Рецензенты:

М. Р. ЕФИМОВА, Е. В. ПЕТРОВА, В. С. КНЯЗЕВСКИЙ



Общая теория статистики: Учебник / А. Я. Боярский, Л. Л. Викторова, А. М. Гольдберг и др.; Под ред. А. М. Гольдберга, В. С. Козлова. — М.: Финансы и статистика, 1985. — 367 с. ил.

В пер.: 1 р. 10 к. 20 000 экз.

Рассмотрены вопросы организации статистики, сбора статистической информации, характеристики статистической совокупности, статистического изучения взаимосвязей, анализа рядов динамики и индексного анализа. Приводятся методология расчета относительных и средних величин, показателей вариации, регрессии, корреляции и динамики, а также выборочного и индексного методов. Используются данные ежегодников «Народное хозяйство СССР». Примеры условные.

0 0702000000—129  
010(01)—85 60—85

ББК 60.6

31

© Издательство «Финансы и статистика», 1985

## ПРЕДИСЛОВИЕ

Успешное выполнение задач, поставленных Основными направлениями экономического и социального развития СССР на 1986—1990 годы и на период до 2000 года (проект) требует дальнейшего улучшения качества подготовки экономистов высшей квалификации, повышения эффективности учебно-воспитательного процесса. В связи с этим в последние годы значительно обновлены и пересмотрены учебные планы и программы, в которых отражены решения партии и правительства по актуальным проблемам экономического и социального развития страны, совершенствования хозяйственного механизма, новейшие достижения в производстве, науке и технике, в экономической теории и практике. Естественно, что эта работа предъявляет повышенные требования к учебникам и учебным пособиям — важнейшим средствам обучения и самостоятельной работы будущих специалистов в области экономики. Это относится в первую очередь к фундаментальным и специальным дисциплинам, формирующим общетеоретический кругозор, профессиональные знания и практические навыки экономиста любого профиля. Важное место среди этих дисциплин занимает общая теория статистики.

На основе учета этих требований и обобщения опыта применения методов статистического исследования, а также преподавания статистики авторским коллективом и подготовлен настоящий учебник.

Исходя из задач повышения уровня и действенности экономико-статистического анализа обстоятельно изложены такие вопросы курса, как группировки, анализ регрессий и корреляций, анализ динамики, статистическое моделирование и прогнозирование, индексы.

Подробно рассмотрены вопросы организации, содержания статистического наблюдения и обеспечения качества статистической информации в условиях применения новейшей вычислительной техники и внедрения автоматизированной системы государственной статистики.

Более углубленное изложение некоторых вопросов, представляющее интерес для студентов специальности «статистика», набрано в пособии петитом. Этим же шрифтом набраны также

исторические сведения и некоторые дополнительные разъяснения по изложенному материалу.

Учебник подготовлен межвузовским авторским коллективом, в состав которого вошли преподаватели статистики Московского государственного университета им. М. В. Ломоносова, Киевского института народного хозяйства им. Д. С. Коротченко и Одесского института народного хозяйства.

Авторы учебника: проф. А. Я. Боярский — гл. 1; доц. Л. Л. Викторова — гл. 11; проф. А. М. Гольдберг — гл. 2, 3, 5; доц. В. Р. Каменецкий — гл. 7, 8, 9, 10; доц. В. С. Козлов — гл. 12, 13, 14, 15; доц. Г. Н. Козлова — гл. 2, 3, 4, 6; доц. А. С. Кравец — гл. 8, 12, 13; доц. Л. П. Олесевич — гл. 5, 6, 7, 13; доц. Я. М. Эрлих — гл. 12, 13.

Авторы приносят благодарность доцентам кафедры статистики Московского института управления М. Р. Ефимовой и Е. В. Петровой, а также заведующему кафедрой общей теории статистики Ростовского института народного хозяйства профессору В. С. Князевскому за обстоятельное рецензирование рукописи и ценные замечания по улучшению ее структуры и содержания.

## РАЗДЕЛ I

### СТАТИСТИКА КАК НАУКА. ОРГАНИЗАЦИЯ СТАТИСТИКИ

#### ГЛАВА 1

#### СТАТИСТИКА КАК НАУКА. ПРЕДМЕТ И МЕТОД СТАТИСТИКИ

##### 1.1. СОВОКУПНОСТЬ И ВАРИАЦИЯ

Статистика в ее современном понимании связана со счетом каких-то объектов, с их количественными характеристиками. Из этого, однако, не следует, что всякий счет — статистика. Некто подсчитал, что в доме, в котором ему предстоит поселиться, 8 этажей, что в отведенной ему квартире 36 м<sup>2</sup>, т. е. на одного члена семьи придется по 9 м<sup>2</sup>. Назвать это статистикой нельзя. Но если подсчитать, что из всех построенных в данном пятилетии и в данном городе домов столько-то 5-этажных, столько-то 6-этажных и т. д., что в них живет столько-то человек и, следовательно, сколько-то квадратных метров приходится на одного жителя, — это статистика.

Или другой пример. С заводского конвейера вчера сошло 110 автомобилей. И это еще не статистика. Наблюдением, которое велось изо дня в день, установлено число автомобилей, сошедших с конвейера в каждый день; из этих чисел составлен ряд, рассмотрение которого показало, что при общей тенденции к некоторому возрастанию числа автомобилей отмечаются периодические колебания этого числа по дням недели. Это — статистика и не совсем простая.

Из этих примеров видно, что отличает статистику от простого счета. Количественные характеристики, устанавливаемые статистикой, не являются зафиксированными раз и навсегда, одинаковыми для всех объектов. Они меняются (варьируют) от объекта к объекту (от дома к дому, от человека к человеку, от дня ко дню) и вообще во времени. Но чтобы такие изменения (вариации) могли наблюдаться, необходимо множество объектов — домов, людей, дней. Таким образом, *статистика исследует количественные характеристики, которые при определенных обстоятельствах могут варьировать*. Важно заметить, что вариации, о которых идет речь в статистике, это не просто различие. Например, в часах определенной марки шестеренки различаются диаметром, числом зубцов или дни года различаются величиной светового дня. Эти изменения к статистике не имеют отношения: размер данной шестеренки в данном механизме, величина светового дня такого-то числа и месяца (в данном месте) могут быть только такими, а не иными. С другой стороны, длительность службы шестеренки до пол-

ного износа или количество выпавших в данный день осадков могут рассматриваться в статистике: то и другое наперед не задано.

В общей форме это выражается так: индивидуальные значения рассматриваемых в статистике количественных характеристик (признаков) не определены жестко их принадлежностью к данной совокупности или их местом в ней; они могут быть в ней теми или иными, следовательно, содержать элемент случайности. Из количественной характеристики объектов, образующих совокупность, статистика выводит характеристики для всей совокупности в целом или для некоторой ее части, иначе говоря, получает обобщающие показатели («сводные признаки»). Указанное выше свойство вариации делает невозможным получение таких показателей без массового наблюдения.

## 1.2. СТАТИСТИЧЕСКАЯ НАУКА

Массовое наблюдение и получение на основании его результатов обобщающих показателей — отнюдь не такое простое дело, каким оно может показаться с первого взгляда. Это сложный процесс познания, требующий применения выработанных научных методов и правил.

Статистическая практика началась еще в глубокой древности со счета населения. Самые ранние свидетельства имеют более чем 4-тысячелетнюю историю (Китай). В Древнем Риме в цензах (переписях), кроме счета населения (только свободных граждан), отмечались и некоторые элементы их имущества. Однако применение научных методов для получения более точных результатов переписи населения относится к XIX в.

Как общие, так и приспособленные к изучению объектов определенного рода и решению определенных задач методы и правила наблюдения, получения показателей, их анализа разрабатываются статистической наукой. Ее основы были заложены трудами Уильяма Петти, в частности его знаменитой «Политической арифметикой» (1683 г.), что дало К. Марку основание считать Петти в некотором роде изобретателем статистики. Нельзя не упомянуть и работы Дж. Граунта, собравшего за ряд лет сведения об умерших в Лондоне и составившего первую весьма, конечно, примитивную таблицу смертности в виде ряда относительных чисел доживающих до того или иного возраста. Проявление же в статистических числах некоторой закономерности явления было осознано лишь в середине XVIII в. Зюсмильхом, увидевшим в этой закономерности проявление «божественного порядка» (например, в устойчивом соотношении чисел мальчиков и девочек среди новорожденных 21:20).

Другим источником статистической науки, которому статистика обязана и самым своим названием, явились работы немецких ученых по государственному учету, начавших преподава-

ние этого курса в университетских городах Галле, Гельмштадта, Геттингена в конце XVII в. Особо следует выделить труд Ахенвала «Очерк государственного управления европейских государств». Эти исследования получили название описательной школы, а впоследствии статистики.

Государствование в России получило свое выражение в ряде трудов М. В. Ломоносова, в которых он рассматривал вопросы населения, просвещения, народного богатства и производства, государственных финансов и внешней торговли, а также вопросы, впоследствии отнесенные к так называемой моральной статистике. В конце XVIII в. в России появились первые теоретические труды и руководства по статистике. В их числе отметим работы Д. Бернулли, посвященные анализу смертности и браков.

Таким образом, статистика сложилась из элементов политической арифметики и государственного учета. Если от политической арифметики статистика получила комплексный анализ количественных характеристик массовых явлений с целью познания их закономерностей, то от государственного учета — систему количественного описания социально-экономических явлений.

Дальнейшее развитие статистики как науки характеризовалось совершенствованием методов и способов сбора и обработки данных, необходимых для анализа разнообразных массовых социально-экономических процессов и явлений. В XIX в. большой вклад в развитие статистической теории и методологии внесли труды русских ученых Д. П. Журавского, В. Я. Буныковского, А. И. Чупрова, Ю. Э. Янсона и ряда других.

В XIX и XX столетиях Россия, а позднее СССР стал главным центром разработки проблем теории вероятностей, тесно связанной с математическим аппаратом статистики.

Неоценимый вклад в дальнейшее развитие статистической науки и практики представляють труды В. И. Ленина, на методологических основах которых зиждется советская статистика.

## 1.3. ПОЛЕ ДЕЙСТВИЯ СТАТИСТИКИ. СТАТИСТИКА КАК ОБЩЕСТВЕННАЯ НАУКА

Изучаемые статистикой количественные отношения всегда и неразрывно связаны с качеством, с природой соответствующих явлений, с логикой специфического предмета. Поэтому устанавливаемые статистикой закономерности входят в систему знаний соответствующей науки. В. И. Ленин говорит о статистике как о важнейшем орудии социального познания. Вместе с тем применение статистики как метода в различных областях знания обогащает науку. Это можно видеть на примере анализа зависимости между признаками. Появление теории Дарвина стимулировало применение метода статистики в биологии. В частности, сам процесс естественного отбора яв-



ляется массовым, в нем обнаруживаются все свойственные таким процессам черты. В изучении естественного отбора существенное значение имеет вопрос о наследовании потомками особенностей родителей, т. е. о зависимости между их признаками, характеристиками. Для этой цели был найден подходящий инструмент в виде методов анализа корреляции и регрессии (см. гл. 13). Впоследствии эти методы стали применяться и в социально-экономических исследованиях.

Сказанное отнюдь не означает, что методы статистического изучения разных сфер явлений не имеют своих специфических особенностей. Так, в собирании сведений о производстве широко используются записи в разного рода бухгалтерских документах, что исключается в переписях населения. Плотность населения страны измеряется приходящимися на 1 км<sup>2</sup> числом жителей. Однако аналогичное измерение интенсивности загрязнения атмосферы автотранспортом, приходящимися на 1 км<sup>2</sup> территории страны, было бы малопродуктивным ввиду наличия на территории огромных малонаселенных и малодоступных для автотранспорта пространств и т. д.

Результаты статистического измерения, устанавливаемые статистикой числа получают разное значение в естествознании и в общественных науках. В свое время Гегель говорил, что эти числа важны только своими качественными результатами, т. е. закономерностями, получающимися в них свое отражение. Для естествознания это действительно так. Вместе с тем статистические методы широко используются для изучения и анализа социальных и экономических явлений и процессов, что позволяет статистику отнести к кругу общественных наук.

Всесторонне характеризуют различные общественные явления, их количественные и качественные соотношения и изменения, статистика служит методом, орудием познания для экономической науки, истории, социологии и других общественных наук. Вместе с тем статистика — самостоятельная наука, рассматривающая количественные характеристики общественных явлений и их соотношения как таковые.

В конце прошлого века была установлена закономерность: с ростом доходов семьи уменьшается доля расходов на питание. Пока речь идет только об этой закономерности статистика выступает как метод, но для практических действий нужно не только понимание этой общей закономерности, но и конкретные числа, в которых эта закономерность проявляется. Если в плане намечается увеличение доходов и определенное изменение распределения семей по их размеру, то важно определить, как именно изменится структура спроса, как лучше сбалансировать с ним структуру предложения, отсюда — как спланировать производство различных предметов потребления и т. д. И здесь — при исследовании конкретного количественного выражения закономерности — статистика выступает уже как самостоятельная общественная наука.

Подведем итог. Статистика — наука, изучающая массовые явления, применяется как метод познания закономерностей в любой области, где массовые явления имеют место; в то же время статистика — это самостоятельная общественная наука, исследующая конкретные количественные размеры и соотношения общественных явлений в конкретных условиях места и времени.

В этой формулировке общественные явления надо понимать широко. Технология к общественным наукам не относится. Но применение в ней статистики не может ограничиваться только установлением закономерностей. Здесь нужно и научное познание конкретных параметров, так как техника применяется для своих целей людьми. Медицина в целом — не общественная наука (за исключением социальной гигиены). Однако для исследования состояния здоровья населения важно знать конкретно уровень заболеваемости различными видами болезней, степень медицинского обслуживания населения, и т. д., а для конкретных практических расчетов нужно знать и длительность лечения болезни, обеспеченность населения больничными койками и т. д. Все это относится к статистике как общественной науке.

#### 1.4. СТАТИСТИКА И КАЧЕСТВЕННЫЙ АНАЛИЗ

Как сказано выше, конкретные приемы статистического изучения явления тесно связаны с его качественной природой. Следовательно, без анализа этой природы нельзя правильно определить ни способы собирания материала наблюдений, ни способы его обработки, ни тем более направление и способы анализа результатов. Критикуя земских статистиков, не видевших в массе тогдашних крестьянских хозяйств совершенно различные экономические типы, неизбежность возникновения которых следует из экономической теории К. Маркса, В. И. Ленин писал: «Будущий историк русской экономической литературы с удивлением отметит тот факт, что предрассудки народничества привели к забвению самых элементарных требований экономической статистики, обывающих строго разделять хозяев и наемных рабочих, какой бы формой землевладения они ни были объединены, как бы ни были многочисленны и разнообразны переходные типы между ними»<sup>1</sup>. Одним из важных методов статистики является группировка, которой посвящена ниже специальная глава. Земские статистики, придерживавшиеся народнических взглядов, были не способны применять статистику на базе правильного понимания качества явления. Это проявилось в ряде их работ в том, что в основу группировки хозяйств была положена величина земель-

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 3, с. 164.

ного надела, зависевшая не от экономической мощности хозяйства, а в основном от числа членов семьи («едоков»). В. И. Ленин резко критиковал такие группировки. «Пользуясь группировкой по наделу,— писал Ленин,— мы складываем вместе бедняка, который сдает землю, и богача, который арендует или покупает землю;— бедняка, который забрасывает землю, и богача, который «собирает» землю;— бедняка, который ведет самое плохое хозяйство с ничтожным количеством скота, и богача, который имеет много скота, удобряет землю, вводит улучшения и пр. и пр. Мы складываем, другими словами, сельского пролетария с представителями сельской буржуазии»<sup>1</sup>. Вместо этого В. И. Ленин использовал при группировке крестьянских хозяйств посевную площадь, количество скота и другие признаки, действительно отражающие тип хозяйства, что и дало ему возможность раскрыть процесс классового разложения крестьянской массы.

В этой связи следует помнить, что в статистике неизменно проявляется борьба двух методологических концепций. Первую коротко можно сформулировать как принцип примата качественного анализа, о котором только что была речь. Другая— концепция чистого («голого») эмпиризма. Придерживающийся ее статистик собирает материал наблюдения, чаще всего бессистемно, набирая много лишних сведений и упуская из виду существенное для характеристики явления, нередко даже не умея правильно очертить границы изучаемой совокупности. Такой подход еще объясним для статистической науки XVI или XVII в. Однако такой подход несовместим с теоретической основой статистики—марксистско-ленинской политической экономией, диалектическим и историческим материализмом. Еще хуже, когда это игнорирование теории или подмена ее извращенным пониманием исследуемых явлений преследует прямые цели искажения истины.

При честном отношении к делу не исключено, что в эмпирическом материале обнаружится нечто общее, достойное быть отмеченным и подвергнутым дальнейшему исследованию. Как замечает по поводу эмпиризма Ф. Энгельс, что и слепая свинья может найти свой желудок. Но «желудок» эмпирика гораздо чаще оказывается ложным. Так происходит тогда, когда эмпирический подход ставится на службу апологетики буржуазного строя. В прошлом веке для капиталистического воспроизводства была характерна примерно 11-летняя периодичность кризисов. Как известно, примерно таков же период солнечной активности. Достаточно было такого совпадения, чтобы объявить солнечные пятна причиной кризисов. А когда кризисы стали чаще, примерно раз в 8 лет, на небе было найдено другое явление с совпадающим периодом—взаимное расположение

Земли, Солнца и Венеры. И это—вместо анализа действительной природы кризисов, их связи со сменой основного капитала и т. д., что астрономическими явлениями никак не обусловлено.

## 1.5. ПАРТИЙНОСТЬ СТАТИСТИКИ

Стремление к подлинной истине—неотъемлемое свойство марксистско-ленинской статистики, теоретической основой которой являются *марксистско-ленинская политическая экономия, диалектический и исторический материализм*. Марксистско-ленинская наука указывает человечеству путь к будущему. Ей поэтому нечего скрывать или показывать в неверном свете. Наоборот, чем яснее и точнее будут познаны прошлое и настоящее, тем быстрее и с меньшими издержками (в широком смысле слова) будет движение к этому будущему. Буржуазная же наука в противоположность этому заинтересована в торможении истории, в отдалении этого будущего. С конкретными проявлениями этого приходится то и дело встречаться. И, разумеется, чем правее и сильнее вопрос затрагивает существенные интересы людей, тем больше истина извращается буржуазной статистикой.

Яркий пример этому дает статистика безработицы, являющейся бычком трудового народа в капиталистическом мире. Конечно, речь идет не о том, что публикуются прямым образом фальсифицированные данные. Такой прием в статистике исключается уже по одной той причине, что в получении ее данных участвует большой коллектив людей. Искажение истины получается применением «ошибочных» правил, неправильным определением объекта и т. п. Так, в США безработным считается лишь тот, кто активно ищет работу (состоит на учете) и т. п. Сразу выпадают те, кто уже отчаялся ее найти и прекратил поиски; только что вступающие в самостоятельную жизнь и еще не работавшие молодые люди и т. п. В капиталистических странах не принято учитывать как безработных тех, кто занят на полноту, и ряд других категорий. В результате, как часто отмечают профсоюзные органы, число действительно безработных в официальной статистике бывает преуменьшенным в полтора, а то и два раза.

Другой пример. Во Франции рабочий класс добился того, что заработная плата определяется с учетом роста цен, измеряемого соответствующим индексом. Вследствие этого у господствующего класса появляется двойная заинтересованность в преуменьшении упомянутого индекса как с точки зрения политической (для приукрашивания действительности), так и с точки зрения прямых материальных интересов. Коммунистическая печать не раз отмечала, что индекс цен показывает повышение цен меньше действительного. И это достигается не прямой фальсификацией, а соответствующим выбором способов наблюдения цен и обобщения его результатов в индексе.

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 3, с. 94.



Можно отметить, что излюбленным объектом апологетического искажения истины традиционно является прогноз экономики. Еще Ф. Энгельсом было отмечено, что буржуазные экономисты (и, значит, статистики) предсказывают охотно хорошую «погоду» и очень неохотно ухудшение. В наше время это полностью подтверждается. В капиталистических странах уже ряд лет продолжается кризис. Безработных, даже по преуменьшенным данным, в развитых капиталистических странах насчитывается три десятка миллионов, производство не растет. А между тем нередко высказываются прогнозы об улучшении дел в ближайшие месяцы, через полгода, в новом году и т. п. И, конечно, чем правее деятели, тем беззащитнее такие прогнозы: неслучайно их с особой настойчивостью повторяют руководители самых правых групп буржуазных политиков.

Вследствие всего этого статистика является объектом идеологической борьбы, важным участком в общей идеологической борьбе, ведущейся во всех общественных науках. Борьба против буржуазной идеологии в социально-экономической науке — будь то экономика, социология и т. п. — не может ограничиваться спорами об абстрактных понятиях, она должна доводиться до раскрытия конкретных фактов, их анализа и понимания, а это неизбежно приводит к статистике, делает ее наукой партийной.

Партийность статистики относится не только к самой констатации фактов, зависящей от принятых определений, нередко самой организации сбора информации, как в примерах, приведенных выше (о безработице, ценах и т. д.). Она ясно выступает в последующем анализе собранной информации — трактовке статистических фактов, выводах из них, даже в выработке способов математической обработки информации и, конечно, даваемых на их основании прогнозах. Ряд конкретных примеров этому будет приведен ниже по поводу тех или иных методов, применяемых в статистике.

Идеологическая борьба в статистике в настоящее время протекает не только в печати, но и в ряде организаций — статистических органах ООН, Комиссии европейских статистиков, Международном бюро труда, ЮНЕСКО, Международном статистическом институте и др. Сложность ее обусловлена тем, что в этих организациях, как, впрочем, и в государственных статистических органах стран, сложным образом переплетается воздействие разных групп — консервативных и либеральных, открытых или замаскированных угодников капиталистических монополий и честных исследователей, нередко однако, заблуждающихся вследствие отсутствия марксистско-ленинской базы мышления, представителей различных партий — от крайних правых и до коммунистов (например, во Франции и некоторых других странах). Это делает партийную борьбу в области статистики чрезвычайно сложной, требующей не только ясной

марксистско-ленинской теоретической базы, но и глубокого, детального знания самой статистики, ее теории и практики.

О партийности статистики нельзя забывать и тогда, когда речь идет о природных явлениях, тем более явлениях «на стыке» природы и общества (техника, медицина и т. п.). В конечном счете разные области явлений не отгорожены друг от друга непроницаемыми переголовками. Выше мы видели, что даже планета Венера якобы «причастна» к кризисам. Вообще говоря, не исключено, что ее движение как-то влияет на земную атмосферу (или ионосферу) и на метеорологические условия, а через это на урожай и, следовательно, на экономическую конъюнктуру. Но если б удалось правильно измерить это влияние, скажем, на выплавку стали в США, то оно измерялось бы десятком или, допустим, сотней тонн, что явно недостаточно для объяснения кризисов.

Но главное даже не всеобщая связь явлений, а то, что правильная методология в статистике возможна лишь на философской базе диалектического материализма. А приняв эту базу применительно к одним явлениям, трудно от нее отказаться применительно к другим.

Не надо представлять дело и так, что партийные цели апологетического направления всегда ясно осознаются самими учеными. Здесь часто имеет место сложное воздействие бесознательного на сознание, на самый подход к фактам. Например, К. Маркс писал о А. Кетле: «В прошлом у него большая заслуга: он доказал, что даже кажущиеся случайности общественной жизни вследствие их периодической возобновляемости и периодических средних цифр обладают внутренней необходимостью. Но объяснение этой необходимости ему *никогда* не удавалось... он только расширил материал своего наблюдения и исчисления»<sup>1</sup>. Что же мешало Кетле найти объяснение подмеченных закономерностей? Два главных обстоятельства. Как буржуазный либерал он не сумел разглядеть классы и их взаимоотношения в буржуазном обществе. Отсюда его теория «среднего человека» — внеклассового человека — вообще со всеми средними характеристиками. А без понимания классов и классовых отношений, самого устройства общества невозможно было понять источники, например, преступности, особенно приковывавшей внимание Кетле, понять, что ее сохранение — прямое следствие сохранения буржуазного общества. Второе обстоятельство — метафизический подход к явлениям, перенесенный Кетле из области физики (с ее неизменными, допустим, законами и параметрами) на человеческое общество (главный труд А. Кетле так и называется «Социальная физика»). Отсюда представления о постоянстве найденных им характеристик «среднего человека». Но самого А. Кетле нельзя упрекать в какой бы то ни было предвзятости, нарочитом искажении

<sup>1</sup> Маркс К., Энгельс Ф. Соч., т. 32, с. 495—496.

фактов. Источник его неудач — отнюдь не нежелание познать истину, а отсутствие диалектического подхода, отсутствие понимания классовой структуры общества.

#### 1.6. СТАТИСТИКА И МАТЕМАТИКА. ЗАКОН БОЛЬШИХ ЧИСЕЛ

Исследовать количественную сторону явлений — значит применять математику. Отсюда ее важная роль в статистике. Об этом не было бы надобности говорить, если бы нужная в статистике математика исчерпывалась четырьмя правилами арифметики и исчислением процентов. Если задача требует более сложных расчетов, то нет никаких оснований перед ними останавливаться. Надо лишь придерживаться правила: сложность математических методов не должна никогда и нигде быть самоцелью.

Высшая математика имеет для статистики важное значение, так как это, по определению Энгельса, математика переменных величин, в отличие от элементарной, оперирующей постоянными величинами. Следовательно, с высшей математикой неизбежно приходится иметь дело статистике при углубленном изучении динамики явления, его изменения во времени, а также взаимосвязей, о которых не может быть речи, если исключается изменение связанных величин. Обычно общее суждение о тенденции показателя к росту или снижению можно вынести и на глаз, рассматривая ряд его последовательных значений. Но если тенденция не столько ясна или если требуется ее точное измерение, то не обойтись без сложных вычислений, подбора математической функции, наилучшим образом отражающей тенденцию ряда. То же можно повторить и по поводу взаимосвязей признаков. А исследования тенденций динамики и взаимосвязи занимают в статистике очень большое место.

В ряде случаев статистическое измерение прямо опирается на то, что в последние десятилетия принято называть *математической моделью явления*. Такая модель теоретически отражает количественные соотношения изучаемого явления. При ее наличии задача статистики состоит в численном определении параметров, входящих в модели. Например, при анализе затрат, образующих общую себестоимость изделий, устанавливается, что они состоят из двух частей: затраты, пропорциональные объему производства (сырье, материалы, топливо и др.), и условно-постоянные, которые не зависят от объема производства (амортизация основных средств, управленческие расходы и др.). Следовательно, простая модель себестоимости изделия может быть представлена в виде гиперболической функции:

$y = a + \frac{b}{x}$ , где  $y$  — себестоимость изделия,  $x$  — количество изделий. Задача статистики далее состоит в нахождении двух определяющих ее параметров ( $a$  и  $b$ ), оценке степени соответ-

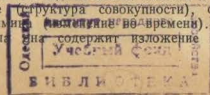
ствия полученной таким путем функции фактическим значениям себестоимости, а также в поиске и измерении действия других факторов, влияющих на себестоимость. Описанная модель при всей простоте принадлежит высшей математике. В других задачах модель может оказаться много сложнее, состоять не из одного соотношения, как в приведенном примере, а из целой их системы и т. д.

Особое значение для статистики, как уже было отмечено выше, имеет теория вероятностей и тесно связанная с ней математическая статистика. Значение для статистики *теории вероятностей* определяется характером статистической совокупности и вариации изучаемых статистикой признаков. Поскольку значение признака может оказаться любым, примеров, иллюстрирующих сказанное, можно привести сколько угодно. Покупатель вошел в булочную-кондитерскую. Нельзя заранее сказать, что именно он купит: булку, бородинский хлеб или конфеты. Нельзя заранее указать и число покупателей в течение дня; для того и другого можно лишь говорить о некотором распределении вероятностей, откуда вытекает распределение вероятностей для различных значений дневного спроса и его структуры, а это уже общая характеристика совокупности покупателей. Это станет еще яснее, если от суточных характеристик перейдем к итогам за месяц, квартал или год и не по одному магазину, а по всей их совокупности в городе. Свойства всех этих распределений имеют немаловажное значение для статистического исследования и практическое значение для организации самой торговли.

В статистике используются операции, прямым образом рассчитываемые с помощью правил теории вероятностей. Это *выборочный метод наблюдения*. Главное из этих правил — ряд теорем, выражающих *закон больших чисел*. Суть этого закона состоит в *исчезновении в сводном показателе элемента случайности, с которой связаны индивидуальные характеристики, по мере объединения в нем все большего их числа*. Так, возвращаясь к примеру с покупателями, легко понять, что как бы ни было трудно сказать, какой именно товар требуется данному покупателю, можно довольно точно предвидеть величину и структуру общего спроса совокупности покупателей, даже с учетом зависимости от дня недели и т. д.

То, что мы выразили в общей форме, в теории вероятностей получает ряд точных формулировок, зависящих от предположек, отражающих особенности тех или иных явлений. Это теоремы Я. Бернулли, Пуассона, П. Л. Чебышева, А. М. Ляпунова и др.

С теорией вероятностей тесно связана *математическая статистика*. Рассматриваемые в ней задачи можно отнести к трем категориям: распределение (структура совокупности), связи (между признаками), динамика (изменение во времени). Как математическая дисциплина она содержит изложение ряда





методов, правил, из которых исследователь применяет то, что необходимо для решения стоящей перед ним задачи. Здесь и начинается различие между приложением математической статистики в естественных науках (включая медицину, технологию и т. п.) и науках общественных.

В естественных науках или в технике главное назначение статистического показателя — как можно ближе оценить некоторый параметр, доступный наблюдению лишь с некоторой ошибкой. Например, произведено 100 измерений высоты горы. Каждое измерение связано со случайной ошибкой. Но в среднем по всей сотне измерений эти случайные ошибки погашают друг друга (точнее — почти погашают), и эту среднюю можно принять как оценку истинной высоты горы. На этом роль совокупности измерений закончена: найденная оценка высоты зафиксирована, после чего все индивидуальные результаты можно предать забвению. Иначе в общественных науках. Задача оценивания (за исключением, может быть, особых отдельных случаев) состоит не в оценивании некоторого недоступного непосредственному измерению фиксированного параметра, а в оценивании по полученным сводным данным еще более общих характеристик, причем ни эти данные, ни даже лежащие в основе их частичные и индивидуальные характеристики не теряют от этого своего реального значения. Так, по наблюдениям структуры спроса за несколько дней в той же булочной можно оценить его общую структуру, но при этом важное значение имеют и данные частичного характера — по дням недели, по часам суток и др., да и индивидуальный спрос отдельного покупателя надо знать не только для получения сводных показателей, но и для удовлетворения именно его спроса.

Статистические показатели в общественных науках имеют существенное значение не только как оценки отличных от них параметров, но и сами по себе. Это делает математическую статистику и для общественных наук ценным аппаратом. Если в естественных науках полученный результат (например, уравнение зависимости, тенденции динамики) при недостаточной пригодности для оценки искомого параметра попросту отбрасывается, то в общественной науке он представляет сам по себе интерес и требует определенных действий. Так, экономист или социолог обычно не может рассматривать полученную тем или иным способом линию общей тенденции ряда как приближенную оценку некоторой действительно существующей. Но сама эта линия, полученная в результате обработки данных наблюдения, имеет для него существенное значение.

С другой стороны, важное значение имеет вопрос о динамике социально-экономических явлений и в том, насколько возможен их прогноз путем prolongации выявленной тенденции на будущий период. А это связано с тем, насколько в ней устранены случайные колебания отдельных лет. Отсюда следует, что «пригодность» для отражения общей тенденции

важна и применительно к социально-экономическим рядам. Для суждения об этом могут быть применены те же расчеты, что и для суждения о пригодности оценок параметров действительной линии динамики. Но даже при отрицательном решении вопроса найденная тенденция сохраняет значение как обобщенное отражение фактического изменения показателя. Да и отдельные конкретные значения ряда служат не только материалом для получения этой линии, но важны и сами по себе, поскольку ими определяется ряд практических выводов. Пожалуй, это особенно ярко проявляется в отношении резко выделяющихся «рекордных» значений, которые в экономическом анализе часто представляют особый интерес.

Впрочем, если в сводных социально-экономических показателях нельзя видеть оценки недоступных наблюдению постоянных величин, то и в естествознании, при более близком рассмотрении, такая их трактовка оказывается не выдерживающей критики. Она вполне приемлема, когда колебания целиком обусловлены ошибками измерения вследствие несовершенства приборов и т. п. Но уже в исследовании живого мира — биологии — она оказывается то и дело неприемлемой. Правильно разобраться во всем этом можно только на базе материалистической диалектики необходимого и случайного.

## 1.7. СТАТИСТИКА И ПЛАНИРОВАНИЕ

Практические задачи, стоящие перед статистикой в условиях социалистического строя, вытекают из планомерности экономики. План устанавливает для народного хозяйства, для его отраслей, для предприятий и объединяющих некоторые их совокупности ведомств задания в виде системы показателей. В процессе выполнения плана необходимо оперативно наблюдать его ход. Это значит, что статистика должна доставлять информацию о выполнении плана, о влияющих на него факторах, об отклонениях от плана, как правило, требующих принятия каких-то мер. По окончании же планового периода должны быть подведены итоги выполнения плана. Все это в целом можно сформулировать коротко: статистика должна осуществлять контроль выполнения плана.

Однако задача статистики не исчерпывается констатацией того, насколько выполнен план, как фактическая величина выполнения плана отличается от заданной в плане. Сопоставляя показатели, дифференцируя их по группам объектов, дополнительно обрабатывая их (в случае необходимости с помощью сложных математических методов), статистика должна показывать, чем и как было обусловлено такое выполнение плана, дать его анализ. Далее, в этом анализе и вообще в статистической информации должен содержаться ряд показателей, которых в плане нет, — либо потому, что планирование не доводится до таких деталей, либо потому, что речь идет о принци-

пально не планируемых явлениях, например стихийные бедствия, пожары и т. п.

Статистика, доведенная до анализа факторов, оказывается часто хорошим средством для их прогноза. С этим тесно связано и то, что, контролируя выполнение планов, статистика в то же время снабжает плановые и другие органы информацией, необходимой для составления планов. Это связано и с прогнозом, так как составление планов должно начинаться по крайней мере до начала планируемого периода, а окончательные итоги за предшествующий период могут быть получены только позднее. Следовательно, незадолго до окончания предшествующего периода на основании имеющихся данных о ходе выполнения планов и их анализа необходимо делать прогноз об окончательных результатах. Так, за несколько месяцев до конца года нужен прогноз годовых итогов, для которого используется информация за прошедшую часть года и т. п.

Наконец, независимо от планов и их составления со статистикой связано решение ряда практических вопросов. Она необходима как для научного исследования, для обогащения и углубления теоретического познания явлений, так и для практических действий. Новые сложные вопросы встают перед статистикой в деле освещения роли интенсивного и экстенсивного факторов роста, что стало особенно актуальным на современном этапе развития социалистического хозяйства. Основой перехода к интенсификации производства служат научно-технический прогресс, анализ которого с помощью целой системы количественных характеристик также является важной задачей статистики. Все большее значение наряду с чисто экономическими характеристиками развития приобретают социальные, что требует основательного расширения и улучшения соответствующих разделов статистики. Без ее данных трудно представить себе составление и выполнение любой программы — производственной, энергетической, повышения жизненного уровня народа и т. д.

В общей теории статистики рассматриваются в общем виде методы сбора, сводки, обработки и анализа статистической информации. Конкретные показатели — их система, способы исчисления и информационное обеспечение рассматриваются в специальных статистических дисциплинах. Они могут быть более широкими, как экономическая статистика, социальная или социально-экономическая статистика, либо более узкими, как статистика промышленности, статистика населения, сельскохозяйственная статистика и т. д., и даже еще более узкими, как, например, статистика текстильной промышленности, статистика животноводства и т. п.

Сами статистические результаты — численные значения показателей — не являются объектом преподавания, но обычно используются в нем в качестве примеров. Они являются также

объектом статистических публикаций — ежегодников, справочников, регулярно публикуемых сообщений ЦСУ СССР, статистических управлений республик, краев, областей и др.

## ГЛАВА 2 ОРГАНИЗАЦИЯ СТАТИСТИКИ

### 2.1. УЧЕТ И СТАТИСТИКА В СОЦИАЛИСТИЧЕСКОМ ОБЩЕСТВЕ

Если спросить любого человека, что такое учет, то его удивит простота и наивность вопроса. Ведь учетом мы занимаемся постоянно и систематически: учтем время, вещи и предметы, считаем заработанные деньги и затраты на покупки. На предприятии или в учреждении ведется учет выпускаемой продукции, отработанного времени и его потерь, выработки и заработка и т. п. Таким образом, учет — один из неотъемлемых видов деятельности человека.

Развитие общества представляет собой непрерывный процесс производства и потребления материальных благ. Этот процесс отражается в разнообразных, взаимосвязанных и тесно переплетающихся экономических, социальных, политических, демографических, моральных и других явлениях, событиях и фактах, посящих как массовый, так и индивидуальный характер. Без систематического анализа этих явлений, их последствий и направлений дальнейшего развития невозможно существование всякого организованного общества, эффективное ведение хозяйства, удовлетворение постоянно растущих потребностей людей. Базой, основой такого анализа являются сведения, данные систематического и всестороннего учета. Следовательно, речь идет о систематическом измерении и изучении разнообразных социальных и экономических явлений и процессов на основе количественного подсчета отдельных элементов, отдельных единиц, из которых состоят эти явления. В этом и заключается сущность учета.

Итак, *учет* — способ непрерывной, систематической регистрации и измерения различных социальных и экономических процессов, явлений, фактов общественного развития, каждый элемент, каждая единица которых обязательно подсчитывается, т. е. имеет определенную количественную характеристику. В зависимости от поставленных перед данным видом учета задач, а также от масштаба и степени сложности самого объекта учета для его количественной характеристики применяются самые разнообразные способы подсчета — от простого элементарного арифметического счета до сложных методов высшей математики. Например, не представляет особой трудности подсчитать наличные трудовые ресурсы предприятия на

определенную дату или наличие уборочной техники агропромышленного комплекса района, ее состав и степень пригодности к работе. Когда же возникает необходимость в определении сложных экономических категорий, например объемов выпускаемой продукции, национального богатства общества, капитальных вложений и др., которые нужно охарактеризовать не только в натуральном, но и в стоимостном выражении, простой, количественный счет отдельных элементов или единиц, входящих в состав этих категорий, не пригоден. В учете тогда используются различные методы пересчета (переводные коэффициенты, индексы и др.) для обеспечения правильного определения величины данного объекта учета.

Важнейшими *требованиями* к организации и ведению учета являются: организация его систематичности и непрерывности; обеспечение объективности и достоверности данных о наличии, состоянии и развитии данного объекта или явления; характеристика объекта учета как на определенный момент времени, так и его изменений за определенный период, т. е. в динамике.

Состояние и развитие учета отражают уровень развития производительных сил и производственных отношений данной общественно-экономической формации. Эту зависимость учета от характера способа производства впервые сформулировал К. Маркс в «Капитале», определив учет как средство контроля и мысленного обобщения процесса производства. Он писал: «...учет как средство контроля и мысленного обобщения этого процесса становится тем необходимым, чем более процесс производства совершается в общественном масштабе и утрачивает чисто индивидуальный характер...»<sup>1</sup>

При капиталистическом способе производства с присущим ему антагонистическим противоречием между общественным характером производства и частнокапиталистической формой присвоения его результатов учет носит ограниченный и классовый характер: он обычно ведется в рамках отдельных предприятий и монополий и не может быть организован как всеобъемлющий общественный (народнохозяйственный) учет; он ведется в интересах господствующего класса капиталистов и поэтому играет апологетическую роль, всячески скрывая и затухивая пороки капиталистического способа производства. Известно, к каким фальсификациям данных учета прибегают капиталистические монополии, чтобы скрыть свои прибыли и сверхприбыли, охраняя «коммерческие» и «торговые» тайны, а попросту мошеннические сделки и спекулятивные махинации. Особенно изощряются буржуазные экономисты в фальсификации данных об уровне жизни трудящихся масс, росте безработицы и обнищании беднейших слоев населения.

Только социалистический способ производства создает объективные предпосылки для организации и ведения подлинно

общественного, общенародного учета, превращения его в единую народнохозяйственную систему.

После победы Великой Октябрьской социалистической революции партия большевиков во главе с великим Лениным осуществляет титаническую работу по созданию новых социалистических принципов организации производства и дисциплины труда, новых форм, функций и содержания работы государственного аппарата в целом и отдельных его звеньев. В результате этой огромной работы была создана единая система социалистического учета и статистики как важнейшая база планового ведения народного хозяйства и способа контроля за производством и распределением продуктов.

Известно, что буквально через несколько дней после победы Великого Октября В. И. Ленин с гениальной прозорливостью провозглашает: «Социализм — это учет», а затем в целом ряде своих блестяще аргументированных работ разрабатывает основополагающие теоретические и методологические принципы организации и ведения учета при социализме. К этим работам следует прежде всего отнести «Государство и революция», «Очередные задачи Советской власти», «Как организовать соревнование?». Обосновывая объективную необходимость создания единой системы учета и контроля при социализме, ее неразрывную связь с самой природой социалистического способа производства, В. И. Ленин писал: «...учет и контроль повсеместный, всеобщий, универсальный, — учет и контроль за количеством труда и за распределением продуктов — в этом суть социалистического преобразования, раз политическое господство пролетариата создано и обеспечено»<sup>1</sup>.

Руководствуясь этими ленинскими указаниями, Коммунистическая партия и Советское государство обеспечили постоянное совершенствование единой системы учета и статистики в соответствии с особенностями каждого этапа развития социалистической экономики. В настоящее время неизмеримо возрастает значение дальнейшего совершенствования всех инструментов государственного управления и планового руководства народным хозяйством, одним из которых является единая система народнохозяйственного учета и статистики. Эта система предусматривает неразрывную органическую взаимосвязь трех видов учета: бухгалтерского, оперативно-технического и статистического. Эта взаимосвязь выражается в диалектическом единстве всех трех видов учета, которые характеризуются общими и отличительными чертами, определяющими их организационные, методологические принципы и практическую значимость. *Общими чертами всех трех видов народнохозяйственного учета являются:*

— единая исходная информационная база в виде различных данных первичного учета (первичная, входная или исход-

<sup>1</sup> Маркс К., Энгельс Ф. Соч., т. 24, с. 153.

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 35, с. 199—200.



ная информация). Эти данные представляют собой систему записей, фиксирующих в соответствующих единицах измерения каждое явление, событие, процесс, происшедший в определенный момент или за определенный период времени. Они содержатся в самых разнообразных документах первичного учета: ведомостях, табелях, нарядах, актах, справках, рапортах, накладных, табелях, перфокартах, магнитных лентах и т.д. Перечень форм и содержание этих документов по всем видам производственной, хозяйственной и финансовой деятельности предприятий, организаций и учреждений определяется «Альбомом типовых междоведомственных форм первичной учетной документации», который разрабатывается, периодически пересматривается, обновляется и утверждается Центральным статистическим управлением СССР и Министерством финансов СССР. Документы первичного учета являются необходимой документальной базой всех видов учета и статистической отчетности, и поэтому от своевременности, точности, полноты и достоверности сбора и обработки данных во многом зависит качество статистической отчетности и действительность экономико-статистического анализа;

— единые принципы организации всех видов народнохозяйственного учета и единое централизованное руководство их развитием. Высшим государственным органом, осуществляющим централизованное руководство всем делом учета и статистики в стране, является Центральное статистическое управление СССР (ЦСУ СССР), местные органы которого имеются в каждой единице административно-территориального деления СССР. Формы документов статистического учета разрабатываются и утверждаются ЦСУ СССР, бухгалтерского учета — ЦСУ СССР и Минфин СССР, оперативно-технического учета — соответствующими министерствами и ведомствами по согласованию с ЦСУ СССР или союзных республик. Таким образом, статистическими органами обеспечиваются и контролируются сбор и обработка всех видов учетных данных на всех уровнях народного хозяйства. Они устанавливают содержание форм учета, сроки и адреса их представления. Такой единый централизованный принцип организации и руководства учетом определяет роль и значение статистических органов как составной части системы управления и планирования социалистического государства;

— методологическое единство всех видов учета. Это выражается в том, что все виды учетных документов и содержащиеся в них данные собираются и обрабатываются по единому плану и программе, в которых предусматриваются применение одинаковых единиц измерения, четкое определение количества и состава объектов учета, сопоставимости по кругу объектов, территории, периодам времени, применение одинаковых стоимостных оценок и т.д.

— аналитичность материалов учета. Это означает, что со-

бранные и правильно методически обработанные данные всех видов учета должны найти широкое применение в экономико-статистическом анализе факторов и резервов повышения эффективности производства, анализе выполнения планов социально-экономического развития страны, повышения материального и культурного уровня жизни народа. Этим и определяются общность и единство задач всех трех видов учета.

Вместе с тем *каждому виду народнохозяйственного учета* присущи свои *отличительные черты*, которые могут быть охарактеризованы следующим образом.

*Бухгалтерский учет* представляет собой сплошную, непрерывную регистрацию наличия и движения всех материальных и денежных средств предприятия, которая фиксирует каждый факт поступления (с указанием источника) и использования (с указанием направления) каждого вида этих средств. Как правило, в бухгалтерском учете отражаются лишь те предметы и процессы, которые имеют денежное выражение: данные обрабатываются на строго документальной основе присущия бухгалтерскому учету способами (калькулирование, двойная запись регистрируемых операций, балансовые подсчеты и т.д.).

Документы бухгалтерского учета имеют не только учетно-статистическое, но и правовое, юридическое значение. Это выражается в том, что каждый документ в обязательном порядке подписывается ответственным исполнителем и руководителем соответствующего подразделения, а в целом по предприятию, организации, учреждению — его руководителем и главным бухгалтером. Эти лица несут полную ответственность за своевременность, достоверность и качество составляемых документов бухгалтерского учета. Поэтому задачей бухгалтерского учета является также выполнение функций контроля за соблюдением требований хозяйственного расчета, строжайшей экономией средств, охраной общественной социалистической собственности.

*Оперативно-технический учет* представляет собой регистрацию отдельных событий и фактов непосредственно в момент их совершения или проявления на рабочих местах и участках (отделах) предприятий и организаций, в которой отражаются данные о состоянии производственной или хозяйственной деятельности. Эти данные называются оперативными, срочными, так как они получены сразу же, как только в них возникает потребность, и используются руководством предприятия или организации для быстрого, т.е. оперативного, реагирования на обнаруженные неполадки и принятия срочных мер по их устранению. Они называются техническими, так как отражают главным образом технологические, организационные, управленческие и экономические процессы и явления в соответствующих натуральных или трудовых измерителях.

Например, в капитальном строительстве, данные оперативно-технического учета представляют собой информацию о соблюдении недельно-суботных и сетевых графиков производства работ, об отклонениях от установленных нормативов расхода различных видов строительных материалов, комплектности и равномерности их поставок, об использовании строительных машин и механизмов по времени и мощности и т.д.

Наконец, третий вид народнохозяйственного учета — *статистический* содержит информацию о массовых социальных, экономических и других общественных явлениях и процессах. На его основе разрабатываются обобщающие количественные и качественные показатели развития всех сторон и форм социального воспроизводства в их единстве и взаимосвязи. Поэтому этот вид учета по своим организационным и методологическим принципам, аналитическому уровню и практической значимости является наиболее сложным и многообразным, играет ведущую и определяющую роль во всей системе народнохозяйственного учета. Статистический учет, или просто статистика, — составная часть системы государственного управления, контроля и планирования народного хозяйства и культуры страны. Обработывая и обобщая данные бухгалтерского и оперативно-технического учетов, статистика обеспечивает всестороннюю характеристику выполнения народнохозяйственных планов, анализирует факторы и резервы повышения эффективности производства, представляет необходимые материалы для перспективного планирования и прогнозирования экономического и социального развития страны. С этой целью составляется и разрабатывается государственная статистическая отчетность, проводятся периодические и единовременные выборочные статистические и социологические обследования (перепись, учет, инвентаризация, изучение общественного мнения и др.), обеспечиваются разработка и анализ единой системы статистических показателей, всесторонне характеризующих процесс расширения социального воспроизводства. Социально-экономическая статистика, по определению В. И. Ленина, — одно из самых могущественных орудий социального познания.

В материалах июньского (1983 г.) и апрельского (1985 г.) Пленумов ЦК КПСС подчеркивается необходимость принятия решительных мер по обеспечению хорошо отлаженной, бесперебойной работы всего хозяйственного механизма, коренному улучшению планирования и управления. Значительный вклад в успешное решение этих первостепенных, программных задач должна внести вся система государственного учета и статистики. Она призвана поднять на новый качественный уровень всю работу по обеспечению своевременности, качества и достоверности информации, ее оперативной обработке, сводке и анализу; представлению своевременных и полных материалов для текущего и перспективного планирования народного хозяйства; анализу пропорций и соотношений в развитии отдельных отраслей и предупреждению нарушений этой пропорциональности. Значительная работа должна быть выполнена в области дальнейшего совершенствования системы плановых и отчетных, отраслевых и народнохозяйственных показателей, улучшения методологии их определения и анализа. Вся работа статистических органов должна быть направлена на оказание практической помощи руководящим и плановым органам в выявлении

и использовании внутренних резервов повышения эффективности производства, безусловного выполнения планов экономического и социального развития страны.

## 2.2. ОРГАНЫ ГОСУДАРСТВЕННОЙ СТАТИСТИКИ, ИХ ФУНКЦИИ И СТРУКТУРА

Система государственной статистики в СССР организована в соответствии с административно-территориальным делением страны и имеет иерархическую структуру, включающую следующие уровни: союзный, республиканский, краевой, областной, городской, районный.

Централизованное руководство всей системой учета и статистики, работой органов государственной статистики в стране осуществляет Центральное статистическое управление СССР (ЦСУ СССР). В союзных республиках аналогичные функции выполняют ЦСУ союзных республик, которые подчиняются ЦСУ СССР. В автономных республиках, краях, областях и национальных округах имеются статистические управления, а в административных районах и городах — районные и городские инспектуры государственной статистики или районные (городские) информационно-вычислительные станции (Р(Г)ИВСы).

На всех уровнях система государственной статистики располагает развитой сетью вычислительных центров (ВЦ) и машинносетевых станций (МСС), которые образуют вычислительную систему ЦСУ СССР и предназначены для обработки статистических данных в своей системе, а также выполнения учетно-вычислительных работ по заказам предприятий, организаций, колхозов и совхозов. Руководство сетью ВЦ и МСС осуществляет Всесоюзное объединение по информационно-вычислительному обслуживанию ЦСУ СССР (Союзмашинформ).

ЦСУ СССР и органы государственной статистики в АССР, краях, областях, городах и районах образуют единую систему, функционирующую на основе общих принципов, единой методологии и организации работ.

Для разработки теоретических и методологических принципов дальнейшего совершенствования статистики и учета, применения экономико-математических методов и ЭВМ в статистических исследованиях при ЦСУ СССР организован Научно-исследовательский институт статистики (НИИ ЦСУ СССР). С целью проектирования механизации и автоматизации учетно-статистических и других информационно-вычислительных работ при ЦСУ СССР был создан Всесоюзный государственный проектно-технологический институт (ВГПТИ), реорганизованный во Всесоюзный научно-исследовательский и проектно-технологический институт по механизации учета и отчетности в народном хозяйстве (ВНИПУчет ЦСУ СССР). При ЦСУ СССР есть Главное управление подготовки и повышения квалификации работников учета (ГУПК ЦСУ СССР).

ЦСУ СССР является союзно-республиканским органом. В его функции входят:



— разработка и дальнейшее совершенствование методологии советской статистики, системы статистических показателей, организация учета и статистики;

— сбор, контроль, обработка и своевременное представление Правительству СССР и местным руководящим органам, а также министерствам и ведомствам научно обоснованных статистических данных, характеризующих ход выполнения народнохозяйственных планов, эффективность общественного производства, научно-технический прогресс и т. д.;

— экономический анализ статистических данных, составление отчетного баланса народного хозяйства, межотраслевого баланса, балансов материальных, финансовых, трудовых ресурсов, перспективный расчет численности населения, выполнение других статистико-аналитических расчетов;

— обеспечение развития механизации и автоматизации информационно-вычислительных работ в народном хозяйстве, участие в разработке и создании Общегосударственной автоматизированной системы сбора и обработки информации для учета, планирования и управления народным хозяйством (ОГАС);

— контроль за состоянием учета и отчетности, достоверностью отчетных данных во всех отраслях народного хозяйства;

— публикация в печати сообщений о выполнении государственных планов экономического и социального развития СССР, издание статистических бюллетеней.

Возглавляет ЦСУ СССР начальник, входящий, как и другие руководители министерств и ведомств, в состав Правительства СССР, утверждаемый Сессией Верховного Совета СССР. Начальник ЦСУ СССР возглавляет коллегию, состоящую из руководящих работников центрального аппарата и начальников ЦСУ союзных республик. Коллегия на своих заседаниях рассматривает важнейшие вопросы методологии, организации и автоматизации статистических работ, итоги основных выполненных статистических работ, заслушивает отчеты руководящих работников ЦСУ СССР, ЦСУ союзных республик и местных статистических органов.

Для обсуждения теоретических и методологических вопросов развития статистики, программ и инструкций по важнейшим статистическим работам при ЦСУ СССР образован Научно-методологический совет. В его состав, кроме руководителей основных подразделений ЦСУ СССР, входят известные ученые и специалисты в области статистики. Состав совета утверждает начальник ЦСУ СССР.

Непосредственное выполнение функций, возложенных на ЦСУ СССР, осуществляется рядом управлений, отделов и других структурных подразделений, организованных в основном по отраслевому признаку. Так, имеются управления статистики промышленности, статистики сельского хозяйства, статистики капитального строительства, статистики материально-техниче-

ского снабжения и переписей, статистики труда и заработной платы, статистики зарубежных стран, Всесоюзной переписи населения, баланса народного хозяйства; отделы сводной статистики и статистической методологии, статистики природных ресурсов и окружающей среды, статистики транспорта и связи, статистики торговли, статистики культуры и др. Внутри крупных управлений функционируют отделы. Так, в рамках управления статистики промышленности имеются отделы: сводно-экономический, статистики основных фондов и производственных мощностей, статистики отраслей тяжелой промышленности и др.

ЦСУ союзных республик и статистические управления АССР, краев, областей, городов осуществляют на своей территории руководство делом учета и статистики и те же функции, что и ЦСУ СССР.

Районное (городское) звено системы государственной статистики в настоящее время представлено районными (городскими) информационно-вычислительными станциями (РИВС и ГИВС). Они образовались в результате объединения районных (городских) инспектур государственной статистики с районными (городскими) МСС системы ЦСУ СССР. В последнее время РИВСы (ГИВСы) оснащаются ЭВМ и в связи с этим преобразуются в районные (городские) информационно-вычислительные центры (РИВЦы, ГИВЦы). На Р(Г)ИВСы и Р(Г)ИВЦы возложены не только функции, связанные с руководством учетом и статистикой на территории района (города), но и обслуживанием вычислительными работами всех предприятий, организаций, колхозов и совхозов, находящихся на территории данного района (города).

Работа органов государственной статистики осуществляется по заранее разрабатываемым планам. ЦСУ СССР ежегодно составляет «План статистических работ Центрального статистического управления», который, по существу, является основополагающим документом, регламентирующим работу всех статистических органов страны. По каждой работе, включенной в этот план, устанавливаются периодичность выполнения, сроки и способ представления выходных и промежуточных данных, характер разработки, информационная база. На его основе ЦСУ союзных республик составляют свои планы, которые в свою очередь используются для разработки планов областными (АССР, краевыми) органами статистики. Последние берут за основу при составлении планов работы Р(Г)ИВСов и Р(Г)ИВЦов.

Наряду с органами государственной статистики статистическая работа осуществляется промышленными и сельскохозяйственными предприятиями, организациями, стройками, министерствами и ведомствами, государственными комитетами, где имеются специальные службы (отделы, группы или отдельные работники), которые ведут статистическую работу в масштабе

своих предприятий, организаций и учреждений. Эти службы и составляют систему *автоматизированной статистики*. Они собирают и обрабатывают данные, необходимые для оперативного руководства, и находятся в двойном подчинении: в административном отношении подчиняются своим ведомствам, а в методологическом — органам ЦСУ СССР. Ведомственная статистика играет важную роль в системе государственной статистики страны и во многих случаях является источником статистических данных для общегосударственной статистики.

При современных огромных масштабах общественного производства, постоянном расширении и усложнении межотраслевых связей, высоком уровне развития общественного разделения труда и возрастающих темпах научно-технической революции управление народным хозяйством невозможно без четко поставленной и научно организованной статистической информации. По подсчетам НИИ ЦСУ СССР общий объем всей статистической информации, обрабатываемой в системе ЦСУ СССР, составляет в расчете на год около 148 млрд. знаков, или свыше 35 млрд. показателей. Совершенно очевидно, что в этих условиях нельзя обойтись без широкой и всесторонней механизации сбора и обработки статистической информации.

Еще в директивах XXIV съезда КПСС была поставлена задача создания единой общегосударственной автоматизированной системы сбора и обработки информации для учета, планирования и управления народным хозяйством на базе государственной сети вычислительных центров и единой автоматизированной сети связи страны. Выполнение этой задачи в области государственной статистики нашло свое выражение в создании автоматизированной системы государственной статистики (АСГС), явившейся одним из функциональных звеньев общегосударственной автоматизированной системы сбора и обработки информации.

*Автоматизированная система государственной статистики* (АСГС) является сложной и развивающейся системой. Поэтому она создается поэтапно в виде отдельных очередей. Работы по созданию АСГС были начаты в 1971 г. В конце 1975 г. была сдана в промышленную эксплуатацию первая очередь АСГС, а в конце 1980 г. — вторая. В одиннадцатой пятилетке должна быть разработана и внедрена третья очередь АСГС.

Каждая из очередей АСГС решала свои задачи по развитию и совершенствованию государственной статистики. *Первая очередь* обеспечила переход на ЭВМ наиболее важных и трудоемких статистических работ. Организационно-технологической формой автоматизированных решений задач по обработке статистических данных явились комплексы электронной обработки информации (КЭОИ). Внедрение КЭОИ позволило объединить все ВЦ системы ЦСУ СССР в единую вычислительную сеть и обеспечить этим высокие скорости обработки и передачи данных в системе. В результате значительно ускорилась разработка статистической отчетности и увеличилась доля рабочего времени статистиков, используемого для аналитических целей.

*Вторая очередь* АСГС являлась дальнейшим развитием первой. Ее основными задачами были расширить объем статистической информации, подлежащей автоматизированной обработке, повысить уровень механизации и автоматизации ее сбора и обработки и на основе этого усовершенствовать методологию и организацию государственной статистики. Особенность второй очереди состояла в том, что она создавалась и развивалась на более совершенной технической базе — ЭВМ третьего поколения, что позволяло по-новому решать вопросы организации статистической информации. Качественно новым в этом плане являлось создание автоматизированных банков данных (АБД).

АБД представляет собой интегрированный программно-технологический комплекс, осуществляющий накопление, хранение и постоянное обновление статистической информации в целях ее использования для оперативного решения экономико-статистических задач, выполняемых органами государственной статистики. Итерация данных в АБД производится в однократной фиксации каждого показателя, многократным и многоцелевым его использованием при решении статистических задач. Создание АБД открывает широкие возможности для наиболее полного и эффективного использования информационного фонда государственной статистики, повышения уровня экономического анализа данных, комплексного анализа экономических и социальных процессов в народном хозяйстве с широким применением экономико-математических методов и ЭВМ. К концу десятой пятилетки в системе ЦСУ СССР были созданы четыре АБД.

В результате создания второй очереди АСГС в 1980 г. общий объем статистической информации, обрабатываемой ЭВМ, на республиканском и областном уровнях достиг в среднем 45—48 %, а на союзном — около 60 % всего объема информации. Совершенствование методов обработки информации и эффективное использование технической базы АСГС позволили за десятую пятилетку условно высвободить более 40 тыс. человек.

Задачей *третьей очереди* АСГС является дальнейшее развитие функциональных подсистем на основе внедрения автоматизированных банков данных и системы телеобработки статистической информации (СТОСИ), которая предусматривает автоматическую передачу данных от источников информации в вычислительные центры.

С внедрением АСГС не только рационализируется процесс сбора и обработки статистической информации, но и совершенствуется сама система статистических показателей, методология их анализа. Объединение задач в комплексы и создание единой информационной базы обеспечили возможность разработки новых выходных таблиц с важной аналитической информацией без увеличения исходных данных. Наличие АБД расширило возможность использования комбинационных группировок, методов корреляции, факторного анализа, повысило аналитичность и сопоставимость статистических данных, позволило полнее обеспечить потребности органов управления и планирования в статистической информации и необходимых справочных материалах.

## 2.3. МЕЖДУНАРОДНАЯ СТАТИСТИКА

Современный этап мирового развития, характеризующийся дальнейшим углублением общего кризиса капитализма, обострением борьбы двух общественно-политических систем, ростом и укреплением мировой социалистической системы, обусловил значительное развитие международной статистики —



одной из составных частей социально-экономической статистики.

*Международная статистика* охватывает статистику стран социалистического содружества и статистику капиталистических стран. Организационными формами развития ее теории, методологии и практики, кроме соответствующих государственных статистических органов и служб в каждой стране, являются органы статистики в рамках СЭВ и различные международные статистические организации.

Во всех странах социалистического содружества за сравнительно короткий период были успешно решены организационные и методологические проблемы статистической теории и практики. Это стало возможным благодаря принятию ленинских принципов научной организации статистики, использованию большого теоретического и методологического опыта советской статистики и огромной практической помощи статистиков СССР.

Для всех социалистических государств характерны строгая централизация статистической службы как важнейшего звена государственного управления, научно организованная база первичного учета, методологическое единство построения и анализа плановых и статистических показателей. Единство целей, задач и содержания работы статистических органов обусловило их принципиально одинаковую организацию и структуру, весьма близкие к органам ЦСУ СССР.

Углубление международного социалистического разделения труда, возникновение и развитие экономического сотрудничества между странами социализма, превратившееся в процесс экономической интеграции, объективно потребовали создания единой системы статистической информации, разработки единых теоретических и методологических принципов и положений для всесторонней характеристики и анализа этих процессов. Так возник новый раздел статистической науки — *международная статистика мировой социалистической системы*. Ее цели, задачи и содержание могут быть сформулированы следующим образом:

- разработка теоретических и методологических проблем обеспечения сопоставимости национальных статистических данных и внедрение соответствующих программ и методов в практику;

- создание и постоянное совершенствование единой системы организации и методологии статистики на основе унификации и стандартизации статистических и плановых показателей, единиц измерения, номенклатур, классификаций, методик их определения и анализа;

- создание и постоянное совершенствование единой системы механизированной и автоматизированной обработки учетной и статистической информации на основе широкого применения новейших средств вычислительной техники и ЭВМ;
- всесторонняя характеристика хода выполнения государственных планов социально-экономического развития каждой страны и в целом стран социалистической системы, динамики развития отдельных отраслей экономики и культуры, повышение жизненного уровня населения стран социализма;

- систематический взаимный обмен статистической информацией и новейшими методологическими материалами и методическими разработками.

Успешное решение этих задач обеспечивается применением всеми социалистическими странами единой марксистско-ленинской научной методологии статистики, эффективными формами организации сотрудничества, координации и взаимного обмена опытом работы.

С созданием в 1949 г. Совета Экономической Взаимопомощи (СЭВ) появились первые формы взаимного общения и координации деятельности в области статистики в виде совещаний начальников статистических управлений стран-членов СЭВ, обмена специалистами, создания рабочих групп и комиссий по отдельным проблемам и отраслям статистики.

Была создана Постоянная комиссия СЭВ по статистике, а в составе Секретариата СЭВ организован отдел статистики.

Усилиями статистиков всех социалистических стран был решен ряд важнейших научных методологических и практических проблем, не имевших аналогов в статистической теории и практике. К этим проблемам, в частности, относятся:

- унификация показателей международного социалистического разделения труда, что позволило значительно улучшить сопоставимость статистических данных и облегчить работу по координации народнохозяйственных планов в рамках СЭВ и долгосрочному прогнозированию развития каждой страны и всей социалистической системы хозяйства (разработка и внедрение единых классификаторов, стандартов, методик, научных принципов выбора и обоснования товаров-представителей и т. д.);

- создание единой автоматизированной системы сбора, обработки, хранения и передачи статистической информации (вопросы программного обеспечения, единых принципов организации и работы ЭВМ и ВЦ коллективного пользования, создания и развития автоматизированного банка данных и т. д.);

- разработка методологических положений и системы показателей по отдельным отраслям статистики.

Теоретические, методологические проблемы, практические вопросы статистики широко освещаются и анализируются в различных изданиях и публикациях СЭВ: Информационном бюллетене по статистике; Информационном бюллетене Секретариата СЭВ «Экономическое сотрудничество стран-членов СЭВ»; статистическом ежегоднике СЭВ.

*Статистика капиталистических стран* по форме своей организации, методологии и содержанию практической деятельности целиком и полностью отражает характерные черты капиталистического способа производства на современном этапе его развития, выступая одновременно в роли защитника капиталистического строя. Используя все многообразие приемов и способов статистической методологии, буржуазные статистики стремятся более утонченно фальсифицировать статистические данные для приукрашивания капиталистического действительности, сокрытия острых экономических потрясений и социальных

конфликтов. В этом основная суть, содержательная позиция всей статистики капиталистических стран, каких бы проблем и вопросов анализа экономических и социальных процессов она ни касалась.

*Международные статистические организации* — это статистические органы Организации Объединенных Наций (ООН), основным из которых является Статистическая комиссия. Статистическая комиссия разрабатывает вопросы методологии, сопоставимости показателей, рекомендации для Статистического бюро Секретариата ООН. Статистическая комиссия координирует статистическую работу специализированных органов ООН, консультирует их по вопросам сбора, анализа и распространения статистических данных. Специализированные органы ООН периодически представляют комиссии доклады о своей деятельности в области статистики. Статистическая комиссия в свою очередь регулярно представляет Экономическому и Социальному Совету ООН доклады о своей деятельности.

Исполнительным органом в области статистики является Статистическое бюро Секретариата ООН. Оно собирает и публикует данные, полученные как от статистических органов государств — членов ООН, так и от специализированных органов ООН, подготавливает доклады по различным вопросам статистики, ведет разработку методологических вопросов.

Кроме того, Статистическое бюро периодически публикует обзоры о переписях и выборочных обследованиях, проведенных в различных странах, составляет доклады по международным статистическим стандартам, исчисляет прогнозы численности населения мира и выполняет другие работы. Бюро имеет ряд периодических изданий, как, например, «Ежемесячный статистический бюллетень», «Демографический ежегодник», «Ежегодник по внешней торговле» и другие.

Вопросами демографии и статистики населения занимается Комиссия по народонаселению Экономического и Социального Совета ООН. Всесторонне занимаются вопросами статистической методологии и анализа ряд международных специализированных органов ООН: Международная организация труда (МОТ), Международная продовольственная и сельскохозяйственная организация (ПСХО), ЮНЕСКО, Всемирная организация здравоохранения (ВОЗ) и др.

Вопросы статистики рассматриваются также региональными экономическими комиссиями для Европы, Азии и Дальнего Востока, Латинской Америки, Африки, подчиненными Экономическому и Социальному Совету ООН. При Европейской экономической комиссии и Экономической комиссии для Азии и Дальнего Востока функционируют постоянные конференции статистиков. Конференция европейских статистиков была создана в 1953 г. в качестве постоянного органа.

Международным статистическим органом является также Международный статистический институт (МСИ). МСИ ведет

работу по обобщению научных исследований в области теории и методологии статистики, обсуждению актуальных проблем народонаселения, охраны окружающей среды, статистическому анализу экономики и культуры развивающихся стран, жизненного уровня населения мира. МСИ проявляет заботу о популяризации статистики, расширении статистического образования и подготовке национальных кадров по статистике особенно для развивающихся стран Азии и Африки. С этой целью созданы два международных центра статистического образования в Калькутте и Бейруте.

Советские статистики и статистики других социалистических стран активно участвуют в работе международных статистических органов. В работе сессии МСИ принимали активное участие видные советские статистики: член-корреспондент АН СССР, профессор Т. В. Рябушкин, профессор А. Я. Боярский, профессор М. А. Королев и др.

## ГЛАВА 3 СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

3.1. СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ,  
ЕГО СОДЕРЖАНИЕ И ЗАДАЧИ

Характеристика количественной стороны любого экономического или социального процесса, явления или факта, его качественной сущности, места, роли и взаимосвязей в общей системе общественного производства невозможна без всестороннего и глубокого исследования, а именно статистического исследования. Прежде чем приступить к использованию различных способов и приемов статистической методологии, необходимо располагать исчерпывающими и достоверными данными об объекте исследования, каким бы сложным и многогранным он ни был. А для этого нужно знать цели и задачи всего статистического исследования, знать, какого рода выводы можно сделать на основе анализа имеющихся данных, каковы их теоретическая и методологическая ценность и практическая значимость. Таким образом, *организация статистического исследования* предполагает предварительную всестороннюю и тщательно продуманную разработку вопросов его подготовки и проведения, а также анализа полученных результатов.

Статистическое исследование заключается в следующих этапах:

- сбор статистической информации (статистическое наблюдение) и ее 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.

Как видим, *первым и исходным этапом статистического исследования является статистическое наблюдение*. Именно в процессе наблюдения формируются первичные статистические данные или исходная статистическая информация, представляющая собой как бы фундамент будущего статистического знания. Чтобы это знание было прочным, его основа должна быть добротной и высококачественной. Если при сборе первичных статистических данных допущена ошибка, если собран недоброкачественный статистический материал, то как бы его

затем хорошо ни сводить и обрабатывать, рано или поздно этот брак скажется и повлияет на правильность, достоверность как теоретических, так и практических выводов<sup>1</sup>. Следовательно, являясь основой всего статистического исследования, в значительной степени определяющей его результаты и успешность проведения, статистическое наблюдение от начальной до завершающей стадии получения итоговых материалов должно быть всесторонне продуманным и четко организованным в полном соответствии с требованиями статистической теории и методологии. Это означает, что статистическое наблюдение должно быть организовано как планомерное, массовое и систематическое.

*Планомерность* статистического наблюдения заключается в том, что оно подготавливается и осуществляется по заранее подробно разработанному плану, который охватывает все вопросы методологии, организации и техники сбора статистической информации, контроля ее качества и достоверности, оформления итоговых материалов для дальнейшей их сводки и обработки. Следует подчеркнуть, что *план статистического наблюдения* должен представлять собой неразрывную часть плана всего статистического исследования.

*Массовый характер* статистического наблюдения означает, что оно организовано и направлено на охват возможно большего, массового числа случаев проявления данного явления или процесса, на получение правдивых статистических данных для характеристики не только отдельной единицы изучаемой совокупности, но — самое главное — массы единиц совокупности. Объективная и всесторонняя характеристика современных сложных экономических и социальных явлений в жизни общества может быть дана лишь на основе сбора разнообразных массовых данных, являющихся совершенно необходимым условием научно обоснованных плановых решений, аналитических и прогнозных расчетов, выводов и рекомендаций.

*Систематичность* статистического наблюдения определяется тем, что оно должно производиться не стихийно, от случая к случаю, а систематически, либо непрерывно, либо регулярно, по возможности через равные промежутки времени. Ведь все экономические и социальные явления динамичны, они протекают в постоянном развитии, характеризуются непрерывными количественными и качественными изменениями, структурными сдвигами и усложняющимися соотношениями и взаимосвязями. Изучение тенденций, направлений и закономерностей этого развития, происходящего под влиянием действия самых разнообразных обстоятельств и факторов, а также прогнозирование возможно и необходимо лишь на основе систематического наблюдения. Таким образом, статистическое наблюдение пред-

<sup>1</sup> Вопросы обеспечения достоверности данных и высокого качества статистической информации рассматриваются в гл. 4.



ставляет собой научно организованный, планомерный и систематический процесс сбора массовых данных о различных экономических и социальных процессах, явлениях и фактах. От его правильной, научно обоснованной организации и проведения зависят последующие этапы статистического исследования и его результаты, эффективность в целом. На основе материалов статистического наблюдения, их последующей обработки и анализа разрабатывается система показателей для управления и планирования народного хозяйства, для прогнозирования экономического и социального развития страны и отдельных его регионов.

Из сказанного следует, что *главной задачей статистического наблюдения в СССР является получение достоверных статистических данных* о процессе расширенного социалистического воспроизводства для решения актуальных задач совершенствования развитого социализма. Исходя из этой задачи для каждого вида статистического наблюдения определяются более конкретные, детальные цели и задачи, охватывающие все необходимые компоненты: единицы и объекты наблюдения, объем подготовительной работы, круг исполнителей и сроки выполнения отдельных видов работ, методы сбора данных, порядок контроля их достоверности, конечные результаты наблюдения и т. д.

Таким образом, статистическое наблюдение — большая, кропотливая и трудоемкая работа, требующая привлечения квалифицированных кадров, всесторонне продуманной ее организации, планирования, подготовки и проведения.

### 3.2. ВИДЫ И СПОСОБЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Статистическое наблюдение различается по видам и способам проведения. Виды статистического наблюдения можно классифицировать по ряду признаков (см. схему 1).

По степени охвата исследуемой совокупности статистическое наблюдение подразделяется на два вида: сплошное (полное) и несплошное (частичное).

При *сплошном наблюдении* обследованием охватываются все единицы изучаемой совокупности. Например, при Всесоюзной переписи населения 1979 г. регистрации по основной программе подлежало все без исключения население нашей страны. Примером сплошного наблюдения служит текущая отчетность предприятий и организаций, которая охватывает все подотчетные хозяйственные объекты и служит источником сведений для контроля и анализа выполнения государственного плана.

Сплошное наблюдение обеспечивает полноту информации об изучаемых фактах и явлениях. Однако проведение сплошного наблюдения не всегда возможно. В частности, сплошное наблюдение невозможно в случае физического уничтожения и порчи обследуемых единиц или когда обследуемая совокуп-

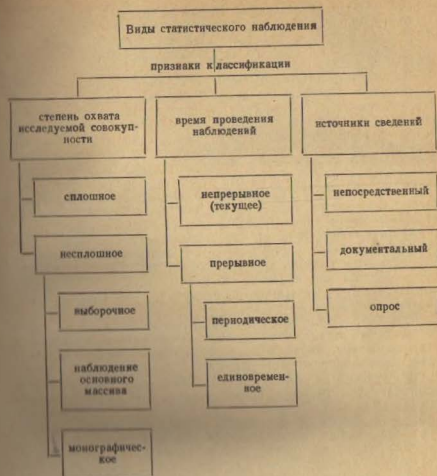


Схема 1. Классификация видов статистического наблюдения

ность слишком велика, а также когда сплошное наблюдение требует больших затрат времени и средств. По этой причине проводятся несплошные наблюдения.

При *несплошном наблюдении* обследованием охватывается только определенная часть изучаемой совокупности. В свою очередь несплошное наблюдение подразделяется на выборочное, наблюдение основного массива, монографическое.

*Выборочным* называют наблюдение части единиц исследуемой совокупности, выделенной методом случайного отбора. Этому виду обследования будет посвящена гл. 11.

*Наблюдение основного массива* охватывает собой обследование определенных, наиболее существенных по значимости изучаемых признаков единиц совокупности. В качестве примера

можно назвать обследование 15—20 % крупных предприятий, выпускающих 90—95 % определенной продукции. Или наблюдение хода весеннего сева только по итогам работы колхозов и совхозов в связи с тем, что на их долю приходится 95 % всей посевной площади страны.

Для *монографического наблюдения* характерно глубокое и всестороннее исследование лишь отдельных единиц совокупности, обладающих какими-либо специфическими особенностями. При этом не ставится цель охарактеризовать всю совокупность в целом. Примером такого вида наблюдения может служить изучение передового опыта предприятий, организаций, колхозов, совхозов, бригад и т. д. с целью его распространения или выявления недостатков в работе отстающих предприятий.

В зависимости от временного фактора наблюдение может быть непрерывным и прерывным. Прерывное в свою очередь включает наблюдение периодическое и единовременное.

*Непрерывное* (текущее) наблюдение осуществляется путем непрерывной регистрации фактов по мере их возникновения. При таком наблюдении прослеживаются все изменения изучаемого процесса или явления. Примером текущего наблюдения может служить учет выработки продукции, учет отпуска материалов в производство и т. д.

*Периодическое наблюдение* производится либо регулярно через определенные промежутки времени (периодическое наблюдение), либо нерегулярно, однократно, по мере надобности (единовременное наблюдение). Примером периодического наблюдения служат ежегодные переписи скота или учет товарных остатков, осуществляемый одновременно, т. е. на момент переписки.

В зависимости от источников собираемых данных различают наблюдение непосредственное, документальное и опрос.

*Непосредственное наблюдение* осуществляется путем фиксации фактов, лично установленных регистраторами в результате осмотра, измерения, подсчета признака изучаемого явления. Примером может служить снятие товарных остатков при инвентаризации, замеры времени на выполнение трудовых операций при изучении их трудоемкости и т. п.

*Документальное наблюдение* основано на использовании в качестве источника информации данных различных документов. Так, для составления статистической отчетности о работе предприятия используются документы бухгалтерского учета.

*Опрос* базируется на получении данных в форме ответов опрошиваемых лиц. Такой вид наблюдения характерен для переписей населения, проведения различных социологических обследований и опросов общественного мнения.

Статистическое наблюдение, как отмечено выше, различается не только по своим видам, но и по способам проведения, иначе говоря, способам сбора статистических данных. В статисти-

ческой практике применяются следующие способы наблюдения: *отчетный, экспедиционный, саморегистрации, анкетный.*

*Отчетный способ* заключается в представлении всеми предприятиями, организациями, колхозами, совхозами и т. д. отчетов в установленные сроки, адреса и по определенной форме. Это наиболее распространенный в нашей стране способ сбора статистических данных. Именно отчетный способ обусловил необходимость и возможность создания в стране автоматизированной системы государственной статистики (АСГС), что диктуется массовостью, большими объемами данных и строго регламентированными сроками обработки отчетности, поступающей в систему государственной статистики. Возможность автоматизации обработки статистической отчетности обеспечивается унификацией ее содержания, форм и однородностью алгоритмов обработки ее данных.

Суть *экспедиционного способа* состоит в том, что к каждой единице наблюдения посылаются специальные лица, называемые счетчиками или регистраторами, которые в специальных формулярах фиксируют сведения о наблюдаемом явлении. Такой способ применяется при сборе данных у населения, в частности при его переписях.

При *саморегистрации* характерно то, что специальные работники снабжают опрошиваемых бланками и дают инструкции о порядке их заполнения, заполняют же бланки сами опрошиваемые.

При *анкетном наблюдении* определенному кругу лиц вручаются (или публикуются в периодической печати) специальные анкеты (вопросники). Заполнение этих анкет носит добровольный характер и осуществляется, как правило, анонимно. Такой порядок снижает полноту и достоверность получаемой информации. В связи с этим названный способ наблюдения применяется в обследовании, где не требуется высокая точность, а нужны приближенные, ориентировочные результаты, например при изучении работы органов связи, торговых предприятий, издательств периодической печати.

Выбор видов и способов статистического наблюдения зависит от целого ряда факторов, среди которых: цели и задачи наблюдения, специфика наблюдаемого объекта, срочность представления результатов, наличие подготовленных кадров, возможность применения технических средств сбора и обработки данных и др. В настоящее время последний фактор — определяющий. Он выражается в создании и функционировании АСГС, предусматривающей комплексный подход к организации процессов сбора, передачи и обработки статистической информации на базе использования автоматизированного банка данных. В этих условиях появляется возможность значительно сократить перечень регистрируемых в статистических формулярах данных, а следовательно, упростить и уменьшить трудоемкость процедуры сбора и повысить достоверность статистиче-

ской информации. При этом вследствие автоматизации передачи информации в АСГС сокращается и время прохождения данных от источника сведений к месту обработки данных и потребителю результатов обработки, т. е. управленческому органу. Это повышает общую эффективность функционирования системы государственной статистики и эффективность управления народным хозяйством.

### 3.3. ПРОГРАММНО-МЕТОДОЛОГИЧЕСКОЕ ОБЕСПЕЧЕНИЕ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Одной из важнейших задач, которую необходимо решить при подготовке статистического наблюдения, является определение объекта и единицы наблюдения.

Объектом наблюдения называется совокупность общественных явлений или процессов, подлежащих исследованию. Объектом наблюдения может быть промышленность (совокупность промышленных предприятий), сельское хозяйство (совокупность колхозов, совхозов и других сельскохозяйственных предприятий) и т. д. Объектом наблюдения при переписи населения 1979 г. являлось население (совокупность жителей страны).

При установлении объекта наблюдения следует строго и точно определить границы изучаемой совокупности. Для четкого ограничения наблюдаемой совокупности необходимо указать существенные признаки объекта наблюдения, отличающие его от других объектов. Например, при переписи населения требуется установить, какое именно население подлежит регистрации: постоянно проживающее в данной местности или личное, т. е. фактически находящееся там в момент переписи.

Определяя объект наблюдения, необходимо точно указать единицу наблюдения и единицу совокупности.

Единицей наблюдения является составной элемент объекта наблюдения, представляющий собой источник информации. В качестве единицы наблюдения могут быть приняты объекты: промышленное предприятие, колхоз, совхоз, рабочий, учащийся и т. д., в зависимости от конкретных задач статистического наблюдения.

Единицей совокупности называют составной элемент объекта наблюдения, который служит основой счета и обладает признаками, подлежащими регистрации в процессе наблюдения. Например, при проведении переписи плодово-ягодных насаждений единицей совокупности будет отдельное дерево (например, груша), так как признаки, регистрируемые в этом обследовании (возраст насаждений, их породный состав и т. п.), относятся не к колхозу, совхозу, а к каждому дереву.

Правильное определение единицы совокупности и единицы наблюдения имеет большое значение, поскольку дает возможность обеспечить не только полноту и точность учета всех эле-

ментов объекта обследования, но и правильность обработки и анализа результатов статистического наблюдения.

На необходимость строгого и точного определения единицы наблюдения указывал В. И. Ленин. Критикуя дореволюционную государственную фабрично-заводскую статистику, он показал, как народник Карышев, используя статистические данные, в которых неопределенно толковались понятия фабрика и завод, неправильно истолковал неуклонное сокращение их числа в России в конце XIX в.<sup>1</sup>

Единицы наблюдения называют отчетными единицами, когда наблюдение проводится в форме собрания отчетности. Отчетными единицами в этом случае являются промышленные предприятия, организации, стройки, колхозы, совхозы и т. д.

Учитывать все множество признаков, характеризующих объект наблюдения и его единицы, не всегда является возможным и часто не представляет интереса. Поэтому при подготовке наблюдения необходимо решить вопрос о том, какие признаки будут подлежать регистрации в соответствии с поставленными целями и задачами наблюдения. Все это находит свое отражение в программе статистического наблюдения, представляющей собой перечень вопросов, на которые должны быть получены ответы в ходе наблюдения.

Разработка программы наблюдения — одна из наиболее значимых и ответственных задач. От того, насколько правильно разработана программа, зависит общий успех проведения наблюдения.

Программа статистического наблюдения должна разрабатываться исходя из задач наблюдения, с учетом потребностей в определенных статистических данных для целей государственного управления, хозяйственного руководства и научных исследований. При разработке программы следует учитывать ряд предъявляемых к ее содержанию требований. К основным из них относятся следующие:

программа должна по возможности содержать только те признаки, которые необходимы и значения которых будут использованы или при разработке материалов наблюдения, или в контрольных целях. При этом надо стремиться к полноте сведений, обеспечивающих получение доброкачественных результатов. При отсутствии уверенности в получении полных и достоверных данных целесообразно ограничить объем собираемой информации, чтобы получить хотя и небольшой, но, безусловно, достоверный материал. В. И. Ленин, формулируя положение о широте программы, писал: «...лучше получить немного сравнительно достоверных, полных и единообразных сведений,

<sup>1</sup> См.: Ленин В. И. Полн. собр. соч., т. 4, с. 30.



чем много сведений отрывочных, сомнительных и несравнимых»<sup>1</sup>;

программа наблюдения должна быть по возможности построена так, чтобы ответы на один вопрос могли использоваться для контроля ответов на другие. Так, программой переписи населения СССР 1979 г. предусматривалось включение таких вопросов, ответы на которые были бы взаимно контролируемы. Например, для выяснения возраста населения ставились два вопроса: о дате рождения и о числе исполнившихся лет (см. приложение);

вопросы программы наблюдения должны быть четкими и предельно ясными, не допускающими разных толкований.

Разработка программы — дело довольно сложное. Разработчики должны иметь достаточно глубокое представление об объекте наблюдения и его свойствах. Поэтому программа должна разрабатываться коллективно, с привлечением практических и научных работников, хорошо знающих сущность изучаемых явлений.

Разрабатывая программу, она должна быть всесторонне обсуждена в прессе, на соответствующих совещаниях. В. И. Ленин, продолжая критику дореволюционной фабрично-заводской статистики в части неразработанности программы собирания сведений, писал: «Если такая программа будет вырабатываться в канцеляриях, не подвергаясь критике специалистов и (что особенно важно) всестороннему обсуждению прессы, то сведения никогда не могут быть сколько-нибудь полными и единообразными»<sup>2</sup>.

Функционирование третьей очереди АСГС, в частности внедрение автоматизированных банков данных, обуславливает ряд изменений в методологии разработки программы наблюдения. В условиях функционирования АБД реализуется так называемая *регистровая форма статистического наблюдения*. Основой ее является создание регистра, своего рода автоматизированной картотеки однородной совокупности единиц статистического наблюдения определенного типа. Назначение регистра — обеспечение синтетической характеристики объекта наблюдения на стадии получения первичной информации. Регистр объединяет показатели целого комплекса статистических отчетов, т. е. обеспечивает взаимосвязь целой совокупности показателей.

В настоящее время функционируют регистры промышленных предприятий, сельскохозяйственных организаций, строев. Рассмотрим некоторые особенности статистического наблюдения на примере регистра промышленных предприятий (РПП). Известно, что производственная деятельность определяется множеством показателей, отражающих все стороны этой деятельности в динамике. Поэтому система показателей, включаемых в РПП,

обеспечивает возможность получения данных о годовых, полугодовых, квартальных и более коротких периодах производственно-хозяйственной деятельности; возможность получения соответствующих производных показателей, выполняющих характеристику производственно-хозяйственной деятельности, проведение комплексного анализа деятельности.

Для формирования содержания РПП используется система показателей ф. № 22 годового отчета «Сводная таблица основных показателей», в наибольшей мере соответствующая перечисленным требованиям. Вся совокупность показателей в РПП разбивается на группы, такие, как «Продукция», «Фонды», «Труд и заработная плата», «Численность работающих», «Оборотные средства», «Прочие расходы». Формирование таких групп обусловлено необходимостью в ряде случаев их автономного использования при решении соответствующих задач. Помимо показателей ф. № 22, в РПП фиксируются и показатели других форм, а также ряд справочных данных (тогда ввода предприятия в действие и др.). Показатели по каждому предприятию располагаются по годам, образуя динамические ряды. Динамическое отображение данных обеспечивается постоянным их накоплением по каждому предприятию с момента включения его в регистр на все время функционирования.

РПП включает всю совокупность показателей, необходимых для комплексного анализа хозяйственной деятельности предприятий и их групп по отраслям, министерствам, территориальной подчиненности. В связи с этим в значительной мере сокращается перечень данных, подлежащих фиксации в статистической отчетности. В частности, отпадает необходимость в предоставлении данных по годам, полугодиям, кварталам, расистым величинам (она будет рассчитываться машиной на основе хранившихся и введенных показателей).

Для проведения наблюдения необходим свой инструментариум, в частности формуляры и инструкции. *Формуляром* называется специальный документ, в котором фиксируются ответы на вопросы программы. В зависимости от конкретного содержания проводимого наблюдения формуляр может называться формой статистической отчетности, переписным или опросным листом, анкетой, картой, карточкой или просто бланком. Различают две разновидности формуляров — карточную и списочную. *Формуляр-карточка* предназначается для отражения одной единицы статистической совокупности. В *формуляре-списке* фиксируются сведения о нескольких единицах совокупности. Примером может служить формуляр «Список проживающих в помещении», заполнявшийся при переписи населения 1979 г. Когда для обработки данных применялись клавишные машины и группировка документов производилась вручную, использовались в основном карточные формуляры. В настоящее время, в условиях широкого применения электронных и перфорационных машин (ЭВМ и ПЭВМ), предпочтение отдается списочным формулярам, так как они менее трудоемки для перфорации — самой трудоемкой операции машинной обработки, а также позволяют экономить бумагу. Кроме того, списочный формуляр менее трудоемок и для заполнения, поскольку ряд признаков, общих для всей совокупности единиц наблюдения, отражаемых в формуляре, фиксируется однократно, а не на каждой карточке. Однако в ряде случаев даже при наличии всех названных преимуществ списочный формуляр становится неприемлемым. В частности, при наличии большого количества вопросов

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 4, с. 32.

<sup>2</sup> Там же, с. 30.



он очень разрастается. Это наблюдается, когда в формуляре выделяются зоны, подлежащие заполнению машиночитаемыми знаками, как, например, в переписном листе переписи населения СССР 1979 г. (см. приложение). В этом случае используются карточный формуляр.

Внедрение и функционирование АСГС предъявляет определенные требования к содержанию и форме статистического формуляра. Основным направлением совершенствования работы по проектированию статистических документов является унификация их форм, т. е. соблюдение единых требований к содержанию и построению форм статистических документов. Унификация предполагает прежде всего типизацию состава включаемых показателей исходя из целей и задач статистического исследования. Учитывается и способ организации обработки данных. Так, функционирование АБД позволяет минимизировать число включаемых в формуляр показателей вследствие того, что постоянно действующие плановые данные и данные прошлых периодов хранятся в АБД, а расчетные показатели (темпы роста, процент выполнения плана и др.) будут подсчитываться автоматически вычислительной машиной. Унификация статистических документов определяет правила размещения показателей на бланке, требования к их заполнению, размеры строк и граф и т. д. Бланк статистического формуляра включает три части: заголовочную, содержательную (предметную) и оформительную. Так, в заголовочной части статистического отчета указываются: коды<sup>1</sup> формы отчета и отчитывающегося предприятия; наименование, почтовый адрес, ведомственная подчиненность и отраслевая принадлежность этого предприятия; наименование и адрес получателя; способ и периодичность представления отчета; организации, в которые посылается отчет, и сроки его представления; название формы отчета.

Основная, содержательная часть отчета оформляется обычно в виде таблиц и содержит наименование, коды и числовые значения статистических показателей. Оформительная часть содержит сведения о должностных лицах, подписывающих документ, их подписи и даты подписей.

Заполнение любого статистического формуляра осуществляется в соответствии с инструкцией. Инструкция содержит письменные указания и разъяснения по выполнению программы наблюдения. В инструкции отражаются цель и задачи статистического наблюдения, сведения об объекте и единицах наблюдения, о времени и сроках проведения наблюдения, об оформлении результатов и сроках их представления в соответствующие организации. В инструкции даются указания, как следует понимать тот или иной вопрос программы, приводятся примеры возможных ответов и порядок заполнения формуляра. Так, в ин-

струкции о порядке проведения переписи и заполнения переписных листов указываются метод переписи, время и порядок заполнения переписных листов, категории переписного населения (наличие, постоянно проживающее, временно проживающее население), единые правила записи ответов.

От четкого составления инструкции во многом зависит качество статистического исследования, поэтому она должна быть написана кратко, просто и ясно.

При организации статистического наблюдения должен быть решен вопрос о времени наблюдения. При этом должен быть установлен период, в течение которого будет проводиться наблюдение — срок наблюдения, и точно определено время, к которому относятся регистрируемые сведения, — объективное время наблюдения. Оно может быть либо определенным моментом, либо тем или иным периодом. Так, за определенный период (квартал, месяц, декаду, сутки) фиксируются сведения о выпуске продукции, о расходе материальных ресурсов и т. д., а по постоянно на определенный момент подсчитывается численность населения, размер имеющегося жилищного фонда и т. д. Момент времени, к которому приурочены регистрируемые в процессе наблюдения сведения, называют критическим моментом статистического наблюдения. Например, критическим моментом при переписи населения 1979 г. было 12 часов ночи с 16 на 17 января 1979 г. Поэтому родившиеся после 12 часов ночи регистрации не подлежали и в переписные листы не вносились, а умершие после 12 часов ночи записывались как живые. Таким образом, устанавливая критический момент, можно с точностью фотоснимка отразить истинное состояние изучаемого явления на определенный момент времени.

Срок наблюдения характеризует период, в течение которого должна осуществляться регистрация сведений об изучаемом явлении. Например, во Всесоюзной переписи населения СССР 1979 г. это был период с 8 часов утра 17 января по 24 января 1979 г. включительно, т. е. 8 дней. Обычно стремятся к тому, чтобы срок наблюдения был не слишком удален от критического момента, с тем чтобы можно было без особых затруднений восстановить истинное состояние объекта наблюдения на тот момент. Иногда устанавливается срок (число месяца, день недели, количество дней после критического момента), не позднее которого следует представить сведения.

#### 3.4. ВОПРОСЫ ОРГАНИЗАЦИОННОГО ОБЕСПЕЧЕНИЯ ПОДГОТОВКИ И ПРОВЕДЕНИЯ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Для успешной подготовки и проведения статистического наблюдения должны быть решены и вопросы его организационного обеспечения. Это делается при составлении орга-

<sup>1</sup> Код — условное цифровое или буквенно-цифровое обозначение какого-либо признака. Коды устанавливаются едиными общесоюзными классификаторами, утвержденными в централизованном порядке.

низационного плана наблюдения. В нем устанавливаются цели и задачи наблюдения, объект наблюдения, место, время, сроки наблюдения, круг лиц, отвечающих за проведение наблюдения; решается также ряд практических вопросов, связанных с подготовкой, организацией и проведением статистического наблюдения. Таким образом, организационный план наблюдения содержит перечень мероприятий, а также описание характеризующих их обстоятельств места и времени, направленных на успешное проведение работы по сбору необходимых сведений.

Обязательным элементом организационного плана является указание органа наблюдения. Орган наблюдения — это организация, учреждение или их подразделение, осуществляющие подготовку и проведение наблюдения и несущие ответственность за эту работу. Статистическое наблюдение осуществляется в масштабе всей страны ЦСУ СССР, а в пределах своих регионов — ЦСУ союзных республик и местными статистическими органами. В качестве органов наблюдения могут также выступать учетно-статистические отделы министерств и ведомств, различные общественные организации. В организационном плане определяются права и обязанности каждого органа наблюдения, взаимоотношения их между собой.

Для обеспечения полноты охвата объекта наблюдения, достоверности собираемых сведений в организационном плане четко устанавливается и место наблюдения. Под местом наблюдения понимается пункт, где непосредственно будет осуществляться регистрация наблюдаемых фактов, заполнение статистических формуляров. Часто место проведения наблюдения совпадает с местонахождением единиц наблюдения, например статистическая отчетность о деятельности промышленного предприятия, колхоза, совхоза, какой-либо организации составляется по месту их нахождения.

Однако при специально организованном наблюдении, в случае, когда единицы наблюдения меняют место своего пребывания (например, при переписи населения), приходится специально указывать место наблюдения. Так, в инструкции по проведению переписи населения записано, что оно учитывается по месту жительства, а не по месту работы.

К числу вопросов организационного обеспечения проведения статистического наблюдения относится выбор организационной формы наблюдения. В статистической практике реализуются две организационные формы наблюдения — отчетность и специальное статистическое обследование.

Отчетность — это такая организационная форма, при которой единицы наблюдения (предприятие, организация, стройка и т. д.) представляют сведения о своей деятельности в виде формуляров, жестко регламентированных ЦСУ СССР по содержанию, форме и срокам. В настоящее время отчетность является важнейшей формой статистического наблюдения в СССР. ЦСУ СССР утверждает общесоюзный минимум ста-

тистических показателей и форм документации статистической отчетности для предприятий, организаций и строек.

Отчетность как форма наблюдения имеет свои преимущества, вытекающие из ее особенностей: обязательность представления и документально обоснованная и юридически подтвержденная достоверность. Достоверность гарантируется юридической ответственностью руководителей отчетывающихся предприятий и организаций, а также тем, что данные статистической отчетности получаются в результате обработки первичных документов, при которой обеспечивается возможность сопоставимости и контроля данных.

Статистические отчеты являются одним из основных источников сведений о состоянии развития народного хозяйства страны и его отдельных отраслей. В них содержатся данные о деятельности предприятий и организаций, колхозов, совхозов, строек, научных учреждений и т. д., которые используются как для практических задач управления, так и в научных целях.

Например, «Отчет промышленного предприятия о выполнении плана по труду» (ф. № 9) содержит такие показатели, как среднесписочная численность работников промышленного предприятия, движение численности работников и фонд их заработной платы; объем продукции, среднегодовая численность промышленно-производственного персонала и выработка на одного работающего; размер и состав фонда заработной платы промышленно-производственного персонала, фонд материального поощрения; единовременные премии, прочие денежные выплаты работникам и служащим, не входящие в фонд заработной платы, а такжечисления страховых взносов по обязательному страхованию. Эта форма служит основным источником данных для получения сводных отчетов по использованию трудовых ресурсов и фонда заработной платы по отраслям промышленности.

Виды статистической отчетности различаются по степени охвата объектов наблюдения, периодичности и способу представления, порядку прохождения.

Прежде всего статистическая отчетность делится на типовую и специализированную.

Типовая отчетность имеет одинаковое содержание и форму для всех предприятий (организаций) данной отрасли народного хозяйства или всех его отраслей. Специализированная отчетность отражает специфику той или иной отрасли.

По периодичности представления различают отчетность недельную, двухнедельную, месячную, квартальную, годовую. По способу представления отчетность делится на почтовую и телеграфную (срочную). Такое деление определяет различие в организации не только сбора данных, но и их сводной обработки. Так, сводная обработка почтовой отчетности осуществляется на областных ВЦ с записью результатов в сводные таблицы и на магнитную ленту, которые передаются на ВЦ республиканского статуправления для дальнейшей сводки и т. д. Срочная отчетность обрабатывается по другой схеме: формируются два потока — поток плановых данных (на год, с начала года, на квартал и на отчетный период) и фактических данных за предыдущий отчетный период (почтовые данные) и поток фактических дан-

ных за отчетный период (телеграфные данные). Полученные в результате сводной обработки на областном ВЦ данные передаются в вышестоящие уровни по каналам связи. В целях сокращения сроков прохождения статистической информации по уровням системы государственной статистики (от областного к республиканскому, от республиканского к союзному) телеграфная отчетность будет получать все большее развитие.

В зависимости от установленного порядка прохождения отчетность делится на централизованную и нецентрализованную. Централизованная отчетность от предприятий и организаций поступает в систему государственной статистики, где обрабатывается и отсюда рассылается соответствующим хозяйственным и советским органам управления. Нецентрализованная отчетность поступает и обрабатывается в министерствах и ведомствах с последующей передачей полученных результатов обработки в органы ЦСУ СССР.

При проектировании и создании АСГС решаются вопросы предоставления статистической информации предприятиями и организациями в систему ЦСУ на машинных носителях (перфолентах, магнитных лентах) или обеспечения ввода данных в ЭВМ непосредственно по каналам связи, соединяющим ВЦ предприятий и ВЦ ЦСУ.

Машинная обработка предъявляет свои требования к форме статистического отчета. Так, для обеспечения возможности переноса данных на машинные носители в бланке отчета предусмотрена необходимость кодирования признаков. С этой целью в верхнем правом углу документа выделена рамка, в которой проставляют коды признаков заголовочной части отчета. Коды проставляют экономисты отделов (секторов) подготовки и выпуска статистических материалов. Как правило, во всех статистических отчетах, подлежащих обработке на ЭВМ, вводится место для записи контрольных сумм (по строкам, графам, страницам). Введение контрольных сумм обеспечивает автоматизацию процесса контроля перфорации машинных носителей.

При оснащении системы ЦСУ машинным ЕС ЭВМ, имеющим большой объем памяти на магнитных дисках, появляется возможность отказаться от наличия в статистических отчетных плановых данных, а также данных за прошлые периоды, общий объем которых в ряде статистических отчетов составляет до 40 %. Эти данные будут вводиться в память ЭВМ один раз на специальных бланках, накапливаться и храниться в автоматизированном банке данных. При оснащении системы ЦСУ считывающими устройствами, способными автоматически вводить информацию в ЭВМ с бумажных формуляров, можно будет использовать в качестве средств ввода специальные машиночитаемые бланки. Это упростит операции по перфорации данных и повысит эффективность применения ЭВМ.

Помимо отчетности, в нашей стране реализуется и другая организационная форма наблюдения — *специальное статистическое обследование*. Примерами такой формы являются специально организованные переписи, осуществление специального учета (посевных площадей, скота, тракторов), обследования

бюджетов семей рабочих, служащих, колхозников и т. д. Способы проведения такого обследования различны. В частности, оно может осуществляться путем использования данных бухгалтерского учета и отчетности предприятий (например, переписи фактических остатков материальных ценностей) или путем специально организованной регистрации (посевных площадей, скота, тракторов и др.).

Классическим примером второго способа специального статистического обследования является проведение переписей населения. В СССР осуществлено шесть таких переписей: в 1920, 1926, 1939, 1959, 1970 и 1979 гг. В результате проведения и обобщения данных переписей населения получались сведения о численности, размещении, составе населения по различным признакам. Все эти сведения крайне важны для изучения вопросов социального и экономического развития, текущего и перспективного планирования деятельности отраслей народного хозяйства. Они также необходимы для изучения демографических процессов в стране.

Примером специального статистического обследования является также проведение выборочного социально-демографического обследования населения по состоянию на 1 января 1985 г. Материалы этого обследования являются важным источником для планирования дальнейшего развития страны.

## ГЛАВА 4 ОБЕСПЕЧЕНИЕ КАЧЕСТВА МАТЕРИАЛОВ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

### 4.1. ДОСТОВЕРНОСТЬ И СВОЕВРЕМЕННОСТЬ СТАТИСТИЧЕСКИХ ДАННЫХ — ОСНОВНАЯ ЗАДАЧА СТАТИСТИЧЕСКИХ ОРГАНОВ

Важнейшей задачей органов советской статистики является обеспечение достоверности и своевременности статистических данных. Объективность и строгая достоверность — закон нашей статистики.

Под *достоверностью* понимается степень объективного отображения статистическими данными сущности явлений и процессов. *Своевременность* характеризует поступление данных в сроки, соответствующие целям проводимого наблюдения.

В современных условиях значение достоверности и своевременности информации возрастает. Это обусловлено тем, что с усложнением экономики значительно возрастает «цена» принятого на основе недостоверных данных ошибочного решения, кроме того, запоздавшие данные могут оказаться бесполезными.

Широкое применение экономико-математических методов и ЭВМ предъявляет в свою очередь повышенные требования к качеству обрабатываемой информации. Ущерб от ошибок в ней



наносится прежде всего самому вычислительному центру, так как приходится затрачивать время и средства на их устранение, а иногда повторять обработку и расчеты на ЭВМ. Не менее существенный ущерб наносится не обнаруженными на ВЦ ошибками, которые, как правило, вызывают отрицательные последствия в случае принятия на их основе управленческих решений.

Все это обуславливает необходимость реализации системы мер, обеспечивающих достоверность и высокое качество статистической информации на всех этапах ее прохождения: от источника возникновения до органа, принимающего на ее основе управленческое решение. Особое значение имеет контроль на стадиях наблюдения и подготовки данных к обработке, поскольку ошибки, допущенные на этих стадиях, трудно, а порой и практически невозможно исправить на последующих этапах.

Одной из мер, осуществляемых в государственном масштабе для обеспечения достоверности исходной статистической информации, является борьба с приписками и другими умышленными искажениями отчетных данных. Принятое ЦК КПСС и СМ СССР постановление «О мерах по предотвращению фактов обмана государства и по усилению контроля за достоверностью отчетов о выполнении планов и обязательств», а также Указ Президиума Верховного Совета СССР устанавливают, что «приписки в государственной отчетности и представление других умышленно искаженных отчетных данных о выполнении планов должны рассматриваться как противогосударственные действия, наносящие вред народному хозяйству СССР, и лица, виновные в этом, наказываются лишением свободы на срок до трех лет. Персональную ответственность за своевременность представления и достоверность отчетных данных несут руководители предприятий».

Органам государственной статистики поручено проверять состояние учета и отчетности и достоверность отчетных данных в народном хозяйстве, а также проверять в министерствах и ведомствах состояние работы по организации первичного учета и обеспечению достоверности отчетных данных на подведомственных им предприятиях и обсуждать в необходимых случаях результаты проверок с участием соответствующих министерств и ведомств. В случае обнаружения в ходе проверки приписок и других умышленных искажений отчетных данных органы государственной статистики вносят предложения о привлечении виновных лиц к ответственности в соответствии с действующим законодательством.

В случаях, когда искажения отчетности являются следствием неудовлетворительной постановки учета на предприятии (недостаточная квалификация учетно-статистических кадров, слабая производственная дисциплина и др.), от органов государственной статистики совместно с министерствами и ведомствами требуется кропотливая и настойчивая работа по совершенствованию первичного учета и отчетности на предприятиях путем органи-

зации соревнования работников учета, проведения общественных смотров, организации школ передового опыта, проведения совещаний-семинаров и т. д.

Контроль достоверности статистической информации осуществляется на всех этапах обработки данных на ВЦ: от приема первичных статистических отчетов до выпуска сводной отчетности. Несмотря на высокие требования к достоверности и качеству отчетных данных, направляемых в статистические органы предприятиями и организациями, поступающая на ВЦ информация не всегда является безошибочной. Поэтому при приеме отчетов ВЦ обязаны проконтролировать поступающие данные и в случаях обнаружения ошибок возвратить отчет предприятию (организации) для исправления.

Ошибки могут возникать также при выполнении машинных и ручных технологических операций обработки информации, поэтому и здесь необходимы меры контроля для предотвращения или устранения ошибок.

#### 4.2. ОШИБКИ НАБЛЮДЕНИЯ

Опыт показывает, что при изучении экономических явлений степень точности данных наблюдений может быть различной. Каждое конкретное измерение величины данных, осуществляемое в процессе наблюдения, дает, как правило, приближенное значение величины явления, в той или иной мере отличающееся от истинного значения этой величины. Расхождения между результатом наблюдения и истинным значением величины наблюдаемого явления называются *ошибками наблюдения*.

В зависимости от характера, стадии и причин возникновения различают несколько типов ошибок наблюдения. Классификация ошибок приведена в табл. 4.1.

По своему характеру ошибки делятся на случайные и систематические. *Случайными* называются *ошибки*, возникновение которых обусловлено действием случайных факторов. Например, при опросе опрашиваемый мог оговориться или опрашиваемый при записи результата опроса мог допустить опisku. Такие ошибки могут быть направлены как в сторону преувеличения, так и в сторону преуменьшения. При сводной обработке результатов массового наблюдения они, как правило, взаимопогашаются и на конечном результате не отражаются.

Большую опасность представляют *систематические ошибки*. Например, ошибки измерения плохо настроенным измерительным прибором; ошибки, являющиеся следствием неясной формулировки вопроса программы наблюдения, и др. Их наличие обуславливает направленное искажение конечного результата наблюдения.

В зависимости от стадии возникновения различают ошибки регистрации; ошибки, возникающие в ходе подготовки данных

к машинной обработке; ошибки, проявляющиеся на стадиях машинной обработки данных на ЭВМ.

К ошибкам регистрации относятся те неточности, которые имеют место при записи данных в статистический формуляр наблюдения (первичный документ, статистический отчет, бланк переписи т. д.) или при вводе данных в ЭВМ через терминальное устройство (телетайп, дисплей). Например, случайные опи-

ция, установленным требованиям), либо, наконец, сбоями в работе оборудования (перфораторов, устройств для записи на магнитную ленту, контрольных, сортировальных машин, ЭВМ и др.).

Для выявления и предупреждения ошибок наблюдения важно знать их виды и причины возникновения.

Различают следующие виды ошибок:

*ошибки измерения* связаны с определенными погрешностями, которые возникают при однократном статистическом наблюдении социально-экономических явлений;

*ошибки репрезентативности* характерны для несплошного наблюдения и связаны с тем, что величина изучаемого признака в отобранной совокупности отличается от величины этого признака во всей изучаемой совокупности. Подробнее о таких ошибках говорится в гл. 11;

*преднамеренные ошибки* возникают по причине сознательного искажения данных с целью приукрасить действительное положение дел, чтобы получить премию, прибавку к зарплате или другое дополнительное вознаграждение. Например, приписки к отчетности с целью показать более высокий процент выполнения плана, занижение плановых заданий, сокрытие (резервирование) той или иной части изготовленной продукции, материальных ресурсов, скота и т. п. Как отмечалось ранее, такие действия рассматриваются как преступления и наказываются в уголовном порядке;

*непреднамеренные ошибки*, как правило, носят случайный характер и могут рассматриваться как промахи в работе. Такие ошибки являются результатом низкой квалификации работников, небрежности, недобросовестности в работе. Иногда они связаны с некоторыми субъективными явлениями, например, стремлением лиц пожилого возраста иногда округлять свой возраст до чисел, кратных 5 или 10, и т. п. К непреднамеренным следует отнести и ошибки запоминания.

Таблица 4.1. Классификация ошибок наблюдения

Привязки классификации	Виды ошибок
Характер ошибок	Случайные
	Систематические
Стадия возникновения	Ошибки регистрации
	Ошибки при подготовке данных к машинной обработке
	Ошибки в процессе машинной обработки
Причины возникновения	Ошибки измерения
	Ошибки репрезентативности
	Преднамеренные ошибки
	Непреднамеренные ошибки

ски или преднамеренное искажение значений отдельных показателей; несоблюдение формы бланка, т. е. запись не в установленную строку или графу документа. Нередкими являются ошибки, возникающие при наборе данных на клавиатуре телетайпа или дисплея, если информация вводится непосредственно в ЭВМ, минуя документ.

Особую категорию ошибок представляют те ошибки, которые возникают на ВЦ в ходе подготовки данных наблюдения к обработке или в процессе самой обработки. Возникновение ошибок на ВЦ может быть вызвано либо некачественным состоянием первичных документов (неправильное, нечеткое, небрежное заполнение данных в документах) или машинных носителей информации — перфокарт, перфолент, магнитных лент (деформация носителя, физический дефект, ошибки в записях, некомплектность массива), либо неправильным хранением информации (несоответствие помещения, где хранится информа-

#### 4.3. МЕТОДЫ ПРОВЕРКИ ДОСТОВЕРНОСТИ ДАННЫХ НАБЛЮДЕНИЯ

Для обеспечения необходимой достоверности данных наблюдения в государственной статистике реализуется система мер, предусматривающая тщательную проверку качества поступающей в статистические органы первичной отчетности или данных специальных наблюдений. Важным элементом такой системы является, как указывалось выше, проведение проверок достоверности отчетных данных непосредственно на местах, т. е. на предприятиях, в организациях, колхозах, совхозах и т. д. Целью таких проверок является установление достоверности отчетных данных, правильности ведения первичного учета, выявление и устранение причин несвоевременного представления

государственной отчетности и фактов искажения отчетных данных.

Проверки могут быть тематические, комплексные и повторные. При тематических проверках проверяется достоверность отчетных данных по одной или нескольким формам отчетности одной отрасли статистики.

Комплексные проверки предусматривают одновременную проверку достоверности отчетных данных по двум или нескольким отраслям статистики. Такие проверки дают возможность полнее вскрыть состояние учета и отчетности в целом по проверяемому объекту или ведомству и сделать по этому вопросу более глубокие выводы. Повторные проверки проводятся через год после проведения предыдущей проверки с целью выяснения, устранены ли недостатки, выявленные в ходе предыдущей проверки.

Перед проведением проверки составляется программа, в которой указывается перечень показателей, подлежащих проверке; количество предприятий, организаций, учреждений, колхозов, в которых должна быть проведена проверка; срок ее проведения; за какой период проверяются данные; какими инструкциями и указаниями следует руководствоваться при проверке; перечень основных первичных документов, на основании которых составляется проверяемая отчетность; порядок сопоставления и взаимосвязи показателей проверяемых отчетов с показателями других отчетов.

Результаты каждой проверки оформляются актом, в котором указываются обнаруженные недостатки в учете и отчетности, приписки и другие искажения отчетных данных по каждому факту искажений, особенно по припискам (сокрытиям); указывается, в каком отчете, по какому показателю, за какой период допущено искажение данных, размер приписки (сокрытия); влияние искажения данных на степень выполнения плана; последствия приписок, была ли в связи с припиской и другими искажениями незаконно выплачена премия, кому и в каких размерах; причины недостоверности данных с указанием виновных в этом лиц.

Результаты проверок излагаются в докладной записке на имя руководителя соответствующего статистического органа, проводившего проверку, с приложением акта проверки по каждому проверенному объекту. Во всех случаях установления фактов недостоверности отчетности статистические органы дают предприятию указания о внесении исправлений в отчете. В случаях установления приписок и других умышленных искажений отчетных данных органы статистики вносят предложения о привлечении виновных лиц к ответственности в соответствии с действующим законодательством.

Характерной чертой современного этапа развития государственной статистики является, как отмечалось в гл. 2, широкое применение ЭВМ и других технических средств, обеспечивающих однократный сбор и комплексное использование собираемых данных при решении большого числа взаимосвязанных задач. В связи с этим задачи обеспечения достоверности статистических данных нельзя рассматривать в отрыве от некоторых вопросов,

связанных с организацией машинной обработки информации на вычислительных центрах статуправлений. Здесь необходимо учитывать два обстоятельства. Во-первых, большинство контрольных операций в условиях применения ЭВМ могут выполняться автоматизированно. При этом по ряду задач реализуется процесс автоматического исправления ошибок (автокоррекция). Во-вторых, возникает проблема учета и устранения ошибок, связанных с самим процессом машинной обработки информации.

Основными видами контроля достоверности данных являются синтаксический, логический и арифметический (табл. 4.2).

Таблица 4.2. Виды и содержание контроля

Способы контроля	Что проверяется
Синтаксический (вручную)	Структура документа
	Полнота документа (наличие реквизитов)
	Полнота заполнения строк
Логический (на ЭВМ)	Соответствие кодов и наименований признаков
	Наличие отклонений от заданных значений
	Наличие логических связей между показателями
Арифметический (на ЭВМ)	Соответствие построчных контрольных сумм документа и контрольных сумм ЭВМ
	Соответствие пографных контрольных сумм документа и контрольных сумм ЭВМ

При синтаксическом контроле проверяются правильность структуры документа, наличие необходимых реквизитов, полнота заполнения строк документов в соответствии с установленными правилами. Например, в ходе приема переписных листов проверяется внешнее качество заполнения документа, соблюдение правил использования информационного поля, наличие ответов на все вопросы программы, правильность оформления ответов (в частности, чтобы графические отметки, оформляющие ответы в переписном листе, были сплошными, без разрывов, четкими и черными).

Значимость и необходимость синтаксического контроля обусловлены применением для обработки данных вычислительной техники, которая предъявляет жесткие требования к соблюдению правил заполнения формуляров, поскольку они либо подлежат непосредственному вводу в ЭВМ (например, переписные листы во время Всесоюзной переписи населения 1979 г.), либо



используются для последующего переноса данных на машинные носители (перфокарты, перфоленты, магнитные ленты).

Логическим контролем проверяется правильность записи кодов, соответствие их наименованиям и значениям показателей. Проверяется отсутствие недопустимых отклонений значений показателей от наиболее вероятных и наличие необходимых взаимосвязей между показателями. Так, при контроле правильности заполнения переписных листов сопоставляются ответы на различные вопросы, выявляющие логически несовместимые сочетания (например, десятилетний мальчик показан женатым). В случае обнаружения подобных сочетаний истинный ответ находится путем дальнейших сопоставлений с ответами на другие вопросы (например, если далее в переписном листе записано, что этот мальчик имеет высшее образование и работает преподавателем в школе, становится очевидной ошибка в регистрации возраста).

При арифметическом контроле проверяется значение показателей путем арифметических действий. В большинстве случаев арифметический контроль базируется на сравнении полученных итогов с предварительно подсчитанными контрольными суммами по строкам или по графам. Такой контроль, в частности, осуществляется при проверке правильности перфорации данных, проверке достоверности информации для передачи и ввода в ЭВМ.

Одним из видов арифметического контроля является балансовый метод. Он основывается на взаимном балансировании итоговых показателей одной совокупности при расчленении ее на отдельные элементы. Так, например, ресурс какого-то материала на предприятии на конец месяца складывается из остатка на начало месяца плюс поступление в течение месяца минус расход на производственные и другие нужды в этом же периоде.

#### 4.4. ОРГАНИЗАЦИЯ КОНТРОЛЯ ДОСТОВЕРНОСТИ ДАННЫХ В УСЛОВИЯХ МЕХАНИЗИРОВАННОГО СБОРА И ОБРАБОТКИ ИНФОРМАЦИИ

Сегодняшний этап функционирования АСГС основывается на использовании современных ЭВМ, а также автоматических средств регистрации и передачи данных, обеспечивающих широкую автоматизацию всех этапов сбора и обработки статистической информации. В этих условиях контроль достоверности данных является неотъемлемым элементом технологического процесса машинной обработки статистической информации.

В зависимости от этапа машинной обработки статистических данных можно выделить следующие стадии контроля достоверности информации: а) контроль в процессе приема входной информации; б) контроль в процессе подготовки машинных носителей входной информации; в) контроль ввода данных в ЭВМ; г) контроль результатов обработки данных.

Статистическая информация может поступать на ВЦ в различной форме: в виде статистических отчетов, присланных по почте; в виде перфолент, полученных в процессе передачи данных по каналам связи; в виде перфолент и магнитных лент, заготовленных на других ВЦ; по каналам связи в режиме автоматической передачи информации от ЭВМ одного ВЦ в ЭВМ другого.

Статистические отчеты от предприятий и организаций поступают в статуправление области. Приемка отчетов осуществляется экономистами отделов подготовки и выпуска статистических материалов. При этом отчеты подвергаются регистрации и контролю, организация и техника которого зависит от способа получения отчета (почтовый или телеграфный).

При контроле почтовых отчетов проверяются полнота их заполнения и наличие необходимых подписей, соответствие плановых и фактических данных, итоги с начала года, допускаемые пределы относительных и других рассчитываемых показателей, их взаимосвязи, правильность подсчета итоговых показателей и исчисления относительных величин и т. п. Как отмечалось выше, многие из этих операций могут выполняться на ЭВМ. Функции экономиста сводятся к внешнему осмотру документа на предмет наличия всех необходимых показателей и подписей.

При контроле телеграфной отчетности экономист проверяет правильность представления в телеграмме всех кодов и наличие необходимых показателей. В случае обнаружения ошибки делается запрос предприятию, вносятся исправления в телеграмму, заполняется бланк корректировки.

При поступлении информации в виде перфолент в процессе автоматической передачи данных по каналам связи на ВЦ республиканского и союзного уровней одновременно поступает и текстовая дубликат в виде телеграммы. После регистрации принятой информации перфолента передается на ЭВМ для контроля. Отчеты на магнитных лентах также контролируются на ЭВМ. Контроль производится путем сравнения полученных итогов с данными одновременно присланных сводных таблиц.

Контроль правильности переноса данных с поступивших статистических отчетов на машинные носители (перфокарты, перфоленты) осуществляются операторами в отделе подготовки машинных носителей ВЦ. Суть контроля заключается в следующем. До переноса данных на машинные носители по каждому статистическому отчету подсчитываются контрольные суммы по каждой строке или графе. Суммирование сопровождается занесением всех реквизитов на бумажную ленту, которую прикрепляют к отчету. После этого отчеты поступают на перфорацию, т. е. на операцию переноса данных на перфокарты или перфоленты. Полученные машинные носители контролируются. Метод контроля зависит от вида носителя.

Перфокарты проверяются методами счетного контроля, верификации, комбинированным методом.



## ГЛАВА 5

СОВОКУПНОСТЬ КАК ОБЪЕКТ  
СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ5.1. СТАТИСТИЧЕСКАЯ СОВОКУПНОСТЬ И ЗАДАЧИ  
ЕЕ СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ

Статистика всегда изучает массовые явления, определенное множество объектов, которые принято называть совокупностью единиц. Совокупности можно называть статистическими, если они обладают следующими характерными чертами: объективностью существования; качественной однородностью образующих их явлений, единиц; варьированием изучаемых признаков в пространстве и во времени.

Решая практические или научные задачи, статистика оперирует фактическими данными, при этом устанавливается, какой процесс, какая общая совокупность должна быть предметом ее изучения; но границы этой совокупности, круг единиц, входящих в нее, их качественное своеобразие, характерные признаки объективно заданы, т. е. предмет статистики реально существует. Например, ставится задача определения структуры рабочих кадров на конкретном предприятии. В процессе ее решения устанавливается лишь набор признаков, по которым будет изучаться структура рабочих кадров (пол, возраст, образование, стаж работы, и др.), а исследуемая совокупность — коллектив рабочих предприятия — объективно существует, границы совокупности предопределены рамками данного предприятия. Изучаемые признаки, характеризующие структуру рабочих кадров, также реально существуют.

Отдельно взятые единицы любой статистической совокупности характеризуются множеством различных свойств и особенностей, участвуют в различных по своей природе процессах, но в определенном отношении обладают общими свойствами, объединены общим законом развития и в этом отношении являются однокачественными, выступают как единицы соответствующей совокупности. Так, совокупность студентов многообразна с точки зрения социальных, биологических, семейных, экономических и других условий, но качественно однородна с точки зрения процесса учебы, в результате которого формируются специалисты высокой квалификации. Еще пример. Деревообрабатывающие предприятия страны можно рассматривать как однокачественную статистическую совокупность с точки зрения вида обрабатываемого материала и специфических отраслевых

Счетный метод контроля заключается в том, что сравниваются итоги контрольной табуляграммы, полученной в результате распечатки содержания и суммирования контролируемых реквизитов полученных перфокарт, с контрольными суммами, зафиксированными ранее на бумажных лентах. Этот метод является наиболее эффективным, так как дает возможность выявить до 99 % ошибок.

Методом верификации осуществляется контроль перфокарт на специальных машинах — контроллинках. Он имеет свои достоинства: меньшая (почти вдвое) по сравнению со счетным методом контроля трудоемкость, простота организации работы. Однако степень точности этого метода ниже, чем при счетном контроле. В среднем при этом методе обнаруживается до 93 % ошибок.

Комбинированный метод контроля перфокарт основывается на одновременном применении двух вышеуказанных методов. Например, менее важные реквизиты проверяются на контроллинках, остальные — методом счетного контроля.

Правильность переноса данных на перфоленты проверяется методами визуального контроля, повторного набора данных, сравнения дублей перфолент, «нуль-контроля», контроля на ЭВМ.

При визуальном контроле контрольная лента, содержащая распечатку содержания перфоленты, сравнивается с данными статистического отчета. Этот способ довольно трудоемок и сравнительно мало надежен, поэтому его применяют только для проверки небольшого объема информации.

Контроль методом повторного набора осуществляется по принципу контроля перфокарт методом верификации с помощью устройства подготовки данных на перфоленте (УПДЛ).

Контроль дублей перфолент заключается в сравнении на УПДЛ пробитой перфоленты с ее копией, отперфорированной другим оператором.

«Нуль-контроль» проверяет правильность пробитки ленты в процессе самой перфорации. Суть контроля заключается в достижении нулевой разности между суммой набранных на клавиатуре показателей и контрольной суммой, зафиксированной в одном из счетчиков машины.

Правильность переноса данных на перфокарты и перфоленты можно проверить и с помощью ЭВМ по определенной программе. В этом случае в составе контролируемых перфокарт или перфолент должны быть машинные носители с контрольными суммами. Контроль осуществляется при вводе информации в ЭВМ. Принцип его аналогичен реализации метода «нуль-контроля». При вводе в ЭВМ данные каждой строки (графы) суммируются машиной и сравниваются с контрольными суммами. В случае несовпадения машина печатает протокол ошибок.

Организация и техника контроля ввода данных в ЭВМ, а также контроля результатов обработки данных рассматриваются в гл. 6.

характеристик, но каждое предприятие отличается от другого численностью персонала, объемом производства, техническим уровнем, составом производимой продукции и др. Следовательно, существование каждой из единиц совокупности связано с множеством процессов, но только в одном из них они выступают как единицы соответствующей статистической совокупности.

Изучая объективно существующие совокупности, статистика решает различные познавательные задачи, что вызывает необходимость рассматривать каждую единицу совокупности в разных аспектах, с различных позиций. Так, в процессе исследования производительности труда на предприятии статистика рассматривает совокупность рабочих с точки зрения их стажа работы, квалификации, уровня образования, возраста и др. Круг перечисленных признаков присущ всем единицам совокупности, но при этом количественные значения признака у каждого рабочего различны, что в конечном итоге влияет на формирование уровня производительности труда. Поэтому совокупность рабочих предприятия с позиций их участия в едином производственном процессе качественно однородна, а в отношении результата труда, выражающегося в производительности, единицы совокупности неоднородны. Другой пример — социалистические сельскохозяйственные предприятия. Их природа едина. Однако существуют различия в отношении форм собственности между совхозами и колхозами, существуют предприятия разного производственного направления, разной эффективности и т. д.

Приведенные примеры позволяют сделать вывод об объективном подразделении предмета статистических исследований — совокупностей явлений, подчиненных единым закономерностям, на своеобразные по своему характеру так называемые частные совокупности или типы. Предметом статистического изучения являются не просто совокупности явлений, а общие совокупности, состоящие из нескольких частных совокупностей. Единицы, относящиеся к отдельным частным совокупностям, имеют единый круг признаков, определяющих их качество, а количественные значения этих признаков оказываются близкими друг к другу. Таким образом, частные совокупности качественно и количественно однородны. При этом под качественной однородностью следует понимать принадлежность единиц к одному определенному типу, что обеспечивается наличием в равной степени у всех единиц совокупности основных, характерных признаков данного типа. Под количественной однородностью понимается близость количественных значений основных признаков единиц совокупности друг к другу.

Однако однородность единиц, составляющих совокупность, не является чем-то абсолютным, навсегда заданным. Согласно закону динамического развития изменение количественных значений признаков меняет качественное содержание единицы совокупности и со временем состав совокупности изменяется: отдельные единицы, не соответствующие данному качеству, исклю-

чаются из состава совокупности, другие, наоборот, вновь входят в ее состав.

Существование частных совокупностей всегда имеет место, и игнорировать этот факт нельзя. Чтобы правильно понять и раскрыть своеобразие природы общей совокупности, необходимо иметь полное представление о системе образующих ее частных совокупностей, без этого не могут быть решены важнейшие задачи статистики — характеристика структуры совокупности, изучение взаимосвязей между явлениями. Так, характеризуя процесс выполнения производственной программы предприятиями отрасли, общую совокупность подразделяют, как правило, на три частные: предприятия, не выполнившие производственную программу, выполнившие ее на 100 % и перевыполнившие программу. А далее, в пределах выделенных частных совокупностей выясняются причины отклонений фактического выполнения производственной программы от установленных плановых заданий.

Следует иметь в виду, что единицы общей совокупности по-разному объединяются в частные статистические совокупности в зависимости от характера изучаемого процесса. Частных совокупностей может быть несколько, при этом либо все без исключения единицы общей совокупности распределяются по определенным частным, либо выделение частных совокупностей охватывает лишь часть единиц всей совокупности. Например, при изучении трудовых ресурсов выделяются лица трудоспособного возраста; при характеристике уровня образования выделяется население от 10 лет и старше и т. п.

Как отмечалось ранее, единицы совокупности обладают многими свойствами и особенностями, выражающими их характер и связи с окружающим миром. Эти свойства, особенности единиц называются в статистике признаками. Признаки — непосредственный объект статистического измерения. Только через них можно охарактеризовать природу совокупности, уловить ее закономерности.

Статистические признаки подразделяются на существенные и несущественные (второстепенные), статические и варьирующие, меняющиеся в пространстве и времени, атрибутивные (качественные) и количественные, первичные и вторичные.

Существенные признаки выражают сущность совокупности, определяют ее однокачественность и тем самым обособляют от окружающих явлений. Несущественные признаки не характерны для качества совокупности, не показательны для определяющих ее закономерностей. Но они воздействуют на существенные признаки, связаны с ними, дают дополнительные сведения о свойствах явлений. Так, например, для совокупности промышленных предприятий, входящих в состав определенной отрасли, существенными признаками являются: назначение продукта, вырабатываемого предприятием, вид перерабатываемого сырья, характер технологического процесса. Для выделения из этой совокупности крупных предприятий необходимо знать размеры основных

фондов, выпуск продукции, численность промышленно-производственного персонала. Перечисленные признаки определяют однокачественность совокупности, т. е. принадлежность их к группе крупных предприятий. В этом отношении их следует считать существенными. Одновременно деятельность предприятий характеризуется и другими признаками (степень выполнения плана по объему производства, уровень технической оснащенности и производительности труда, степень использования основных производственных фондов, уровень эффективности производства, текучесть рабочих кадров и пр.), которые оказывают значительное воздействие на размеры производства, численность персонала, стоимость основных производственных фондов. Однако четкой сопряженности значений этих несущественных признаков с существенными нет. Это характеристики, которые отображают условия функционирования предприятий, но не показывают прямо их масштаб производства.

Устойчивость и изменчивость предмета статистического изучения обуславливает деление признаков на варьирующие и статические. Варьируют признаки могут в пространстве, при переходе от одной единицы совокупности к другой, и во времени. Например, стаж работы отдельных рабочих предприятия может принимать различные значения от 0 до 30 и более лет, изменяется он и с течением времени.

Статические признаки имеют неизменные значения у всех единиц совокупности и характеризуют изучаемый объект в относительной устойчивости. Характеристика статических признаков не требует обобщения, она может быть получена измерением один раз у одной единицы. Такие признаки статистику не интересуют. Статистика изучает только *варьирующие признаки*. Значения признака у отдельных единиц совокупности называются *вариантами*; они принимают любые значения в пределах границ вариации данного признака и могут повторяться при переходе от одних единиц к другим.

Свойства общественных явлений могут характеризоваться атрибутивными (качественными) и количественными признаками.

Атрибутивным называется такой признак, варианты которого, характеризуя особенности отдельных единиц, не имеют количественного выражения. Примером таких признаков являются: национальность, социальная принадлежность людей, профессия рабочего и т. п.

Количественными называются признаки, варианты которых выражаются числовой характеристикой или мерой. Примером этого рода признаков могут служить: процент выполнения плана предприятием, размер посевных площадей колхозов, заработная плата рабочих, стаж работы и др.

Признаки подразделяются также на первичные и вторичные. Такое деление уточняет структуру совокупности.

Первичные признаки непосредственно характеризуют единицу совокупности в целом, как нечто единое, неделимое. Так, первичными признаками предприятия являются: число рабочих, стоимость основных производственных фондов, мощность электромоторов, общие затраты на производство продукции и т. д.

Вторичные признаки определяют структуру единиц совокупности, отражают различные стороны их существования и выступают либо как относительные характеристики, либо как средние из значений того или иного сложного элемента. Для промышленного предприятия такими признаками будут: стоимость основных фондов на одного рабочего, затраты на единицу продукции и т. д.

Рассмотренные выше свойства и особенности статистической совокупности являются отправными пунктами при решении конкретных задач статистического изучения массовых общественных явлений и процессов. Основываясь на принципах и методах научного познания, разработанных диалектическим и историческим материализмом и на положениях марксистско-ленинской политической экономии, статистика решает множество разнообразных задач. Исследовать методами статистики массовое явление это значит, во-первых, вести наблюдение, учет и сбор разнообразных фактов, элементов совокупности или наблюдать само явление в множестве его повторений в пространстве и во времени, регистрировать результаты наблюдений с помощью специфических характеристик. Во-вторых, обобщать и анализировать собранные сведения с целью выявления и измерения закономерностей, которые, как правило, непосредственно не улавливаются, скрыты в многообразии массы первичных данных. В результате обобщения и анализа первичных данных статистика получает сведения о размерах совокупности или средних показателях, о процессах, характерных для исследуемой совокупности, о тенденциях их изменения, о структуре совокупности, о взаимосвязях между процессами.

## 5.2. ТРЕБОВАНИЯ К ФОРМИРОВАНИЮ СТАТИСТИЧЕСКОЙ СОВОКУПНОСТИ

Количественная характеристика массового явления не может быть основана на механическом образовании совокупности. При ее формировании очень важно соблюдать определенные требования, в противном случае выводы, полученные в процессе анализа, не будут реальными.

Важным требованием при сборе статистических данных об объектах однородной совокупности является их *сопоставимость*. Причем сопоставимость в широком смысле слова. Это значит, что объекты, образующие совокупность, должны быть тождественны по изучаемым показателям, методологии их исчисления и единицам измерения, сравнимы во времени и пространстве.



Так, сравниваемые показатели должны быть получены на основе данных, собранных по одинаковой программе, аналогичным способом (опрос, непосредственное или документальное наблюдение и т. п.); исследуемые показатели должны быть получены в результате применения одинаковых приемов счетной обработки — ручных или механизированных, с применением арифметического или логического контроля, использованием тождественных алгоритмов расчета и т. д. Очень важно, чтобы показатели, характеризующие единицы совокупности, были исчислены по единой методологии. Нельзя сравнивать уровни производительности труда по предприятиям отрасли, если этот показатель на одних предприятиях исчисляется в расчете на одного рабочего, а на других в расчете на одного работающего. Урожайность сельскохозяйственных культур может быть исчислена на 1 га площади, занятой посевами к концу весеннего сева, и на 1 га фактически убираемой площади. Такие показатели несопоставимы по колхозам региона, особенно в тех случаях, когда летом была гибель сельскохозяйственных культур.

Важным требованием тождественности показателей по объектам является одинаковость единиц измерения. Если данные имеют различную размерность, ни о какой согласованности единиц совокупности не может быть и речи. При этом одинаковые по названию единицы измерения не совпадают по содержанию в явлениях главного компонента. Так, могут оказаться неравноценными одинаковые количества молока, сахарной свеклы, руды и других видов продукции по содержанию в них главного компонента, например количества жира (в молоке), сахара (в свекле), металла (в руде). Не тождественны даже тонны нефти и угля, так как у них различен физико-химический состав. В целях обеспечения сопоставимости используют условные натуральные единицы измерения (гл. 8).

Объекты, имеющие стоимостное измерение, оцениваются в рублях, но при этом могут использоваться денежные измерители различного содержания. Так, цены на промышленную продукцию могут быть оптовыми ценами предприятий либо промышленности; государственные розничные цены могут подразделяться по зонам. Кроме того, цены могут быть фактически действующие и сопоставимые («неизменные»), которые со временем подвергаются пересмотру. Перечисленные признаки тождественности показателей относятся к сопоставимости единиц совокупности во времени и пространстве. Наряду с этим имеются требования, относящиеся к сопоставимости либо во времени, либо в пространстве.

Одним из требований сопоставимости единиц во времени выступает одинаковость территории, охватываемой совокупностью. Это требование нельзя понимать буквально. Вопрос о сопоставимости по территории решается конкретно, в зависимости от задач исследования. Скажем, чтобы установить, за счет каких факторов изменилось население области, границы кото-

рой менялись в изучаемом периоде, необходимо сравнивать численность населения, проживающего в данной области, в различные последовательные промежутки времени в одних и тех же границах. Если же нас интересует изменение численности населения области как административной единицы и, следовательно, с учетом изменения ее территории, то надо сравнивать численность населения на определенные моменты времени, взятые для различных территориальных границ.

Сопоставимость единиц совокупности в пространстве вызывает требования *исчисления показателей на одну и ту же дату либо за одно и то же время*. В. И. Ленин считал это правило самым элементарным условием всякого добросовестного употребления статистики<sup>1</sup>. Так, очевидно, не могут быть сопоставимы результаты деятельности предприятия, если в одном случае они характеризуются показателями, исчисленными за один месяц, а в другом — за полугодие.

Второе важное требование формирования статистической совокупности — *полный, исчерпывающий охват объектов наблюдения*. Анализ отдельных случайных фактов приводит к ложным выводам, искажает существующие закономерности. Важнейшая черта предмета статистического познания — исчерпывающее воспроизведение исследуемой области жизни. Фундамент фактов, который создает статистика для социального познания, должен, как указывал В. И. Ленин, включать в себя «*всю совокупность*» относящихся к рассматриваемому вопросу фактов, *без единого исключения*, обеспечивая учет «объективной связи и взаимозависимости исторических явлений в их целом»<sup>2</sup>. «В области явлений общественных нет приема более распространенного, как *выхватывание отдельных фактиков*, игра в примеры. Подобрать примеры вообще — не стоит никакого труда, но и значения это не имеет никакого, или чисто отрицательное, ибо все дело в исторической конкретной обстановке отдельных случаев. Факты, если взять их в их *целом*, в их *связи*, не только „упрямая“, но и безусловно доказательная вещь. Фактики, если они берутся вне целого, вне связи, если они отрывковы и произвольны, являются именно только игрушкой или кое-чем еще похуже...»<sup>3</sup>.

В случае использования неполных данных надо добиваться, чтобы они были *представительными*, типичными, т. е. включали примерно пропорциональное число представителей всех рассматриваемых типов. Особенно важно, чтобы типичными были динамические данные (см. гл. 14).

Третье требование, предъявляемое к формированию статистической совокупности — *достоверность и своевременность* данных, используемых в процессе анализа совокупности, о чем подробно говорилось в гл. 4 данного учебника.

<sup>1</sup> См.: Ленин В. И. Полн. собр. соч., т. 23, т. 34.

<sup>2</sup> Там же, т. 30, с. 351.

<sup>3</sup> Там же, с. 350.

Природа и содержание статистических показателей соответствует тем экономическим и социальным явлениям и процессам, которые они отражают. Процесс общественного воспроизводства определяется экономическими и социальными законами, действие которых раскрывается в соответствующих экономических и социальных категориях. Например, для характеристики действия основного экономического закона социализма используют такие категории, как совокупный общественный продукт, национальный доход, экономическая и социальная эффективность производства, потребление материальных и духовных благ и др. Все экономические и социальные категории или понятия, открытые и научно обоснованные марксистско-ленинской теорией, носят абстрактный характер, отражают наиболее существенные черты социалистического воспроизводства, общие взаимосвязи и взаимозависимости его элементов. Однако для того, чтобы измерить размеры и соотношения явлений или процессов, т. е. дать им соответствующую количественную характеристику, нужно разработать экономические и социальные показатели, соответствующие каждой категории (понятию). Именно соответствием показателей сущности экономических категорий обеспечивается единство количественной и качественной характеристик экономических и социальных явлений и процессов.

Различают два вида показателей экономического и социального развития общества: плановые и отчетные (статистические). Плановые показатели представляют собой определенные конкретные задания по развитию народного хозяйства и культуры страны в целом или отдельных отраслей, объединений, предприятий и организаций на определенные периоды времени. Отчетные показатели характеризуют реально сложившиеся условия экономического и социального развития, фактически достигнутый уровень за определенный период.

Итак, статистический (отчетный) показатель — это объективная количественная характеристика (мера) общественного явления или процесса в его качественной определенности в конкретных условиях места и времени. Каждый статистический показатель имеет качественное, социально-экономическое содержание и связанную с ним методологию измерения. Так, при проведении переписи населения нужно четко установить, какие именно лица учитываются в составе населения данного населенного пункта: т. е. кто оказался в данном пункте на критический момент (наличное население) или постоянно проживающим (постоянное население), кого считать временно проживающим и кого — временно отсутствующим и т. п.

Статистический показатель имеет также ту или иную статистическую форму (структуру): он может выражать общее число единиц совокупности либо общую сумму значений количественного признака этих единиц, а также среднюю величину

признака, величину данного признака по отношению к величине другого и т. п.

Статистический показатель имеет также определенное количественное значение или численное выражение. Это численное значение статистического показателя, выраженное в определенных единицах измерения, называется его *величиной*.

*Величина показателя* обычно варьирует в пространстве и колеблется во времени. Поэтому обязательным атрибутом статистического показателя являются также указание территории (или совокупности) и момента либо периода времени. Известно, например, что производство молока в личных подсобных хозяйствах населения СССР в 1981—1983 гг. составляло в среднем ежегодно 24,6 млн. т. Здесь качественное содержание показателя — производство молока; территория и совокупность — личные подсобные хозяйства населения СССР; числовые значения (величина) показателя — 24,6; единица измерения — млн. т; время — 1981—1983 гг.; статистическая форма показателя — среднегодовой уровень производства. Есть и определенная методология исчисления этого показателя: в данные о производстве молока включается все падающее молоко коров, овец, коз, кобыл, в том числе молоко, израсходованное на выпойку телят; молоко, высосанное телятами у матери, в объем производства молока не включается. Из этого примера видно, насколько важными и сложными являются четкое определение качественного содержания и методология измерения статистического показателя.

Статистические показатели (как и признаки) можно условно подразделить на первичные (объемные, количественные, экстенсивные) и вторичные (производные, качественные, интенсивные). *Первичные* характеризуют либо общее число единиц совокупности, либо сумму значений какого-либо их признака (см. выше). Взятые в динамике, в изменении во времени они характеризуют экстенсивный путь развития экономики и культуры. По статистической форме эти показатели являются суммарными абсолютными величинами (см. гл. 8). *Вторичные*, производные показатели обычно выражаются средними и относительными величинами и взятые в динамике обычно характеризуют путь интенсивного развития (повышение эффективности использования ресурсов, например рост производительности труда, снижение материалоемкости и трудоемкости единицы продукции и ее себестоимости и т. п.).

Показатели, характеризующие размер сложного комплекса социально-экономических явлений и процессов, часто называют *синтетическими* (совокупный общественный продукт, национальный доход, производительность общественного труда и др.).

В зависимости от применяемых единиц измерения различают показатели: натуральные, стоимостные и трудовые (выраженные в человеко-часах, нормо-часах). В зависимости от *сферы применения* различают показатели, исчисленные на народнохозяй-

ственным, региональном, отраслевом уровнях и т. д. По *точности отражаемого явления* различают ожидаемые, предварительные и окончательные величины показателей. Например, в сельскохозяйственном производстве определяют ожидаемый, предварительный и окончательный размеры валового сбора (урожая) и урожайности отдельных сельскохозяйственных культур. При расчете показателей экономического эффекта определяют их ожидаемую и фактически полученную (окончательную) величины.

В зависимости от *объема и содержания объекта* статистического изучения различают *индивидуальные* (характеризующие отдельные единицы совокупности) и *сводные (обобщающие)* статистические показатели.

Если исследуются такие сложные экономические и социальные категории, как эффективность производства, качество работы, уровень жизни народа и другие, то объективное и достоверное измерение их возможно только на основе системы статистических показателей, которые в единстве и взаимосвязи характеризуют различные стороны и аспекты состояния и динамики развития этих категорий. В настоящее время в соответствии с решениями апрельского и октябрьского Пленумов ЦК КПСС (1985 г.) ставится задача дальнейшей разработки системы показателей хозяйственного и социального механизма, совершенствование которого является важнейшим условием улучшения планирования и управления народным хозяйством, обеспечения роста эффективности общественного производства и благосостояния советского народа.

Статистические показатели, объективно отражая действие экономических законов социализма, единство и взаимосвязи экономических и социальных явлений и процессов, не могут быть надуманными или произвольно сконструированными, какими-то догмами, установленными раз и навсегда. Наоборот, динамичное развитие общества, возрастание роли научно-технического прогресса в процессе постоянного развития производительных сил и совершенствования производственных отношений с объективной необходимостью приводит к тому, что устаревшие, потерявшие свое значение показатели изменяются либо исчезают и появляются новые, более совершенные показатели, объективно и достоверно отражающие современные условия общественного развития. Так, только за последние годы внедрены новые показатели измерения производительности труда (на основе нормативно-чистой продукции), оценки деятельности предприятий (по реализованной продукции с учетом выполнения плана поставок) и т. д. Таким образом, процесс совершенствования методов определения статистических показателей, разработки их системы является постоянным и это составляет одну из главных задач статистической теории и методологии.

Построение и совершенствование статистических показателей должны основываться на соблюдении двух основных принципов:

1) объективности и реальности (показатели должны правдиво и адекватно отражать сущность соответствующих экономических и социальных категорий (понятий) и 2) всесторонней теоретической и методологической обоснованности (определение величины показателя, его измеримость и сопоставимость в динамике должны быть всесторонне научно аргументированы, четко и доступно сформулированы и однозначно, в единообразном толковании применимы). Это предполагает обеспечение полной достоверности показателей, величины которых исчисляются и измеряются только на основе достаточного объема и качества первичных данных учета, отчетности и других источников информации. При этом величины показателей должны правильно количественно измеряться с учетом уровня, масштабов и качественных признаков состояния или развития соответствующего экономического или социального явления (отраслевой и региональной úrovни; отдельное предприятие или работник и т. п.). Вместе с тем построение показателей должно носить сквозной характер, позволяющий не только суммировать соответствующие показатели, но и обеспечивать их качественную однородность в группах и совокупностях, переход от одного показателя к другому для полной характеристики объема и структуры более сложной категории или явления. Наконец, построение статистического показателя, его структура и сущность должны предусматривать возможность всесторонне анализировать изучаемое явление или процесс, характеризовать особенности его развития, определять влияющие на него факторы и выявлять резервы повышения эффективности.

Наряду с перечисленными выше общими требованиями к каждому статистическому показателю могут предъявляться специфические требования, обусловленные сущностью и особенностями развития изучаемого объекта или явления, а также конкретными задачами его исследования.

Между показателями плана и статистики существует органическое, неразрывное единство, так как, по существу, они отражают количественную сторону одних и тех же конкретных массовых экономических и социальных явлений и процессов, происходящих в общественном развитии. Это единство заключается также и в том, что плановые и статистические (отчетные) показатели исчисляются и анализируются по единой научной методологии, они соизмеримы и сопоставимы. Вместе с тем между плановыми и статистическими показателями имеются и различия. Если плановые показатели представляют собой задания по дальнейшему развитию данного объекта или явления и носят директивный, законодательный характер, то статистические (отчетные) показатели отражают фактически достигнутый уровень, реальное положение вещей, ход выполнения плановых заданий. Статистические показатели служат необходимой базой для разработки всех видов народнохозяйственных планов и для различных прогнозных расчетов. С другой стороны, статисти-



ческие показатели более подробно и разносторонне, чем плановые показатели, характеризуют общественные явления, отражают существенные обстоятельства и важные факторы, необходимые для объективного анализа выполнения планов, итогов деятельности предприятий и отраслей. Поэтому круг показателей статистики значительно шире круга показателей плана, что позволяет проанализировать факторы, обеспечившие успехи в работе предприятия или причины невыполнения планов, выявить внутренние резервы повышения эффективности производства и качества труда. Так, предприятиям и стройкам не планируют простой оборудования и строительной техники, прогулы рабочих или брак и недолетки в работе, однако в статистическом учете и отчетности эти явления отражаются и их анализ позволяет сделать необходимые выводы для улучшения работы предприятия. Кроме того, статистические показатели отражают ряд важных социальных и демографических процессов и явлений, которые вообще не могут быть предметом для планирования, например показатели воспроизводства населения, заболеваемость, аморальные явления и др.

Изложение настоящего курса общей теории статистики основано на широком использовании способов и приемов марксистско-ленинской статистической методологии для определения и анализа статистических показателей, характеризующих различные стороны и аспекты экономического и социального развития нашей страны.

## ГЛАВА 6 СВОДКА И ГРУППИРОВКА

### 6.1. СОДЕРЖАНИЕ И ЗАДАЧИ СВОДКИ. ЭТАПЫ СВОДКИ

В результате статистического наблюдения получают материал, характеризующий отдельные единицы совокупности. Эти единицы сложны, обладают многими признаками, изменяются во времени, одни отмирают и исчезают, другие утверждаются и развиваются. В подобных условиях невозможно использовать материал наблюдения для обобщающей характеристики статистической совокупности. Общее теряется в единичном, необходимое и закономерное — в несущественном и случайном. Возникает необходимость в специальной обработке статистических данных — сводке материалов наблюдения.

Сводка представляет собой комплекс последовательных действий по обобщению конкретных единичных данных, образующих совокупность в целях выявления типичных черт и закономерностей, присущих изучаемому явлению в целом. При этом все несущественное должно быть опущено, все существенное и основное — выделено и зафиксировано. Поэтому сводка должна проводиться на основе всестороннего теорети-

ческого анализа изучаемого явления. При неумелой, непродуманной сводке богатый первичный материал теряется, превращается, по словам В. И. Ленина, в «бессмысленные столбы цифр, статистическую «игру в цифирки», вместо осмысленной статистической обработки материала»<sup>1</sup>.

Правильная, научная организация сводки, основанная на предварительном глубоком теоретическом анализе, позволяет выявить сложившиеся или складывающиеся закономерности, обнаружить взаимосвязи, измерить влияние различных факторов на результат и учесть все это в практической работе, при составлении текущих и перспективных планов.

Следовательно, задача сводки — охарактеризовать исследуемый предмет с помощью систем статистических показателей, выявить и измерить таким путем его существенные черты и особенности.

Эта задача решается на трех этапах. Сначала осуществляется систематизация, группировка материалов, собранных при наблюдении. Это дает основу для решения задач исследования. Далее, на втором этапе уточняется предусмотренная планом система показателей, с помощью которых количественно характеризуются свойства и особенности изучаемого предмета. Это трудный и ответственный этап, который требует от исследователя глубоких знаний в области теории и практики статистики. Так, например, показатель производительности труда (выработки) на предприятии может быть рассчитан в различных вариантах и при этом будет иметь разную количественную оценку. Поэтому в соответствии с целью и задачами исследования нужно решить вопрос, какой из показателей выработки отвечает поставленным условиям и должен быть рассчитан по данным наблюдения. Эти вопросы может правильно решить только хорошо подготовленный специалист. Третий этап сводки — исчисление запрокированной системы показателей и их обобщение. При этом выверяются и подготавливаются к обработке (ручной или машинной) исходные статистические данные, проводятся подсчеты и вычисления, а затем исчисленные показатели для наглядности и удобства представляются в таблицах, статистических рядах, графиках. Далее к таблицам и графикам составляются текстовые пояснения, дополняющие их определенными сведениями, разрабатываются аналитические тексты, отражающие общие итоги, выявленные закономерности, характерные черты и связи изучаемого явления.

Перечисленные этапы сводки еще до начала ее проведения отражаются в специально составляемой программе. В самом общем виде программа статистической сводки содержит перечень групп, на которые требуется расчленить совокупность, их границы по установленным признакам группировки;

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 27, с. 182.

систему показателей, характеризующих совокупность, и методику их расчета; систему макетов разработочных таблиц, в которых будут представлены итоги расчетов.

*План проведения сводки* наряду с программой предусматривает ее организацию. В ряде случаев статистическая сводка весьма трудоемкая и многогранная работа, в выполнении которой участвует большое число организаций разной ведомственной подчиненности. Поэтому план проведения сводки крайне необходим. Он должен содержать указания о последовательности и сроках выполнения отдельных частей сводки, ее исполнителях и о порядке изложения результатов.

## 6.2. СТАТИСТИЧЕСКИЕ ГРУППИРОВКИ. РЯДЫ РАСПРЕДЕЛЕНИЯ

*Предмет статистических исследований* — массовые общественные явления — обладают многочисленными признаками и свойствами. Обобщить статистические данные, раскрыть наиболее существенные особенности, формы движения массового явления в целом и отдельных его составляющих невозможно без определенных научных принципов обработки данных.

Так, статистические органы, получив годовые отчеты промышленных предприятий, имеют в своем распоряжении богатый материал, характеризующий с разных позиций деятельность каждого предприятия в отдельности (объем произведенной продукции, ее состав и себестоимость, численность работников, производительность труда, размер фонда заработной платы, прибыли и др.). Для обобщенной характеристики деятельности всех промышленных предприятий в совокупности необходимо знать, как распределяются предприятия по объему выпускаемой продукции и степени выполнения плана, уровню технической оснащенности, по размеру снижения себестоимости и росту производительности труда, по уровню рентабельности и ряду других показателей. Для этой цели выделяются типические группы предприятий по соответствующим признакам. Однотипность предприятий обеспечивает в последующем необходимость для обобщения сопоставимости данных и реальную значимость статистических характеристик. Без преодоления индивидуального многообразия промышленных предприятий общие закономерности развития промышленности в целом теряются в деталях и мелочах, отличающих каждое предприятие одно от другого, а предельное обобщение влечет за собой извращенное представление о действительности. Для разделения совокупности единиц на отдельные группы статистика использует метод группировки.

*Статистические группировки* — первый и важный этап статистической сводки, позволяющий выделить из массы первичного статистического материала, представляющего собой ряды

цифр, однородные группы единиц, обладающие общим сходством в качественном и количественном отношениях. Группировки — это не субъективный технический прием расчленения совокупности на части. В зависимости от того, как произведена группировка, какими принципами руководствовался исследователь, разбивая совокупность на группы, можно сделать различные выводы, иногда крайне противоположные. Поэтому в процессе образования групп важно соблюдать единые научные требования.

Основопологающим принципом применения метода группировки является всесторонний, глубокий анализ сущности и природы изучаемого явления с позиций марксистско-ленинской философии и политической экономии, позволяющей определить его типические свойства и внутренние различия. Как отмечалось в гл. 5, любая общая совокупность является комплексом частных совокупностей, каждая из которых объединяет явления особого типа, однокачественные в определенном отношении. Каждый тип (группа) имеет специфическую систему признаков с соответствующим уровнем их количественных значений. Установить, к какому типу, в какую частную совокупность нужно отнести группируемые единицы общей совокупности, возможно на основе правильного, четкого определения существенных признаков, по которым должна проводиться группировка. Это второе важное требование научно обоснованной группировки. И наконец, третье требование группировки — объективное, обоснованное установление границ групп при условии, что образованные группы должны объединять однородные элементы совокупности, а сами группы (одна по отношению к другой) должны существенно различаться. В противном случае группировка теряет смысл.

Таким образом, на основе применения метода группировки определяются группы по принципу сходства и различия единиц совокупности. Сходство — это однородность единиц в определенных пределах (группах); различие — это их существенное расхождение по группам.

Учитывая перечисленные принципы научного подхода к образованию групп, можно дать следующее определение статистической группировки: *группировка — расчленение общей совокупности единиц по одному или нескольким существенным признакам на однородные группы, различающиеся между собой в качественном и количественном отношении и позволяющие выделить социально-экономические типы, изучить структуру совокупности или проанализировать связи между отдельными признаками.*

Первыми исследователями, применившими группировку в своих научных работах, были известные русские общественные деятели и ученые Д. П. Журавский, П. П. Семенов-Тянь-Шанский, А. П. Радшев, Н. П. Огарев.

Д. П. Журавский в известном исследовании «Об источниках и употреблении статистических сведений» провозгласил принципом статистики еди-

ство количественного и качественного анализа. Выделение и изучение «одно-видных» частей, категорий, групп он считал основным методом статистического анализа, а статистическую науку — наукой «о категорическом исчислении».

Подлинно научный, марксистский метод социально-экономических группировок впервые был разработан В. И. Лениным. В основе этого метода лежит изучение социально-экономической природы изучаемого явления, которая и определяет выбор признаков для группировки.

Приведем пример из работы В. И. Ленина «Развитие капитализма в России». По официальным данным в Европейской части России в 1897 г. по сравнению с 1863 г. численность городского населения увеличилась почти в 2 раза, а численность сельского — почти в 1,5 раза, доля городского населения за эти годы увеличилась с 10 до 13 %.

В. И. Ленин показал, что хотя эти данные для всей Европейской части России в целом правильны, но когда речь идет о развитии капитализма, нужно различать два процесса: во-первых, когда все земли заняты, вся территория заселена, тогда развитие капитализма осуществляется на той же земле «вглубь» и явление роста индустриального населения представлено в чистом виде; во-вторых, когда земельный фонд не весь занят, территория еще не вся заселена, развитие капитализма осуществляется на новой земле, т. е. «вширь», и поэтому увеличивается численность и сельского населения.

В. И. Ленин писал, что смешение этих процессов приведет к ошибочному представлению о том, что отвлекает население от земледелия. Поэтому В. И. Ленин выделил в губерниях Европейской части России три группы: 1) столичные, промышленные и неземледельческие губернии; 2) центральный земледельческий район и 3) земледельческие окраины.

За период с 1860 по 1883 г. в 1-й группе сельское население увеличилось на 25 %, во 2-й — на 35 %, в 3-й — на 87 %. Этот пример, как и многие другие, отчетливо показывает, что В. И. Ленин при выборе признаков группировок исходил прежде всего из социально-экономического существа изучаемых процессов, и это было тем принципиально новым, что он внес в метод статистических группировок.

В своих работах В. И. Ленин неоднократно указывал на необходимость правильного выделения круга существенных признаков, по которым должна проводиться группировка.

Для характеристики размера и типа хозяйства В. И. Ленин решительно отвергает группировку по формальным признакам, таким, как надел, число членов семьи, размер земельной площади. В. И. Ленин осуществляет группировку хозяйств по существенным признакам, характерным для типа и размера хозяйства: количеству рабочего скота, размерам посевов, стоимости продукции, размерам употребленного наемного труда. Это позволило получить яркую картину «придавленности, ни-

щеты, разорения, экспроприации массы мелких землевладельцев»<sup>4</sup>.

Чрезвычайно важным является указание В. И. Ленина на необходимость специализации признаков группировки в соответствии с различиями условий и форм функционирования единиц общей совокупности. Для выделения типов хозяйств В. И. Ленин предлагал брать группировочные признаки в соответствии с различиями местных условий и направлением хозяйств в отдельных районах. Так, в районах, где преимущественное развитие имеют, например, животноводство или огородничество, признаки для характеристики размеров и типов хозяйств должны быть взяты другие, чем в районах, где преобладают посевы зерновых культур.

Не менее важным является указание В. И. Ленина на ограниченность группировки по одному признаку. Он отмечал, что часто экономический тип невозможно выделить по одному основанию, отвлекаясь от других важных признаков; необходимо опираться на существенные и характерные признаки, учитывающие всю совокупность конкретных условий. Так, в статье «К вопросу о нашей фабрично-заводской статистике» В. И. Ленин посредством группировки фабрично-заводских предприятий по двум признакам: количеству рабочих, которое должно быть не меньше 15, и наличию парового двигателя, убедительно показал быстрое развитие капиталистической промышленности в России.

Ценным в методическом плане является ленинский подход к определению интервалов при группировке. Интервалы должны быть выражением развития совокупности, множества типов в ней. Наиболее правильным является выделение типов по сочетанию нескольких интервалов по ряду существенных признаков. Для разных частей совокупности интервалы должны браться специализированные. Все это предполагает, что интервалы должны определяться исходя из сущности изучаемого явления, а в конечном итоге группы должны иметь определенную качественную однородность. Следует отметить, что ленинские указания по вопросам правильного научно обоснованного применения группировок и в настоящее время имеют принципиальное значение в экономических исследованиях.

В статистической методологии, представляющей собой систему методов познания массовых общественных явлений, группировки занимают особое место. Это объясняется тем, что группировки, в отличие от других статистических методов, выполняют две функции: во-первых, являются самостоятельным методом статистического познания социально-экономических явлений, с помощью которого выделяются социально-экономические типы, изучается структура совокупности, вскрываются существующие взаимосвязи; во-вторых, являются приемом,

<sup>4</sup> Ленин В. И. Полн. собр. соч., т. 27, с. 208.



предопределяющим границы и возможности использования других статистических методов (средних, корреляционно-регрессионного анализа, речь о которых будет идти в последующих главах). Применение этих методов предполагает наличие однородных совокупностей, что обуславливает использование статистических группировок в качестве основополагающего, необходимого этапа обработки материалов статистического наблюдения.

Следовательно, группировка — это не просто один из самостоятельных методов, который можно использовать при статисти-

Таблица 6.1. Число семей в СССР и группировка их по размеру (по данным переписи населения 1979 г.)

	Городское и сельское население, млн.
Всего семей	66,3
том числе состоящих из совместно проживающих человек:	
2	19,7
3	19,1
4	15,2
5	6,3
6 и более	6,0

звояет решить три основные задачи: выделить социально-экономические типы, изучить структуру однородной совокупности, обнаружить существующие связи и зависимости между признаками исследуемого явления. В соответствии с этими задачами группировки подразделяются на три вида: типологические, структурные и аналитические. Однако следует отметить, что при решении конкретных задач анализа трудно установить границы между указанными видами группировок. Так, типологическая группировка часто служит основой изучения структуры совокупности, а последняя — базой изучения взаимосвязей между признаками (ниже это будет показано на примерах).

В зависимости от степени сложности изучаемого массового явления и от задач анализа группировки могут производиться по одному или нескольким признакам. Если группы образуются по одному признаку, группировка называется *простой*. Примерами простых группировок могут служить: распределение промышленно-производственного персонала завода по категориям работающих; распределение промышленных предприятий отрасли по степени выполнения плана; распределе-

ние населения по возрастным группам, а семей по их размеру (табл. 6.1).

В экономических исследованиях часто применяются группировки на основе двух или большего числа признаков, взятых в комбинации друг с другом. Группы, образованные по одному признаку, делятся на подгруппы по второму, а последние — на подгруппы по третьему признаку и т. д. Такая группировка называется *комбинационной*. Если приведенную

Таблица 6.2. Распределение семей в СССР по их размеру и принадлежности членов к общественным группам (по данным переписи 1979 г.)

	Число семей, млн.	в том числе семьи, состоящие из совместно проживающих человек:				
		2	3	4	5	6 и более
Все семьи	66,3	19,7	19,1	15,2	6,3	6,0
из них семьи:						
члены которых принадлежат к одной общественной группе	44,3	15,5	12,5	9,4	3,5	3,4
члены которых принадлежат к разным общественным группам	22,0	4,2	6,6	5,8	2,8	2,6

в табл. 6.1 группировку семей по размеру дополнить группировкой по признаку принадлежности членов семей к одной или разным общественным группам, получим пример комбинационной группировки (табл. 6.2).

Комбинационные группировки применяются при изучении сложного социально-экономического явления. Обязательное условие их применения — достаточно большое число наблюдений. Комбинация группировочных признаков резко увеличивает число групп, численность единиц в которых может быть незначительной. Поэтому выводы на основании небольшого числа наблюдений могут быть недостаточными обобщающими, особенно при изучении зависимости между признаками.

В последние годы проводятся исследования по поиску приемов группировки совокупности единиц одновременно по комплексу характеризующих признаков. Такие группировки получили название *многомерных*. Особого внимания заслуживает теория распознавания образов, суть которой состоит в том, что единицы совокупности представляются в виде точек в *n*-мерном пространстве (*n* — число признаков, характеризующих единицы наблюдений). Их объединение в группы (классы, таксоны) производится по принципу близости расположения в многомерном пространстве. Единицы совокупности в силу сходства по ряду признаков образуют как бы сгущение точек в многомерном признаковом пространстве. Эти сгущения, границы которых определяются ЭВМ на основе специально разработанных алгоритмов и критериев, принимаются за искомые группы. Формализация данного подхода к образованию групп по комплексу признаков очевидна, и это

в определенной мере нарушает методический принцип качественного анализа исследуемого явления, что особенно важно при типологической группировке. Однако при решении определенных задач изучения взаимосвязей, а также при использовании группировок как приема, предопределяющего границы применения других статистических методов, теория распознавания образов может иметь практическое применение.

Центральной задачей группировки В. И. Ленин считал выявление и характеристику социально-экономических типов. Это объясняется, во-первых, тем, что всесторонний экономический анализ невозможен без предварительного выделения социально-экономических типов; во-вторых, выявление социально-экономических типов представляет собой основу решения всех задач методом группировок. В статистике *типологической группировкой* принято считать разделение исследуемой совокупности на классы, социально-экономические типы, однородные группы единиц в соответствии с правилами научной группировки.

Решение задачи выявления и характеристики социально-экономических типов является делом далеко не легким. Для ее решения приходится особо учитывать всю совокупность фактов, относящихся к изучаемому явлению, все его свойства и черты. Выбирать из них те признаки и свойства, которые наиболее точно отразили бы его специфику, — значит творчески применять марксистско-ленинскую теорию в статистическом исследовании. Выделить типичное можно не по любому признаку, а только по определенному, который должен изменяться в зависимости от условий места и времени. Блестящие образцы постановки и решения этой основополагающей задачи типологической группировки мы находим в трудах В. И. Ленина. Так, изучая процесс развития капитализма в сельском хозяйстве, В. И. Ленин делит крестьянские хозяйства по экономической мощи на три группы: 1) хозяйства, основанные на эксплуатации, т. е. хозяйства капиталистического типа; 2) хозяйства середняцкие и 3) хозяйства бедняцкие. Это позволило В. И. Ленину убедительно показать процесс расслоения крестьянских хозяйств и зарождения капитализма в сельском хозяйстве России.

Весьма существенным вопросом при выборе группировочных признаков является предварительное установление типов, подлежащих выделению в ходе группировки. Для этого следует учитывать основное направление развития общей совокупности и те условия, в которых существуют единицы совокупности. Установление перечня типов, подлежащих выделению при группировке, требует от статистика глубокого знания конкретной действительности и глубоких теоретических знаний. Поскольку формирование типа связано с конкретными условиями места и времени, в которых существуют группируемые явления, один лишь теоретический анализ не позволяет определить перечень возможных типов. Невозможно устано-

вить его и эмпирически на основании лишь конкретных фактов, так как они не позволяют судить о сходстве и различии единиц совокупности, о существовании особых типов. Поэтому общетеоретические положения, определяющие природу исследуемого явления, при установлении типов должны дополняться конкретными данными статистических разработок, обобщения и анализа. Типы, установленные в ходе предварительного анализа, в последующем корректируются, уточняются. Реально существующие типы в рамках исследуемой общей совокупности определяются лишь в ходе группировки.

Природа типа может проявиться в одном признаке, тогда для выявления частных совокупностей достаточно учесть значение этого признака. Но очень часто своеобразие типов раскрывается в ряде признаков. При изучении сложных явлений в исследуемой совокупности могут быть одни и те же значения группировочного признака, которые в разных сочетаниях с другими признаками будут определять принадлежность единиц то к одной, то к другой частным совокупностям. В подобных случаях использование одного группировочного признака создает неопределенность границ между частными совокупностями. Если при выделении типов используются признаки, присущие только одной из выделяемых частных совокупностей, то для определения природы единицы как целого используется комплекс признаков.

Следует, однако, помнить, что комбинационная группировка по ряду признаков приводит к чрезмерно большому количеству групп, что создает сложности в процессе последующего анализа. Поэтому к отбору существенных признаков надо подходить с особой тщательностью. Так, В. И. Ленин при изучении земледелия в условиях капитализма считал возможным ограничиться комбинационной группировкой по двум признакам: «При современных методах исследования, дающих очень хороший, очень богатый запас сведений о каждом отдельном хозяйстве, было бы, например, достаточно соединить два метода группировки, — скажем, каждую из пяти групп по количеству всей земли подразделить на три, на две подгруппы по размеру употребления наемного труда»<sup>1</sup>.

Ограничивая круг группировочных признаков, надо учитывать их взаимосвязи. В области экономики весьма часто значение одного существенного признака является выражением принадлежности единиц совокупности к определенному типу и по другим существенным признакам. Поэтому даже по неполному кругу признаков часто можно выполнить достаточно точно типологическую группировку. Так, например, весьма показательным признаком для группировки предприятий по признаку передовые и отстающие является процент выполнения плана по выпуску продукции. Несмотря на то что суще-

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 27, с. 207—208.

ственными признаками упомянутого деления предприятий являются уровень производительности труда, степень использования основных производственных фондов, сокращение потерь рабочего времени, качество труда и другие, совсем не обязательно при группировке предприятий использовать весь комплекс признаков. Достаточно ограничиться основным признаком — процентом выполнения плана по выпуску продукции, так как значение этого показателя является выражением достижений предприятий и в других областях.

При выборе группировочных признаков надо учитывать их значение и функции в последующем анализе, а также условия места и времени. В связи с этим немаловажное значение имеет использование в разных условиях функционирования единиц совокупности различных группировочных признаков, т. е. их специализация. Предположим, исследуется технический уровень производства предприятий различных отраслей промышленности. Можно ли в этом случае выбрать единичные признаки для группировки предприятий всех отраслей? Безусловно, нет. Для металлургических предприятий технический уровень производства будет определяться характером технологии производства металла и мощностью печей; в станкостроении — степенью совершенства парка металлорежущего оборудования; на предприятиях швейной и обувной промышленности — наличием поточных линий и их мощностью и т. д.

При выделении социально-экономических типов используются атрибутивные и количественные признаки. Группировки по атрибутивному признаку базируются на положениях марксистско-ленинской политической экономии, и число выделяемых групп соответствует установленному числу разновидностей признака. Так, при группировке совокупного общественного продукта по экономическому назначению вся произведенная за год продукция делится на средства производства и предметы потребления; промышленно-производственный персонал предприятия группируется по общепринятым категориям: рабочие, служащие, инженерно-технические работники и т. д.

В статистической практике часто используются классификации. Провести четкую границу между классификацией и группировкой не всегда возможно, поскольку они выполняют односторонние функции. При классификации, в отличие от группировки, группировочные признаки predetermined, и вопрос об их выборе отпадает; четко сформулированы требования и условия отнесения единиц совокупности к той или иной группе. Иными словами, классификация — это общепринятый методологический стандарт расчленения совокупности на однородные группы, устанавливаемый на определенный промежуток времени. Такова классификация отраслей народного хозяйства, подразделяющая их на отрасли сферы материального производства и непроизводственные; классификация отраслей про-

мышленности и др. В отличие от этого группировка проводится только для целей конкретного исследования.

**Атрибутивный признак** — это качественный признак группировки. В ряде случаев его применение ограничено, и для выделения частных совокупностей используются количественные признаки. Так, для выделения крупных, средних и мелких промышленных предприятий используются признаки, имеющие определенное количественное выражение: объем производимой продукции, численность работающих, стоимость основных производственных фондов. При этом возникает проблема определения числа групп и их границ. Количественные значения признака, на основе которых исследуемые явления разбиваются на группы, называются в статистике интервалами. Смысл и значение интервалов в группировке зависит от ее конечной цели, от функций группировочного признака и взаимосвязи его с другими признаками, которые используются в данной группировке, от задач исследования, от особенностей общей совокупности. Учет всех этих условий предопределяет соответствующие приемы образования интервалов и групп.

Число интервалов при типологической группировке должно соответствовать числу предварительно намеченных частных совокупностей. Границы интервалов устанавливаются с таким расчетом, чтобы достигалась конечная цель группировки, а именно внутри групп единицы должны быть однородными в качественном и количественном отношении, но группы должны существенно отличаться одна от другой. Иными словами, нужно найти такие значения уровней признака, переход через которые означает переход к иному социально-экономическому типу.

Определение границ интервалов — задача трудная, часто требующая исследовательской работы. Для каждого группировочного признака по каждой частной совокупности необходимо провести анализ, основанный на глубоком знании природы совокупности. Примером такого анализа может слу-

Таблица 63<sup>1</sup>

Группы домохозяйств		% ко всему числу дворов		% посевной площади к итогу	
Бедная	{ Без рабочего скота	20,7	37,1%	2,8	8,0%
	{ С 1 голов. раб. скота	16,4		5,2	
Средняя	{ » 2—3 голов. »	» 26,6	38,2%	17,1	28,6%
	{ » 4 » »	» 11,6		11,5	
Богатая	{ » 5—10 » »	» 17,1	24,7%	26,9	63,4%
	{ » 10—20 » »	» 5,8		19,3	
	{ » 20 и более »	» 1,8		17,2	
Всего 100				100	

<sup>1</sup> См.: Ленин В. И. Полн. собр. соч., т. 3, с. 76 (в сокращении).



жить обработку данных земской статистики В. И. Лениным в работе «Развитие капитализма в России». Так, земская статистика по Новоузенскому уезду Самарской губернии подразделяла крестьянские хозяйства по количеству голов рабочего скота на семь групп, что явно скрывало три реально существующих типа хозяйств—бедные, средние и богатые. Объединяя отдельные группы, принятые земскими статистиками, В. И. Ленин выделяет три (см. табл. 6.3).

В отличие от земских статистиков В. И. Ленин выбирал интервалы не произвольно. Выделяя группы бедняцких хозяйств, В. И. Ленин учитывал их особенность—необеспеченность средствами производства, предполагающую продажу рабочей силы. Крестьянское хозяйство, не имеющее рабочего скота или имеющее одну голову, в условиях рассматриваемого района и времени не могло выполнять все сельскохозяйственные работы своими силами. Это и обусловило величину первого интервала. Ясны и мотивы определения границ третьей группы: пять и более голов рабочего скота предполагают использование чужого труда в хозяйстве.

Типологическая группировка на основе количественных признаков, имеющая цель отобразить качественное своеобразие групп, должна учитывать особенности развития и формирования частных совокупностей в конкретных условиях места и времени. В этом случае каждая группа, имея свое особое содержание, выделяется специализированными интервалами. Например, группировка колхозов по размерам валового дохода на крупные и мелкие должна осуществляться на основе специализированных интервалов в зависимости от направлений хозяйств; группировка промышленных предприятий на мелкие, средние, крупные по численности работающих или стоимости основных производственных фондов также предполагает использование различных интервалов в зависимости от принадлежности предприятий к той или иной отрасли промышленности. Так, если в пищевой промышленности предприятие с числом работающих 1000 человек будет крупным, то в тяжелом машиностроении предприятие с таким же числом работающих будет считаться мелким. В этом случае отраслевая принадлежность предприятий является исходным группировочным признаком, а интервалы по численности работающих специализированы по каждой отрасли, что и дает возможность разделить предприятия на крупные, средние, мелкие.

Подводя итог сказанному, можно сделать следующие основные выводы о типологической группировке: 1) типологическая группировка представляет собой исходный, важнейший этап статистической сводки; 2) выделение социально-экономических типов предполагает обоснованный, объективно-научный выбор группировочного признака при условии предварительного определения наметившихся типов в изучаемой совокупности; 3) количество групп и их границы при типологической

группировке определяются на основе глубокого, всестороннего анализа сущности изучаемой совокупности; 4) основным приемом типологической группировки, позволяющим точно выделить реально существующие типы, является комбинационная группировка с использованием при необходимости специализированных группировочных признаков и интервалов.

Социально-экономические типы, выделенные типологической группировкой, могут изучаться с различных позиций, в том числе и с точки зрения их состава, структуры по какому-либо признакам. Группировка, цель которой—расчленение однородной совокупности единиц на группы, характеризующие ее структуру по определенным признакам, называется *структурной*.

В статистической практике и в исследовательской работе структурные группировки находят широкое применение. По результатам регулярно проводимых в СССР переписей населения его структура изучается самым подробным образом: по полу, возрасту, национальности, образованию и другим признакам. Состав предприятий отрасли может изучаться по степени выполнения плана, по уровню рентабельности, степени механизации и автоматизации производственного процесса и др. Применение структурных группировок позволяет не только раскрыть структуру совокупности, но и анализировать изучаемые процессы, их интенсивность, изменение в пространстве, а взятые за ряд периодов времени структурные группировки вскрывают закономерности изменений состава совокупности во времени.

В основу структурных группировок могут быть положены один или более атрибутивных или количественных признаков. Их выбор определяется задачами конкретного исследования и сущностью изучаемой совокупности. Примером структурных группировок по атрибутивным признакам могут служить различного рода классификации, применяемые в статистической практике: распределение работающих в сфере материального производства по отраслям, распределение промышленных предприятий по отраслям промышленности, промышленно-производственного персонала предприятия по категориям работающих, группировка населения страны на городское и сельское и др.

Таблица 6.4. Изменение распределения тачек комбинации «Красная роза» по стажу работы

Группы тачек по стажу работы, лет	Удельный вес численности тачек, % к итогу	
	1976	1980
0—4	21,7	23,2
5—9	15,5	9,1
10—14	13,8	9,1
15—19	20,4	12,1
20—24	22,7	22,7
25—29	4,7	19,2
30—34	1,2	4,6
Итого	100,0	100,0

При структурной группировке по количественному признаку возникает необходимость определения числа групп и их границ. Этот вопрос решается в соответствии с задачами исследования. Один и тот же исходный статистический материал может быть различным образом разбит на группы по одному группировочному признаку, если цели и задачи исследований различны. Например, возрастная структура рабочих предприятия может быть объектом исследования, если изучаются ее динамика, текучесть рабочих кадров, процесс их воспроизводства и другие процессы. В каждом конкретном случае число групп может быть различным. Главное, надо стремиться к тому, чтобы в процессе группировки были ярко отражены особенности изучаемого явления и созданы предпосылки для конкретных выводов и рекомендаций. В табл. 6.4 приводятся структурная группировка, наглядно характеризующая структурные сдвиги в составе ткачей по стажу работы.

Данные табл. 6.4 дают основание для ряда интересных выводов. В составе ткачей комбината преобладают лица со стажем работы 20—24 года и 0—4 года, доля численности этих групп в течение изучаемых лет постоянно превосходит долю численности других возрастных групп. Доля ткачей со стажем работы 5—9 лет снизилась с 15,5% до 9,1%, а доля ткачей с большим стажем работы (свыше 25 лет) увеличилась почти в 4 раза. Это говорит об увеличении степени закрепления кадров ткачей на комбинате.

В приведенном примере интервалы групп равны по своей величине. Если применяются неравные интервалы, то расчет их величины производится по формуле:

$$h = \frac{x_{\max} - x_{\min}}{n}, \quad (6.1)$$

где  $h$  — величина интервала,  $x_{\max}$  — максимальное значение группировочного признака,  $x_{\min}$  — минимальное значение признака,  $n$  — число групп.

Надо отметить, что технически удобнее иметь дело с равными интервалами, но это далеко не всегда представляется возможным из-за свойств изучаемых явлений и признаков. В экономике чаще приходится применять неравные, прогрессивно увеличивающиеся интервалы, что обусловлено самой природой экономических явлений. Примером таких группировок служит распределение промышленных предприятий по численности рабочих, по объему производства; колхозов — по числу колхозных дворов, по размерам посевной площади; электрических станций — по мощности. Так, по численности рабочих промышленные предприятия группируются следующим образом: до 100 человек, 101—200, 201—500, 501—1000, 1001—3000, 3001—10 000, 10 001 человек и выше.

Применение неравных интервалов объясняется главным образом тем, что абсолютное изменение группировочного при-

знака на одну и ту же величину имеет далеко не одинаковое значение для групп с большим и малым значением признака. Между двумя предприятиями с численностью рабочих до 300 человек разница в 100 человек более существенна, чем для предприятий с численностью свыше 5000 человек.

Интервалы групп могут быть *замкнутыми*, когда указав нижнюю и верхнюю границы (см. табл. 6.4), и *открытыми*, когда указана лишь одна из границ групп. Открытые интервалы применяются только для крайних групп (см. табл. 6.5). При группировке с неравными интервалами желательно образовывать группы с замкнутыми интервалами. Это способствует точности статистических вычислений.

Количественные группировочные признаки могут измеряться либо целыми числами (количество членов семьи, число рабочих, колхозных дворов, станков и др.), либо дробными (размер заработной платы, стоимость основных производственных фондов, продукция, возраст). В первом случае рекомендуется верхнюю границу предыдущей группы и нижнюю границу последующей группы обозначать различно с расхождением на одну целую единицу (см. табл. 6.4). Если группировочный признак непрерывный (дробный), то верхняя граница предыдущей группы и нижняя последующей могут различаться на целую единицу, на часть единицы и даже совпадать. При совпадении границ следует давать специальные указания, в какую из групп относить единицу совокупности, если величина ее признака и значение границы группы одинаковы. Так, при распределении жителей района по возрастным группам (лет): 0—20, 20—40, 40—60, 60—80, 80 и старше, возникает вопрос, в какую из групп относить лиц в возрасте 20, 40, 60, 80 лет. Для устранения подобной неопределенности дополнительно указывается «включительно» или «исключительно» считать верхние границы. Если считать верхние границы групп «включительно», то тогда лица в возрасте 20, 40, 60 лет и т. д. должны быть отнесены соответственно к первой, второй, третьей и т. д. группам. Если же считать верхние границы «исключительно», то лиц в указанном возрасте следует отнести во вторую, третью, четвертую и т. д. группы.

Второй важной задачей статистического анализа, проводимого на основе типологической группировки, т. е. в пределах однородных совокупностей, является задача изучения и измерения связи между отдельными признаками. Для решения этой задачи разработаны различные статистические методы, которые будут рассмотрены в гл. 12 и 13. Однако установить факт наличия связи можно и с помощью аналитической группировки.

*Метод аналитических группировок* — традиционный, распространенный прием статистического изучения связей, которые обнаруживаются при параллельном сопоставлении обоб-

ценных значений признаков по группам. Различают признаки зависимые, значения которых изменяются под влиянием других признаков, их обычно в статистике называют *результативными*, и *факторные признаки*, оказывающие влияние на другие. Как правило, основанием группировки служит признак-фактор, а по результативным признакам производится расчет групповых средних, по изменению величины которых определяют наличие связи между признаками. Следовательно, *аналитическими* можно назвать такие *группировки*, которые позволяют установить и изучить связь между результативными и факторными признаками единиц единичной совокупности.

Важная проблема аналитических группировок — правильный выбор числа групп и определение их границ, что в последующем обеспечивает объективность характеристик связи. Поскольку анализ ведется в одноклассовых совокупностях, теоретических оснований для дробления определенного типа нет. Поэтому допустима разбивка совокупности на любое число групп, удовлетворяющее определенным требованиям и условиям конкретного анализа. В процессе аналитических группировок следует соблюдать общие правила группировки, т. е. единицы в образованных группах должны быть статистически однородными по группировочному признаку, группы по величине признака группировки должны быть существенно различны, количество единиц в группах должно быть достаточным для расчета надежных статистических характеристик. Кроме того, групповые средние должны подчиняться определенной закономерности: последовательно увеличиваться или уменьшаться.)

В настоящее время в статистических публикациях предлагаются различные алгоритмы построения аналитической группировки, каждый из которых дает свой вариант группировки как по количеству групп, так и по их границам. Поэтому возникает проблема выбора одного из вариантов образования

групп, который наиболее приемлем для изучения существующих связей между признаками в конкретной исследуемой совокупности. Из возможных вариантов выбирается тот, который удовлетворяет общим требованиям аналитической группировки и частным, вытекающим из задач данного анализа. Для проверки качества выполненной группировки с точки зрения однородности единиц в группах и их различия между группами предлагаются соответствующие критерии, расчет которых позволяет выбрать наиболее оптимальный вариант аналитической группировки. Подробнее о проблемах образования групп и критериях проверки качества группировки будет говориться в гл. 12 и 13.

В качестве примера аналитической группировки рассмотрим группировку строительных организаций области по объему строительно-монтажных работ (табл. 6.5).

Данные табл. 6.5 показывают, что с увеличением размера строительных организаций резко возрастает производительность труда работающих. В крупных строительных организациях, где имеются большие возможности механизировать трудоемкие строительные работы, улучшить организацию труда и т. п., производительность труда более чем в семь раз превышает производительность в мелких строительных организациях.

Непосредственная группировка данных статистического наблюдения — это *первичная группировка*. *Вторичная группировка* — это перегруппировка ранее сгруппированных данных. Необходимость вторичной группировки возникает в двух случаях: во-первых, когда ранее произведенная группировка не удовлетворяет целям исследования в отношении числа групп; во-вторых, для сравнения данных, относящихся к различным периодам времени или к различным территориям, если первичная группировка была произведена по разным группировочным признакам или по разным интервалам.

Метод вторичной группировки был разработан В. И. Лениным и успешно использован в ряде работ. В. И. Ленин использовал два способа вторичной группировки: 1) объединение мелких групп в более крупные и 2) выделение определенной доли единиц совокупности. Первым способом В. И. Ленин пользовался в тех случаях, когда сталкивался с очень дробными группировками, не дающими ясного представления о «главных» группах, скрывающих соответствующие отношения и связи.

Так, например, в работе «Кустарная перепись 1894/95 гг. в Пермской губернии»<sup>1</sup> В. И. Ленин перегруппировал данные о распределении кустарей по доходности заведений. Первоначально имелось 15 групп кустарей. Отметив, что эти данные слишком дробны и их надо свести в более простые и ясные

Таблица 6.5. Зависимость выработки одного работающего от размера строительных организаций

№ п/п	Группы организаций по объему работ, тыс. руб.	Число организаций	Объем строительно-монтажных работ, тыс. руб.	Среднесписочное число работающих	Выработка на одного работающего, руб.
1	До 2500	3	5 959	1 560	3 820
2	2500—5000	5	20 177	2 331	8 656
3	5000—10 000	6	44 602	2 725	16 368
4	Свыше 10 000	6	116 278	4 102	28 347
Итого		20	187 016	10 718	17 449

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 2, с. 379.



рубрики, В. И. Ленин образовал пять групп кустарей по доходности: 1) бедные—с доходом до 50 руб. на семью; 2) малосостоятельные—с доходом 50—100 руб. на семью; 3) средние—с доходом 100—300 руб. на семью; 4) зажиточные—с доходом 300—500 руб. на семью; 5) богатые—с доходом более 500 руб. на семью. Каждая из вновь полученных групп имеет вполне определенное экономическое содержание.

В работе «Развитие капитализма в России» В. И. Ленин использовал второй способ вторичной группировки. Для того чтобы дать общую характеристику социально-экономического строя русской деревни, В. И. Ленин выделил в каждой местности три основные группы крестьянских хозяйств: бедняки, середняки и кулаки. Используя данные земской статистики и

Таблица 6.6<sup>1</sup>

Группы домохозяев	% дворов	% всей посевной площади	% ко всему количеству скота
Без рабочего скота	26,4	2,8	2,9
С 1 гол. раб.	20,3	9,5	8,9
» 2 » » »	14,6	11,8	11,1
» 3 » » »	9,3	10,5	9,8
» 4 » » »	8,3	12,1	11,2
» 5 и более »	21,1	53,3	56,1
Всего	100	100	100

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 3, с. 84 (в сокращении).

военно-конских переписей, В. И. Ленин на основе всестороннего анализа относит к группе бедняцких хозяйств 50 % дворов, к группе середняков—30 % и к зажиточной (кулацкой) группе—20 % дворов. В результате были получены сопоставимые данные по 21 уезду семи губерний.

Рассмотрим пример ленинской группировки данных земской статистики по Камышинскому уезду Саратовской губернии. Результаты первичной группировки данных по уезду приводятся в табл. 6.6.

Чтобы эти данные свести в три социально-экономических типа крестьянских хозяйств, следует руководствоваться ранее установленным процентным соотношением между выделенными типами. Итак, низшая—бедняцкая группа хозяйств должна составлять 50 % всех крестьянских дворов. Значит, в эту группу полностью войдут группы крестьянских дворов без рабочего скота (26,4 %) и хозяйства с 1 головой рабочего скота (20,3 %). В сумме они составят не 50 %, а лишь 46,7 %. Недостающие 3,3 % надо взять из группы дворов с 2 головами рабочего скота.

Посевная площадь в группе бедняцких хозяйств рассчитывается так. Она включает полностью посевную площадь пер-

вых двух групп (2,8+9,5) и часть площади третьей группы. Условно предполагается, что посевная площадь внутри каждой группы распределяется между хозяйствами равномерно, т. е. все хозяйства внутри группы имеют одинаковые посевные площади. В таком случае из третьей группы берется часть посевной площади, соответствующая части дворов, выделенных из нее в низшую группу (3,3 %). Эта часть составляет  $0,226 \left( \frac{3,3}{14,6} = 0,226 \right)$ . Таким образом, часть посевной

площади, которую следует из третьей группы отнест в низшую группу хозяйств, составит:  $0,226 \times 11,8 = 2,7 \%$ . В итоге посевная площадь хозяйств низшей группы будет равна  $2,8 + 9,5 + 2,7 = 15,0 \%$ . Аналогично рассчитывается и процент скота в низшей группе:  $2,9 + 8,9 + 11,1 \times 0,226 = 14,3 \%$ .

Вышая — кулацкая группа хозяйств, в которую должно войти 20 % крестьянских хозяйств, выделяется так. Хозяйства с 5 головами рабочего скота в процентном отношении превышают долю вышей группы уезда на 1,1 %. Следовательно, эта часть хозяйств не войдет в высшую группу, а будет относиться к средней. Она составляет 0,052 (1,1:21,1). Посевная площадь высшей группы соответственно составит:  $53,3 - 53,3 \cdot 0,052 = 50,5 \%$ ; количество скота:  $56,1 - 56,1 \cdot 0,052 = 53,2 \%$ .

Показатели, характеризующие среднюю группу хозяйств, получаются вычитанием из итога (100 %) суммы процентов по показателям низшей и высшей групп, т. е. разностным методом. Окончательные результаты группировки крестьянских хозяйств по социально-экономическим типам приведены в табл. 6.7.

Данные вторичной группировки, приведенные в табл. 6.7, наглядно свидетельствуют о классовой дифференциации крестьянских хозяйств России в пореформенную эпоху. Аналогичная группировка по другим уездам и губерниям дала возможность В. И. Ленину произвести сравнительную характеристику изучаемого процесса и установить известные закономерности в развитии крестьянских хозяйств в России.

Первичная обработка и систематизация материалов статистического наблюдения приводит к образованию упорядоченных рядов цифр.

Таблица 6.7. Распределение крестьянских дворов Камышинского уезда Саратовской губернии по социально-экономическим типам

Группы хозяйства	Процент дворов к итогу	Посевная площадь, % к итогу	Количество скота по группе % к итогу
Низшая	50	15,0	14,3
Средняя	30	34,5	32,5
Вышая	20	50,5	53,2
Итого	100	100,0	100,0

Ряд цифровых показателей, представляющий распределение единиц совокупности по одному признаку, разновидности которого расположены в определенной последовательности, называется *рядом распределения*.

Ряды, построенные по атрибутивному признаку, называются *атрибутивными* рядами распределения. Пример такого ряда приведен в табл. 6.8.

Приведенный ряд распределения содержит три элемента: разновидности атрибутивного признака (социальные группы, по которым распределены члены и кандидаты в члены КПСС); численности единиц в каждой группе, называемые *частотами* ряда распределения; численности групп, выраженные в долях (процентах) от общей численности единиц, называемые *частостями*. Сумма частостей равна 1, если они выражены в долях единицы, и 100%, если они выражены в процентах.

Таблица 6.8. Распределение членов и кандидатов в члены КПСС по социально-му положению (на 1 января 1982 г.)

	Численность членов и кандидатов в члены КПСС, тыс. чел.	% к итогу
Всего	17 769,7	100
в том числе:		
рабочих	7 775,3	43,7
крестьян (колхозников)	2 232,0	12,3
служащих	7 762,4	43,7

в вариационном ряду распределения называются *вариантами* и располагаются в определенной последовательности. Варианты могут выражаться числами положительными и отрицательными, абсолютными и относительными. Примером первого вида может служить распределение рабочих по стажу работы, возрасту, производительности труда и т. д. При группировке предприятий по финансовым результатам варианты могут быть положительными (прибыль) и отрицательными (убытки) числами.

*Вариация признака* может быть дискретной (прерывной) и непрерывной. При *дискретной вариации* значения вариантов отличаются друг от друга на вполне определенную величину (одну или несколько единиц), т. е. варианты дискретного признака обычно выражаются целыми числами. Так, число членов семьи, число человек в бригаде, тарифные разряды рабочих и т. д. могут выражаться только целыми числами.

При построении ряда распределения с дискретной вариацией признака все варианты выписываются в порядке возрастания их величины, подсчитывается, сколько раз повторится одна и та же величина варианта, т. е. частота, и записывается в одной строке с соответствующим значением варианта (см.

табл. 6.1). Частоты в дискретном вариационном ряду, как и в атрибутивном, могут быть заменены частостями.

При *непрерывной вариации* признака его величина может принимать любые значения в определенном интервале. Варианты могут быть целыми и дробными, как угодно мало отличаться друг от друга. Построение непрерывного (интервального) вариационного ряда основано на принципах статистической группировки.

Если построен ряд с равными интервалами, частоты дают представление о степени заполнения интервала единицами со-

Таблица 6.9. Распределение населения района по возрасту (на начало 1984 г.)

Возраст, лет	Число лиц, чел.	Ширина возрастного интервала, лет	Плотность распределения (гр. 1: гр. 2)	Накопленная частота
A	1	2	3	4
Менее 16	13 822	16	864	13 822
16—44	11 456	29	395	25 278
45—59	9 628	15	642	34 906
60 и старше	4 085	40	102	38 991
Итого	38 991	100	390	X

вокупности. При неравных интервалах сравнивать частоты и судить о степени заполнения разных интервалов нельзя. В целях сравнения заполненности интервалов рассчитывается показатель, называемый *плотностью распределения*. Плотность распределения — это число единиц совокупности, приходящееся в среднем на одну единицу ширины интервала. Если плотность распределения определяется отношением частоты на ширину интервала, она будет *абсолютной*, если отношением частоты к ширине интервала — *относительной*.

Возьмем, например, распределение населения района по возрасту (табл. 6.9).

Первый интервал содержит 16 однолетних возрастных групп, начиная от 0 лет и кончая 15 годами. Верхней границей последнего интервала условно будем считать 100 лет (поскольку лица в возрасте старше 100 лет вообще встречаются крайне редко). Тогда его ширина составит 40 лет. Расчет плотности распределения показывает, что среди населения района на начало 1984 г. чаще всего встречаются лица в возрасте до 16 лет и значительно реже лица 60-летнего возраста и старше.

Ряды распределения могут строиться по *накопленным частотам*, которые показывают, какое число единиц имеет вели-

чину варианта, не большую данной. Если вместо абсолютных частот взять частоты, то аналогично получим накопленные частоты. В табл. 6.9 в графе 4 приведены накопленные частоты. Величина накопленной частоты первого интервала равна его частоте, второго интервала — сумме частот первых двух интервалов и т. д. Накопленная частота последнего интервала равна сумме всех частот, т. е. объему совокупности.

Ряды распределения используются в статистике не только как средство систематизации и упорядочения материалов наблюдения, но и для изучения структуры явлений, анализа самих распределений и колеблемости группировочного признака. Подробно об анализе вариационных рядов будет рассказано в гл. 10.

### 6.3. ТЕХНИКА СВОДКИ И ЕЕ КОНТРОЛЬ

С точки зрения техники проведения различают сводку ручную и машинизированную.

*Ручная сводка* осуществляется обычно при небольшом объеме обрабатываемого материала. Она может производиться двумя способами: путем подсчета данных либо с карточек, либо со списков. Ручная сводка с карточек выполняется следующим образом: карточки раскладываются по группам; подсчитывается число карточек, входящих в каждую группу; затем вычисляются итоги по признакам, входящим в разработку; наконец, подсчитанные результаты заносятся в таблицу.

Часто бывает необходимо наряду с количеством карточек рассчитывать и суммы значений отдельных показателей, записанных в карточках. Для облегчения таких подсчетов эти значения записываются на краях карточек. Затем карточки накладываются друг на друга так, чтобы зафиксированные на их краях значения составляли столбик (или строку) чисел. Так считать значительно удобнее и быстрее.

Ручная сводка со списков выполняется двумя вариантами: первый — когда первичный материал занесен в списки в том порядке, в каком он должен сводиться; при втором варианте первичные данные, внесенные в списки, необходимо перегруппировать. В первом случае техника подсчета сводится к простому подведению итогов по отдельным показателям. Второй вариант предусматривает составление по каждой строке списка специальной фишки, т. е. карточки, в которую записывается значение показателя строки. Обработка фишек аналогична обработке карточек.

*Машинизированная сводка* осуществляется с помощью вычислительных машин (клавишных, перфорационных, электронных).

Техника проведения машинизированной сводки с применением клавишных машин незначительно отличается от техники ручной сводки. Дело в том, что клавишные машины выполняют только арифметические действия, лишь

некоторые из них могут печатать результаты в форме таблиц. Поэтому такие наиболее трудоемкие операции, как группировка данных (раскладка карточек, заполнение и раскладка фишек), выполняются вручную.

В условиях применения перфорационных и электронных вычислительных машин операции сводки выполняются механизировано. Однако специфика обработки данных с применением названной техники обуславливает необходимость выполнения целого ряда операций, не свойственных технике ручной сводки и сводки, выполняемой с помощью клавишных машин. К числу таких операций относятся *кодирование признаков и перенос данных с первичных документов на машинные носители*: перфокарты, перфоленты, магнитные ленты. Коды позволяют прежде всего группировать информацию, они обеспечивают в машине подсчет и печать итогов по группировочным признакам. Коды удобны для передачи срочной статистической информации по каналам связи. Например, нет необходимости передавать полные названия отрасли, министерства, предприятия и т. д., когда имеются коды, это значительно сокращает время передачи данных.

При создании Общесоюзной автоматизированной системы обработки информации разрабатываются и внедряются во всех звеньях страны *общесоюзные классификаторы и системы обозначений экономической информации*. В настоящее время в системе ЦСУ ведутся работы по внедрению ряда общесоюзных классификаторов: Общесоюзного классификатора промышленной и сельскохозяйственной продукции (ОКП), Общесоюзного классификатора отраслей народного хозяйства (ОКОНХ), Общесоюзного классификатора строительной продукции (ОКСП) и др.

Все классификаторы статистической информации оформляются в виде специальных *словарей*, где наряду со всеми возможными наименованиями группировочных признаков приводятся их условные обозначения — коды. Кодирование статистического материала заключается в поиске по указанному словарю кода признака по его наименованию и записи этого кода в соответствующие зоны статистического формуляра (как правило, эти зоны выделены в формуляре утолщенными линиями).

Операция переноса данных на машинные носители обусловлена тем обстоятельством, что ввод исходной информации в перфорационную и электронную машину для сводной обработки осуществляется только со специальных технических носителей строго установленного формата. Такими носителями для ПЧМ являются перфокарты, для ЭВМ — перфокарты, перфоленты, магнитные ленты, магнитные карты. В последнее время в качестве носителей исходных данных, приспособленных к непосредственному вводу в ЭВМ, используются специальные *машиночитаемые бланки*. Примером таких носителей могут служить формы переписных листов (ф. № 2С, 3В),



получивших применение в период Всесоюзной переписи населения 1979 г. Эти переписные листы включают две зоны: несчитываемую и считываемую. Считываемая зона предназначена для регистрации сведений, вводимых в ЭВМ. Регистрация этих сведений осуществляется в виде графических отметок, представляемых простым черниграфитным карандашом марки «М».

Машинизированная сводка включает следующие стадии: 1) подготовку данных к сводной машинной обработке; 2) подготовку машинных носителей информации; 3) машинную сводную обработку данных; 4) получение сводных таблиц, размножение их и передачу заказчику.

Подготовку данных к сводной машинной обработке осуществляют экономисты отделов (секторов) подготовки и выпуска статистических материалов. Они проверяют поступившие данные с точки зрения их полноты и правильности оформления. Затем осуществляют кодирование группировочных признаков и комплектовку статистических формуляров для сдачи на ВЦ. Статистические формы комплектуют в отдельные пачки. Скомплектованные документы передают в отдел (сектор) механизированных разработок ВЦ, где проверяется количество документов в каждой пачке, просматривается (на выборку), все ли необходимые данные закодированы. Принятые пачки документов поступают в отдел (сектор) подготовки машинных носителей, где производится перенос данных со статистических формуляров на машинные носители: перфокарты, перфоленты, магнитные ленты. Полученные машинные носители подлежат тщательному контролю. Порядок и содержание работ по контролю достоверности машинных носителей был рассмотрен в гл. 4.

Этап машинизированной сводки — машинный свод данных — включает операции ввода данных в машину, выполнение логических и арифметических действий, получение сводных итогов.

При обработке информации на ПВМ массивы перфокарт с исходными данными поступают на операцию *сортировки*, в процессе которой производится раскладка перфокарт по группировочным признакам, например по территории, форме подчинения, отрасли и т. д. Рассортированные массивы перфокарт поступают на *табуляцию*, в процессе которой осуществляются подсчет и печать сводных итогов. Полученные итоги печатаются на широкой бумажной ленте. Одновременно подсчитанные результаты могут перфорироваться на итоговые карты, которые будут использоваться как носители исходных данных на следующем уровне системы с целью получения соответствующей информации. Например, массив итоговых перфокарт, полученных на ВЦ областного уровня, передается на республиканские ВЦ, где используется для получения сводной отчетности по республике.

Техника сводной обработки статистических данных на ЭВМ значительно отличается от техники свода данных с помощью ПВМ. Принципиальное отличие состоит прежде всего в том, что весь комплекс операций, включающий и группировку, и вычислительный процесс, выполняется ЭВМ автоматически в непрерывном режиме.

Существенным преимуществом ЭВМ является также возможность использования для ввода различных типов носителей: перфокарт, перфолент, магнитных лент, магнитных карт, машиночитаемых бланков. Информация может вводиться в ЭВМ и непосредственно с телеграфных и телефонных каналов связи.

В последнее время получает широкое распространение ввод данных в ЭВМ посредством дисплеев. *Дисплей* — это устройство ввода и вывода информации на электронно-лучевую трубку. Некоторые типы дисплеев снабжены печатающими устройствами типа пишущей машинки.

Большие возможности с точки зрения организации сводной обработки статистических данных обуславливаются и разнообразием форм вывода результатной информации. Вывод результатов может осуществляться на алфавитно-цифровое печатающее устройство в форме *машинограммы*, на перфокарты, перфоленты, магнитные ленты, на графические регистрирующие устройства и устройства вывода в телеграфные и телефонные каналы связи. Графические регистрирующие устройства обеспечивают запись на бумажном бланке любой графической информации, поступающей на ЭВМ (таблиц, графиков, чертежей, схем, контуров рисунков и т. д.).

Организацию и технику сводной обработки статистических данных с применением ЭВМ рассмотрим на примере статистических отчетов.

Статистические отчеты поступают на ВЦ областного управления, где проверяются, кодируются, комплектуются в пачки и сдаются в отдел подготовки машинных носителей. Здесь осуществляется перенос данных с отчетов на машинные носители, в основном перфоленты.

Полученные перфоленты вводятся в ЭВМ. Машина осуществляет логический и арифметический контроль вводимой информации. При выполнении контроля выявляются все смысловые неточности во входящей информации и в зависимости от машинной программы либо осуществляется автоматическое исправление выявленных ошибок, либо выдается на печать протокол расхождений. Он поступает в отдел подготовки и выпуска статистических материалов. Экономист отдела по протоколу находит ошибки, делает в протоколе соответствующие правки, заполняет бланк корректировки и передает его в отдел подготовки машинных носителей для формирования корректирующего массива. Последний вводится в ЭВМ с целью

внесения исправлений в ранее введенный массив исходных данных.

Отконтролированные ЭВМ данные записываются на магнитную ленту. Для обеспечения сохранности информации получают два экземпляра ленты. Полученная магнитная лента с данными первичных статистических ответов будет использоваться для получения всего комплекса соответствующей сводной отчетности. Для разработки свода по области данные с магнитной ленты вводятся в ЭВМ. В соответствии с программой, ранее введенной в машину, осуществляется автоматическое выполнение логических и арифметических операций, в том числе сортировка данных по нужным группировочным признакам, суммирование их по соответствующим группам и подсчет сводных итогов. Результаты подсчетов проверяются ЭВМ автоматически путем увязки одноименных показателей как в пределах одной сводной таблицы, так и между таблицами.

Последующим этапом выполнения сводки на ЭВМ является печать сводных таблиц. Как правило, каждая таблица печатается в нескольких экземплярах в соответствии с количеством адресов рассылки. На заключительной стадии осуществляется контроль сводных таблиц. Проверяется соответствие формы и содержания отпечатанного документа требованиям инструкции, выявляется правильность полученных итогов путем сопоставления их с контрольными суммами. Одновременно просматривается правильность оформления и ясность печатания. На проверенных сводных документах проставляется штамп, указывающий на их достоверность и соответствие требованиям инструкции. Одновременно с выдачей сводных таблиц осуществляется автоматический перенос полученной информации на магнитные ленты, которые будут переданы на ВЦ ЦСУ республики с целью их использования для сводной обработки данных на республиканском уровне.

Техника сводки статистических данных на республиканском ВЦ отличается от рассмотренного варианта тем, что туда поступают данные сразу на машинных носителях (магнитные ленты). Это значительно упрощает технологию сводной обработки и сокращает время сводки.

В условиях АСГС сводка в основном будет осуществляться на ЭВМ. Однако некоторые сводные таблицы простой структуры (с отсутствием строк «в том числе», «из них»), с небольшим числом степеней итогов в подлежащем и при небольшом количестве вычисляемых величин можно разрабатывать на перфорационных машинах. В настоящее время ПВМ находят свое широкое применение на районных и городских информационно-вычислительных станциях (РИБС, ГИВС).

### 7.1. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

После того как данные наблюдения собраны и даже сгруппированы, их трудно воспринимать и анализировать без определенной, наглядной систематизации. Одним из важнейших средств систематизации и обобщения результатов сводки и обработки статистических материалов является *статистическая таблица*. Сведенные в таблицу данные приобретают компактность, наглядность, доходчивость, и статистик получает возможность делать на основании их те или иные выводы. Этим объясняется широкое применение табличной формы изложения статистических данных в практике экономической работы. Статистические формы отчетности — это, как правило, таблицы.

Использование таблиц как средства систематизации данных можно найти в трудах основоположников политической арифметики Дж. Граунта, В. Петти, Г. Кнппа. Эти таблицы носят характер самой простой сводки статистических данных и не дают возможности установить закономерности изучаемого явления. Цель же всякого научного познания действительности состоит именно в раскрытии закономерностей, выражающих те или иные связи между явлениями. Поэтому особую ценность, познавательный интерес имеют статистические таблицы, которые способствуют решению важных задач исследования. Следовательно, важно, какую именно таблицу составляет исследователь: таблицу, являющуюся простой сводкой данных, или же таблицу, которая открывает возможности научного анализа данных.

Отличительная черта статистической таблицы состоит в том, что она дает сводную количественную характеристику исследуемой совокупности. Все одноименные показатели располагаются в таблице в одной и той же строке либо в графе, затем по общим элементам объединяются в разделы, имеющие общий заголовок. Материал приобретает не только удобную для обозрения форму, но и позволяет определить значение каждого признака, их совпадение, последовательность, различие, взаимосвязь. Появляется возможность сравнения, сопоставления и анализа числовых характеристик явления, т. е. возможность сделать научные выводы.

Таким образом, *статистическая таблица представляет собой наиболее рациональную форму изложения результатов сводки и обработки статистических материалов*, позволяющую решить конкретные задачи количественного анализа исследуемого явления.

Статистические таблицы, которые можно рассматривать как вполне научное представление статистического материала, впервые применены в работах русского академика Л. Ю. Крафта (80-е — 90-е годы XVIII столетия).

Понятия составлять сложные аналитические таблицы, предназначенные для раскрытия связей между изучаемыми явлениями, можно найти у А. Кетле. Но А. Кетле не ставил еще вопрос о методологии составления статистиче-

ских таблиц и об их значении как средства раскрытия взаимосвязей между явлениями. Этот вопрос был поставлен немецкими статистиками во второй половине XIX столетия.

Заслугой земских статистиков является широкое применение табличного метода в практической деятельности. Земская статистика сама не разработала теоретических принципов табличного метода, но создала для этого благоприятные условия. Именно практика земских статистиков дала «зачаток» для построения теории табличной разработки и анализа статистических материалов.

Большой вклад в разработку теории табличного метода внесли известные русские статистики А. А. Чупрова и А. А. Кауфман, которые на первое место выдвигали аналитическое значение таблицы как инструмента научного анализа изучаемых явлений.

В конце XIX в. в свет начали выходить труды В. И. Ленина, в которых табличный метод нашел применение на самом высоком научном уровне. В. И. Ленин в своих исследованиях не только использовал готовые таблицы из земско-статистических сборников, но и сам составлял совершенно новые, аналитические таблицы, позволившие объяснить причины экономического разложения крестьянства России.

В экономической работе таблицы играют важнейшую роль. Их ценность становится вполне ощутимой в том случае, если всесторонне исследуется сложное общественное явление.

Грамотный экономист должен уметь правильно составлять таблицы, правильно читать разработанные таблицы и делать на их основе верные выводы.

Статистические таблицы внешне представляют определенного рода пересечения вертикальных граф и горизонтальных строк, которые образуют клетки, предназначенные для вписывания в них статистических данных.

Каждая графа (колонка, столбец) в верхней ее части обязательно имеет пояснение в виде краткого заголовка. Наименования строк помещаются в левой части таблицы. При нанесении только строк и граф без их наименований и статистических данных получается графенная сетка, которая именуется скелетом таблицы. Если скелет таблицы заполнить наименованиями строк и граф, то получится макет таблицы.

Каждая таблица имеет общее подробное название, которое раскрывает читателю ее содержание и назначение, указывая, к какому месту и времени они относятся.

Статистическую таблицу можно рассматривать как форму *логического предложения*, имеющего *статистическое подлежащее* и *сказуемое*. Идея уподобления статистической таблицы грамматическому предложению принадлежит А. А. Кауфману. *Подлежащее* таблицы представляет ту статистическую совокупность, о которой идет речь в таблице, т. е. перечень отдельных или всех единиц совокупности либо их групп. Чаще всего подлежащее помещается в левой части таблицы и содержит перечень строк. *Сказуемое* таблицы — это цифровая характеристика изучаемой совокупности (подлежащего таб-

лицы), представленная перечнем различных признаков в соответствующих графах таблицы.

Рассмотрим в качестве примера табл. 6.5. Подлежащим и ей является совокупность строительных организаций, разбитая на группы по объему выполненных работ. Сказуемое таблицы — это характеристика каждой группы организаций их числом, объемом строительно-монтажных работ, среднесписочной численностью работающих, выработкой на одного работающего.

Подлежащее и сказуемое таблицы могут располагаться по-разному. Это технический вопрос, главное, чтобы таблица была легко обозримой, компактной и легко воспринималась.

В статистической практике и в исследовательских работах используются таблицы различной сложности. Это зависит от характера изучаемой совокупности, объема имеющейся информации, задач анализа. Существуют разные принципы классификации таблиц. Наиболее распространенной в советской экономической литературе является классификация, определяемая разработкой статистического подлежащего или группировкой единиц в подлежащем. Если в подлежащем содержится простой перечень каких-либо объектов или территориальных единиц, *таблица называется простой*. Простые таблицы имеют самое широкое применение в статистической практике. Характеристика предприятий отрасли по технико-экономическим показателям за отчетный квартал, полугодие, год приводится в простой перечневой таблице, в подлежащем которой перечисляются все предприятия отрасли. Характеристика городов Советского Союза по численности населения, размеру жилищного фонда и т. п. также представляется в простой таблице. Подлежащее простой таблицы может содержать перечень территорий, например областей, республик, экономических районов. В этом случае таблица называется *территориальной*. Примером макета простой территориальной таблицы является табл. 7.2.

Простая таблица содержит лишь описательные сведения, ее аналитические возможности ограничены. Глубокий анализ исследуемой совокупности, взаимосвязей признаков предполагает построение более сложных таблиц — групповых и комбинационных.

*Групповые таблицы* в отличие от простых содержат в подлежащем не простой перечень единиц объекта наблюдения, а их группировку по одному существенному признаку. Простейшим видом групповой таблицы являются таблицы, в которых представлены ряды распределения (см. табл. 6.8). Групповая таблица может быть более сложной, если в сказуемом приводятся не только число единиц в каждой группе, но и ряд других важных показателей, количественно и качественно характеризующих группы подлежащего (см. табл. 6.5). Такие таблицы часто используются в целях сопоставления обобща-



ющих показателей по группам, что позволяет сделать определенные практические выводы. Более широкими аналитическими возможностями располагают комбинационные таблицы.

Комбинационными называются статистические таблицы, в подложке которых группы единиц, образованные по одному признаку, подразделяются на подгруппы по одному или нескольким другим признакам. В отличие от простых и групповых таблиц комбинационные позволяют проследить зависимость показателей сказуемого от нескольких признаков, которые легли в основу комбинационной группировки в подложке.

Предположим, задача заключается в том, чтобы по данным сорока предприятий отрасли показать взаимосвязь между объемом производства, электровооруженностью труда, с одной стороны, и средней выработкой продукции на одного рабочего, с другой стороны. Предварительное изучение характера взаимосвязей показывает, что на выработку рабочего влияет объем производства предприятий и электровооруженность труда. Следовательно, результативным будет первый признак, а два последних — факторные. Подлежащее таблицы будет представлять комбинацию двух группировочных признаков, т. е. первоначально выделяются группы предприятий по объему произведенной продукции, а затем последние подразделяются на подгруппы по электровооруженности труда. Затем по группам и подгруппам производится подсчет данных по показателям сказуемого и все результаты оформляются в виде комбинационной таблицы (см. табл. 7.1).

В табл. 7.1 отражена не только структура предприятий с точки зрения выпуска продукции и электровооруженности труда, но и взаимосвязи между признаками. Так, весьма четко прослеживается зависимость выработки рабочего как от объема произведенной продукции, так и от электровооруженности труда. Если на предприятиях с выпуском продукции до 10 млн. руб. в год производительность труда на одного рабочего за год составляет 3,8 тыс. руб., то на более крупных предприятиях с выпуском продукции свыше 20 млн. руб. выработка достигает 7,1 тыс. руб. Аналогично во всех трех группах предприятий с ростом электровооруженности труда увеличивается выработка: в первой группе она возрастает с 3,6 до 4,1 тыс. руб.; во второй — с 5,0 до 5,7 тыс. руб.; в третьей — с 6,8 до 7,4 тыс. руб.

При исследовании зависимости результативных признаков от трех и более факторов комбинационные таблицы усложняются, становятся громоздкими и трудно обозримыми. В этих случаях рекомендуется строить не одну таблицу, а систему взаимосвязанных групповых и комбинационных таблиц.

В. И. Ленин, рассматривая статистические таблицы как важнейшее оружие экономического анализа, особенно подчеркивал значение комбинационных таблиц как метода, позво-

ляющего с наибольшей полнотой освещать изучаемое явление. «Научное значение таких разнообразных и тем учитывающих гигантское разнообразие форм подчинения земледелия и землевладельца рынку комбинационных таблиц,— писал В. И. Ленин,— было бы громадно. Можно сказать без всякого преуве-

Таблица 7.1. Зависимость выработки рабочих группы машиностроительных предприятий от объема производства и электровооруженности труда в 1984 г. (данные условные)

Группы предприятий		Число предприятий	Среднесписочное число рабочих, чел.	Объем произведенной продукции, тыс. руб.	Среднегодовая выработка продукции на 1 рабочего, тыс. руб. (гр. 3 : гр. 2)
по объему произведенной продукции, млн. руб.	по электровооруженности труда, кВт·ч на 1 чел.-ч				
А	Б	1	2	3	4
До 10	5—10	6	11 200	40 440	3,6
	10—20	6	9 100	37 310	4,1
	Итого	12	20 300	77 750	3,8
10—20	5—10	10	27 200	136 000	5,0
	10—20	8	21 000	119 700	5,7
	Итого	18	48 200	255 700	5,3
Свыше 20	5—10	4	15 200	103 360	6,8
	10—20	6	21 100	156 140	7,4
	Итого	10	36 300	259 500	7,1
Всего	5—10	20	53 600	279 800	5,2
	10—20	20	51 200	313 150	6,1
	Итого	40	104 800	592 950	5,7

ления, что они внесли бы целый переворот в науку об экономике земледелия»<sup>1</sup>.

Используя табличный метод и, в частности, групповые и комбинационные таблицы, В. И. Ленин в своих трудах показал характерные особенности развития капитализма в России, причины экономического разложения крестьянства, распределение населения России на классы, открыл закономерности, характеризующие политические отношения между классами. С помощью статистических таблиц В. И. Ленин предельно использовал статистический материал, получил из него максимум информации. Анализируя материалы подворных переписей, проводимых земскими статистиками, В. И. Ленин писал: «Все дело только за обработкой этих превосходных данных...

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 24, с. 281.

обработка подворных данных должна дать как можно больше, как можно рациональнее и детальнее составленных *групповых и комбинационных таблиц* для отдельного изучения *всех* наметившихся или *намечающихся* (это не менее важно) — в жизни *типов* хозяйств. Без разносторонних и рационально составленных групповых и комбинационных таблиц богатейшие подворные данные прямо-таки пропадут»<sup>1</sup>.

*Построение статистических таблиц* начинается с разработки макета будущей таблицы. Это основной и ответственный этап, требующий определенных знаний в области изучаемой проблемы, четкого представления целей и задач исследования в целом и назначения конкретно разрабатываемой таблицы.

*Макет таблицы* включает следующие элементы: общий заголовок таблицы, ее скелет, полное наименование подлежащего и всех его составляющих (перечень единиц, групп, подгрупп), наименование всех граф сказуемого, итоговые строки и графы. После построения макета таблицы следует техническая операция — разнесение статистических данных в соответствующие клетки и подсчет итогов. Не следует увлекаться слишком громоздкими макетами таблиц, их делать сложно заполнять данными, они потеряют обзорность, возникнут сложности при анализе.

Как правило, макет таблицы строится после того, как четко определены границы исследуемой совокупности, произведена ее группировка по одному или нескольким признакам, т. е. известна структура подлежащего таблицы, необходимо решить вопрос о его удобном размещении. Сложнее решаются вопросы разработки сказуемого таблицы.

При детальном изучении сложного явления может быть большое число признаков сказуемого, которые могут быть представлены в свернутом либо развернутом виде. Развернутое сказуемое содержит перечень показателей, которые при необходимости подразделяются на составляющие. Свернутые сказуемые содержат, например, табл. 6.4, 6.5, 6.7, 6.8. Развернутое сказуемое может иметь простую и сложную разработку. Если составные части показателя в сказуемом не подразделяются на части, разработка сказуемого считается простой (см. табл. 7.2).

Если в табл. 7.2 студентов стационарной, заочной и вечерней форм обучения подразделить в свою очередь по полу, разработка сказуемого будет считаться сложной. Сложная разработка сказуемого дает более полную характеристику исследуемого объекта, но приводит к увеличению размеров таблицы и осложняет ее восприятие. Это обстоятельство надо учитывать при разработке макета таблицы и выделять следует лишь те составляющие показатели, которые действи-

тельно необходимы в процессе исследования, а другие объединить в одну графу под рубрикой «прочие». Однако группа «прочие» не должна охватывать более десятой части общего итога. Если таблица получается крайне громоздкая и труднодоступная для обозрения, ее рекомендуется разделить на несколько таблиц. Комбинированная разработка сказуемого позволяет одну сложную таблицу представить в виде системы таблиц с простой разработкой сказуемого.

Таблица 7.2. Высшие учебные заведения УССР на начало 1981/85 учебного года

Число вузов	В них студентов, тыс. чел.				Число преподавателей, чел.
	стационарной формы обучения	заочной формы обучения	вечерней формы обучения	иного	
УССР					
в том числе по городам:					
Днепропетровск					
Донецк					
Киев					
Одесса					
Харьков					
Другие города					

Показатели сказуемого могут давать количественную оценку изучаемого объекта на какой-либо один момент времени (см. табл. 7.2) или за определенный промежуток времени (см. табл. 7.1). Такие таблицы называются *статическими*. Если же показатели сказуемого характеризуют изменение объекта во времени, то таблицы называются *динамическими* (см. табл. 6.4).

Оформление таблиц не должно быть произвольным. Существуют определенные правила, которыми необходимо руководствоваться при оформлении таблиц.

Каждая таблица должна иметь название (заголовок), которое в лаконичной форме должно раскрывать ее содержание. В названии четко и ясно следует указать границы статистической совокупности, период или момент времени, к которому относятся данные. Если единица измерения для всех данных таблицы одинакова, ее целесообразно вынести в общий заголовок. При сопоставлении данных сказуемого по всем элементам подлежащего с одним и тем же годом в заголовке таблицы указать год сопоставления, например по сравнению с 1913 г. или по сравнению с 1980 г. В общем заголовке должны быть сделаны оговорки о полноте приводимых данных и их характере (плановые, фактические, расчетные).

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 24, с. 277.

Кроме общего заголовка таблицы, ее графы и строки тоже должны иметь свое название. Слова в заголовках подлежащего и сказуемого желательно писать полностью и ставить единицы измерения, если они неодинаковые. Если каждая строка имеет свою особую единицу измерения, то для их обозначения целесообразно отводить специальную графу.

Графам таблицы, если их много, желательно давать нумерацию. Это облегчает пользование таблицей и, кроме того, дает возможность показать способ расчета некоторых показателей (см. табл. 7.1). Первая графа, если она предназначена для подлежащего, обозначается буквой «А», графы сказуемого нумеруются арабскими цифрами.

Обязательным атрибутом статистической таблицы являются итоговые строки и графы. Без итогов ряд таблиц нельзя признать законченными. В сложных таблицах следует различать «итого» и «всего». «Итого» — это характеристика, относящаяся к определенной части совокупности, а «всего» — это итог в целом для всей изучаемой совокупности (см. табл. 7.1).

Часто в таблицах приводятся данные, связанные друг с другом. В таких случаях эти данные следует располагать в рядом стоящих графах. Например, абсолютные данные и соответствующие им относительные показатели, данные о плановом задании и о выполнении плана, абсолютные приросты и темпы роста.

Большое значение в статистических таблицах имеет запись цифр в графах с соблюдением определенных правил. Статистика часто оперирует многозначными числами, которые для простоты восприятия приходится округлять. Основное правило округления, как известно, заключается в том, что, если заменяемая цифра больше или равна 5, стоящую от нее слева цифру увеличивают на единицу. При округлении надо помнить о назначении таблицы. Чрезмерное округление может скрыть существующие связи и закономерности, исказить действительную картину. Крайне нежелательно, чтобы после округления оставалась лишь одна цифра, а иногда и две. Если данные приводятся в процентах, округление следует производить до долей процентов. Многозначные числа, состоящие из четырех и более цифр, необходимо записывать, отделяя каждые три цифры друг от друга, т. е. выделяя классы миллионов, тысяч, единиц. В такой записи многозначные числа легче читаются и сопоставляются. Например, число 73872458 надо записать 73 872 458, а при округлении — 73,9 млн.

В статистической таблице каждая клетка должна быть заполнена. Однако в ряде случаев числа в клетках могут отсутствовать. Причины отсутствия должны быть определенным образом показаны в таблице. Если сведения о данном факте отсутствуют, но сам факт имеет место, рекомендуется ставить три точки (...). В случае если отсутствует само явление, ставится тире (—). Если клетка не подлежит заполнению, ста-

вится знак Х. Число 0,0 ставится в тех случаях, когда изучаемое явление наблюдается, но в очень малых размерах и составляет менее половины последней значащей цифры в условиях принятой точности.

Числа в табличных клетках могут сопровождаться определенными знаками. Так, если число получено на основании условных расчетов, т. е. является не совсем точным, его рекомендуется брать в скобки. Сомнительные числа должны сопровождаться вопросительным знаком, а предварительные — знаком \*.

В ряде случаев статистические таблицы сопровождают сносками и примечаниями. Сноски относятся либо к отдельным строчкам, либо к графам, иногда и к отдельным клеткам. Сносками пользуются для того, чтобы указать на ограниченные обстоятельства, которые надо принять во внимание при чтении таблицы. Примечания, которые относятся ко всей таблице, обычно даются сразу же после таблицы. Они дают пояснения методам расчета показателей, указывают источники получения данных, определенные ограничения, принятые при составлении таблицы.

Грамотное, корректное оформление таблиц требует обязательного соблюдения всех отмеченных правил, они вносят единообразие в разработку статистических таблиц и облегчают их чтение.

## 7.2. СТАТИСТИЧЕСКИЕ ГРАФИКИ

Полученные в результате сводки либо статистического анализа в целом числовые показатели могут быть представлены не только в табличной, но и в графической форме. При этом используются различные графики, многообразие видов которых обусловлено различиями в их статистическом содержании, способах построения и широтой круга изображаемых ими общественных явлений и процессов.

Графиками в статистике называются условные изображения числовых величин и их соотношений в виде различных геометрических образов — точек, линий, плоских фигур и т. п. Использование графиков для изложения статистических показателей позволяет придать последним наглядность и выразительность, облегчить их восприятие, а во многих случаях и анализ. Статистический график дает возможность сразу оценить характер изучаемого явления, присущие ему закономерности и особенности, тенденции развития, взаимосвязь характеризующих его показателей.

В. И. Ленин придавал большое значение графическому выражению статистических данных и сам часто прибегал к нему в своей работе. Так, например, анализируя работу II съезда РСДРП в книге «Шаг вперед, два шага назад», он отмечал, что разделение делегатов по различным обсуждавшимся на



съезде вопросам дает единственную в своем роде картину внутренней борьбы в партии. «Чтобы сделать эту картину наглядной,— писал В. И. Ленин,— чтобы получить настоящую картину, а не груду бессвязных ... фактов и фактиков ... я решил попытаться изобразить все основные типы «разделений» нашего съезда в виде *диаграммы*. Такой прием покажется, наверное, странным очень и очень многим, но я сомневаюсь, можно ли найти другой способ изложения, действительно обобщающего и подводящего итоги, возможно более полного и наиболее точного»<sup>1</sup>. В работе «Очередные задачи Советской власти»<sup>2</sup> В. И. Ленин призвал путем широкой популяризации статистических данных нести статистику в массы.

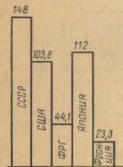


Рис. 7.1. Выпуск стали в отдельных странах в 1980 г., млн. т

В наши дни статистические графики прочно вошли в практику статистической работы. Они применяются для иллюстрации статистических данных, содержащихся в Программе КПСС, в материалах съездов партии и Пленумов ЦК КПСС, используются в агитационной работе для популяризации успехов Советского государства в деле социально-экономического и культурного строительства.

Каждый график состоит из графического образа и вспомогательных элементов. *Графический образ*—это совокупность точек, линий и фигур, с помощью которых изображаются статистические данные. *Вспомогательные элементы* графика включают общее название графика, пояснения условных знаков и смысла графического образа, оси координат, шкалы, числовые сетки и числовые данные, дополняющие или уточняющие изображаемые показатели. Вспомогательные элементы облегчают чтение графика и его истолкование. Так, на рис. 7.1 графический образ представляет собой ряд столбиков, общее название дает представление о явлении, надписи над столбиками и внутри них показывают, в какой стране и сколько стали выплавлялось в 1980 г.

Название графика должно кратко и точно раскрывать его содержание. Пояснительные тексты могут располагаться в пределах графического образа или рядом с ним, как на рис. 7.1 (ярлыки), либо выноситься за его пределы, как на рис. 8.2 и 8.5 (ключ).

Оси координат с нанесенными на них шкалами и числовые сетки необходимы для построения графика и пользования

ним. *Шкалы* могут быть прямыми или криволинейными (круговыми), равномерными (линейными) и неравномерными. Примером неравномерной шкалы является логарифмическая шкала (рис. 7.2). Логарифмические шкалы в статистических графиках используются реже, чем линейные (в основном в тех случаях, когда изображаемый показатель изменяется в очень широких пределах).

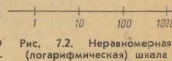
Иногда оказывается целесообразным применять так называемые *сопряженные шкалы*, построенные на одной или двух параллельных линиях. Чаще всего одна из сопряженных шкал используется для отсчета абсолютных величин, а вторая—соответствующих им относительных (см. рис. 8.1). Числа на шкалах проставляются равномерно, при этом последнее число должно несколько превышать максимальный уровень показателя, значение которого отсчитывается по этой шкале. Числовая сетка, как правило, должна иметь базовую линию, роль которой обычно играет ось абсцисс.

Статистические графики можно классифицировать по разным признакам: назначению (содержанию), способу построения и характеру графического образа.

По содержанию или назначению можно выделить: графики сравнения в пространстве, графики различных относительных величин (выполнение плана, структуры, т. е. состава, динамики и др.), графики вариационных рядов, графики размещения по территории, графики взаимосвязанных показателей. Возможны и комбинации этих графиков, например графическое изображение вариации в динамике или динамики взаимосвязанных показателей и т. п. Различные по назначению или по содержанию изображаемых показателей графики рассматриваются в соответствующих главах учебника.

По способу построения графики можно разделить на диаграммы, картодиаграммы и картограммы. По характеру графического образа различают графики точечные, линейные, плоскостные (столбиковые, полосовые, квадратные, круговые, секторные, фигурные) и объемные.

Рассмотрим правила построения *столбиковой диаграммы*, которая используется чаще всего для сравнения одноименных показателей, характеризующих различные объекты или территории. Значения сравниваемых показателей изображаются при этом в виде прямоугольных вертикальных столбиков, имеющих одинаковую ширину и расположенных на общей горизонтальной базовой линии (см. рис. 7.1). Высота каждого столбика и определенном масштабе соответствует величине изображаемого показателя. Столбики, каждый из которых относится к определенному объекту, могут располагаться либо на одинаковом расстоянии друг от друга, либо вплотную. Разновидностью столбиковой диаграммы является *полосовая* (ден-



<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 8, с. 321—322.

<sup>2</sup> Там же, т. 36.

точная) диаграмма (рис. 7.3), для которой характерны горизонтальная ориентация столбиков (полос) и вертикальное расположение базовой линии. Полосовая диаграмма особенно удобна в тех случаях, когда отдельные объекты сравнения характеризуются противоположными по знаку показателями.

Иногда сравниваемые объекты характеризуются резко различившимися значениями показателей. Пусть, например, численность населения областного центра составляет 900 тыс. человек, а районного центра — 25 тыс. человек. Изобразить эти данные с помощью столбиковой диаграммы практически не-

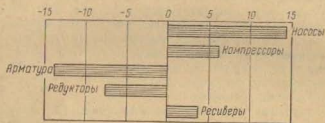


Рис. 7.3. Изменение себестоимости отдельных видов продукции промышленного предприятия в 1982 г. по сравнению с 1975 г., %

возможно, так как высота одного столбика должна в 36 раз превышать высоту другого ( $900:25=36$ ). В подобных случаях используются такие виды плоскостных диаграмм, как *квадратные* или *круговые*. Их построение основано на том, что величины изображаемых показателей должны быть пропорциональны площадям квадратов или кругов, а корни квадратные из сравниваемых величин — пропорциональны линейным размерам этих фигур (сторонам квадратов или радиусам кругов).

В данном примере (рис. 7.4) стороны квадратов, расположенных на горизонтальной базовой линии, пропорциональны  $6:1$  ( $\sqrt{900}:\sqrt{25}=30:5=6:1$ ).

*Квадратные и круговые диаграммы* менее наглядны, чем столбиковые и полосовые, что связано с трудностью визуальной оценки соотношения площадей. Поэтому внутри квадратов и кругов обязательно следует проставлять величины изображаемых показателей. Еще меньшей наглядностью отличаются *объемные диаграммы* (например, в виде кубов), в которых линейные размеры графического образа пропорциональны корням кубическим из сравниваемых величин. Употребляются они поэтому весьма редко.

*Фигурные диаграммы* сравнения предназначены в основном для агитационных целей, для популяризации. В них показатели изображаются в виде определенного количества стандартных фигур, представляющих собой упрощенные изо-

бражения объектов, характерных для соответствующих явлений. Так, силуэт человека может быть принят за определенную численность, например за 100 тыс. чел., упрощенное изображение буровой вышки может обозначать 100 тыс. т добычаемой нефти и т. п. Такие диаграммы очень доходчивы и обладают высокой степенью эмоционального воздействия на читателя. Недостатком их следует считать некоторую неточность, связанную с необходимостью округления изображаемых

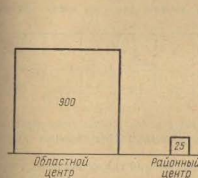


Рис. 7.4. Численность населения областного и районного центров, тыс. чел.



Рис. 7.5. Распределение квартир по числу живущих в них лиц

показателей либо с необходимостью изображения не только целых фигур, но и их частей.

Для изображения *вариационных рядов* применяются *линейные* и *плоскостные диаграммы*, построенные в прямоугольной системе координат. При дискретной вариации признака графиком вариационного ряда служит *полигон распределения*. Он представляет собой замкнутый многоугольник, абсциссами вершин которого являются значения варьирующего признака, а ординатами — соответствующие им частоты или частости. Для примера рассмотрим построение полигона распределения по следующим данным (табл. 7.3).

Отложив в выбранном масштабе число живущих в квартире по оси абсцисс, а соответствующие им количества квартир по оси ординат, получим совокупность точек. Соединив эти точки отрезками прямых и опустив перпендикуляры из крайних точек на ось абсцисс, получим полигон распределения (рис. 7.5).

При непрерывной вариации признака часто используют, как известно, интервальные вариационные ряды, графическим изображением которых служит так называемая *гистограмма*. Для построения гистограммы по оси абсцисс в соответствии с принятым масштабом откладывают границы интер-

валов. Эти интервалы служат основаниями прямоугольников, высота которых пропорциональна плотности распределения соответствующих интервалов.

**Плотность распределения** — это число единиц совокупности (частота, частость), приходящееся на единицу ширины интервала. При равных интервалах плотности распределения пропорциональны частотам или частостям, которые и используются для построения прямоугольников. При неравных интер-

Таблица 7.3. Распределение квартир жилого дома по числу проживающих в них лиц

Число лиц, живущих в квартире	1	2	3	4	5	6	7	Всего
Число квартир	4	6	22	46	18	3	1	100

валах гистограмма строится только по плотности распределения. На рис. 7.6 изображена гистограмма распределения электродвигателей по мощности.

Для иллюстрации рядов распределения также используется **кумулятивная кривая** (кумулята). Для ее построения на оси

Таблица 7.4. Распределение электростанций по мощности

Группы электростанций по мощности, МВт	В процентах к итогу			Нарастающие проценты к итогу		
	число электростанций	отпуск электроэнергии	потребление топлива	число электростанций	отпуск электроэнергии	потребление топлива
До 50	20,5	1,8	0,7	20,5	1,8	0,7
50—100	12,9	2,9	1,5	33,4	4,7	2,2
100—200	36,5	20,1	12,8	69,9	24,8	15,0
200—400	11,3	15,0	12,0	81,2	39,8	27,0
400—1000	12,0	33,4	29,4	93,2	73,2	56,4
1000—3000	3,2	11,9	14,7	96,4	85,1	71,1
Более 3000	3,6	14,9	28,9	100	100	100
Всего	100	100	100	—	—	—

абсцисс откладываются значения дискретного признака (или границы интервала), а на оси ординат — нарастающие итоги частот или частостей, соответствующие этим значениям признака (или верхним границам интервалов). Кумулята распределения квартир по числу живущих в них людей приведена на рис. 7.7.)

Если при группировке подсчитать не только число единиц, но и общий объем какого-либо признака, то полученные данные будут характеризовать его распределение по группам и

могут быть использованы для **графического изображения концентрации**. Пусть, например, имеются следующие данные по электростанциям одной из энергосистем (табл. 7.4).

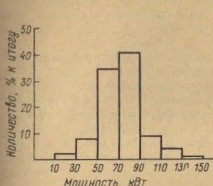


Рис. 7.6. Распределение электродвигателей по мощности

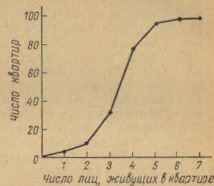


Рис. 7.7. Кумулята распределения квартир по числу живущих в них лиц

Из таблицы видно, что самые мощные электростанции, составляющие 3,6 % общего количества, выработали 14,9 % всей отпущенной за год электроэнергии и израсходовали 28,9 % всего потребленного топлива. В то же время относительно маломощные станции, количество которых составляет 20,5 % общего, выработали 1,8 % всей электроэнергии, затратив 0,7 % всего израсходованного топлива. Эти данные свидетельствуют о концентрации выработки электроэнергии и расхода топлива на крупных электростанциях.

Для графического изображения явления концентрации используются нарастающие итоги показателей. Для этого в групповой таблице (см. табл. 7.4) находятся суммы накопленных значений частостей, а также значения сумм накопленных итогов интересующих нас показателей (в процентах к итогу). В нашем при-

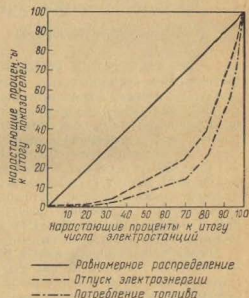


Рис. 7.8. Концентрация отпуска электроэнергии и потребления топлива электростанциями



мере это значения нарастающих процентов к итогу отпуска электроэнергии и потребления топлива. Отложим нарастающие итоги частостей по оси абсцисс, а соответствующие нарастающие итоги двух других показателей — по оси ординат. Соединив отрезками прямых найденные таким образом точки, получим ломаные линии, которые и будут являться *кривыми концентрации* (рис. 7.8).

Если бы распределение выработанной электроэнергии между электростанциями было равномерным (т. е. пропорциональным их числу), то концентрация отсутствовала бы, и график представлял бы собой прямую, выходящую из начала координат и расположенную под углом  $45^\circ$  к оси абсцисс. То же самое, конечно, наблюдалось бы и при отсутствии концентрации потребления топлива. Из рис. 7.8 видно, что концентрация выработки электроэнергии менее заметна, чем концентрация потребления топлива.

## ГЛАВА 8 АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

### 8.1. АБСОЛЮТНЫЕ ВЕЛИЧИНЫ

В итоге сводки статистических данных образуются обобщающие статистические показатели, характеризующие результаты познания количественной стороны массовых общественных явлений. Важной разновидностью таких показателей являются абсолютные величины, отражающие в определенных условиях места и времени размер данного явления в соответствующей ему конкретной форме. Например, по данным ЦСУ СССР в 1983 г. в стране перевезено железнодорожным транспортом 3851 млн. т грузов; розничный товарооборот государственной и кооперативной торговли составил 305,8 млрд. руб., а численность населения Советского Союза на 1 января 1984 г. — 273,8 млн. человек. Все приведенные статистические показатели являются абсолютными величинами.

Абсолютные величины имеют большое практическое и познавательное значение. В них выражаются основные показатели экономической мощи страны (валовой и национальный доходы, объемы промышленного и сельскохозяйственного производств, выпуск важнейших видов продукции и т. д.). В абсолютных величинах устанавливаются плановые задания на всех уровнях управления: народное хозяйство, отрасль, объединение (предприятие).

Различают два вида абсолютных величин: индивидуальные и суммарные. *Индивидуальные* называют абсолютные величины, характеризующие размеры отдельных единиц совокупности (например, количество деталей, изготовленных одним

рабочим за смену, число детей в отдельной семье). Они получают непосредственно в процессе статистического наблюдения и фиксируются в первичных учетных документах.

*Суммарные абсолютные величины* получаются, как правило, путем суммирования отдельных индивидуальных величин. Так, в процессе переписи населения органы государственной статистики получают итоговые абсолютные данные о численности населения СССР, о распределении его по союзным республикам, по полу, возрасту, национальности.

Вместе с тем встречаются ситуации, когда итоговые абсолютные показатели определяют в результате специальных расчетов. Например, при составлении долгосрочных прогнозов развития экономики нашей страны учеными были получены данные о предположительной численности трудовых ресурсов, необходимом количестве различных видов топлива и сырья в 2000 г.

*Абсолютные величины отражают естественную основу явлений. Абсолютные величины выражают либо численность единиц изучаемой совокупности, ее отдельных составных частей, либо их абсолютные размеры в натуральных единицах, вытекающих из их физических свойств (вес, длина и т. п.), или в единицах измерения, вытекающих из их экономических свойств (стоимость, затраты труда). Следовательно, абсолютные величины всегда имеют определенную размерность.*

— *Именованные числа, в которых выражаются абсолютные величины (килограммы, метры, рубли, человеко-дни и т. д.), представляют собой измерители размеров различных явлений и процессов в соответствующей им конкретной форме: выплавку стали — в тоннах, лекционное время — в часах, расстояние перевозки груза — в километрах, отработанное время — в человеко-часах, заработную плату — в рублях.*

При всем разнообразии единиц измерения их обычно подразделяют на три вида: натуральные, трудовые, стоимостные. *Натуральные измерители* характеризуют явления в собственной им натуральной форме и выражаются в мерах длины, веса, объема и т. п. или количеством единиц, числом событий. Так, например, выпуск тракторов характеризуется в штуках либо суммарной мощностью в лошадиных силах; сбор зерна измеряется в центнерах, тоннах.

В ряде случаев используются комбинированные единицы измерения, представляющие собой произведение двух величин, выраженных в различных размерностях. Так, например, производство электроэнергии измеряется в киловатт-часах, грузооборот — в тонно-километрах, затраты труда — в человеко-часах, человеко-днях и т. п.

Для получения суммарных абсолютных величин в случае, когда индивидуальные величины характеризуют отдельные разновидности продукции, близкие по своим потребительным свойствам, применяются *условные натуральные измерители*:

грузовые вагоны пересчитываются в двухосные, мыло — в 40 %-ное, топливо — в 7000-калорийное и т. д. При этом одна из разновидностей продукции принимается за условный натуральный измеритель и к ней с помощью переводных коэффициентов, выражающих соотношение потребительских свойств (иногда трудоемкости, себестоимости и т. д.) отдельных разновидностей, приводятся все разновидности этого продукта.

Допустим, что мыловаренный завод выпустил 7 тыс. т мыла с 40 %-ным содержанием жирных кислот, 5 тыс. т — с 60 %-ным содержанием жирных кислот и 3 тыс. т — с 80 %-ным содержанием жирных кислот. Определим сначала, что переводные коэффициенты будут равны: для 40 %-ного мыла 1,0 (40:40), для 60 %-ного мыла — 1,5 (60:40), а для 80 %-ного мыла — 2,0 (80:40). Тогда общий объем продукции завода в пересчете на условное 40 %-ное мыло составит:  $7 \cdot 1,0 + 5 \cdot 1,5 + 3 \cdot 2,0 = 20,5$  тыс. т.

Трудовые единицы измерения используются для характеристики показателей, отражающих наличие, распределение и использование трудовых ресурсов (численность работников, отработанный человеко-день, человеко-час; время, необходимое для выполнения заданного объема работ исходя из действующих норм, и т. д.).

Натуральные, а иногда и трудовые измерители, не позволяют получить суммарные абсолютные показатели в условиях разнородной продукции. В этом плане универсальными являются стоимостные (денежные) единицы измерения, характеризующие стоимость определенной продукции или объема выполненных работ. В денежной форме исчисляются такие важнейшие народнохозяйственные показатели, как национальный доход, объем продукции промышленности, сельского хозяйства, строительства, транспорта, которые по своему содержанию и методам определения являются суммарными абсолютными статистическими величинами.

Абсолютные величины могут быть измерены с разной степенью точности, которая определяется сущностью измеряемого явления и задачами, поставленными перед измерением. Например, глубину скважины можно вполне измерить с точностью до одного метра, а глубину протектора автомобильного колеса следует определять с точностью до 1 мм. Сущность измеряемого явления определяет также и степень укрупнения единиц измерения. Так, население земного шара показывается в миллионах человек, а численность жителей города — в тысячах.

## 8.2. ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

Наряду с абсолютными величинами одной из важнейших форм обобщающих показателей в статистике являются относительные величины.

Относительные величины представляют собой обобщающие показатели, выражающие меру количественных соотношений,

присущих конкретным явлениям общественной жизни. При расчете относительной величины мы рассматриваем отношение двух взаимосвязанных величин (преимущественно абсолютных), т. е. измеряем их соотношение, что особенно важно в статистическом анализе. Например, общественные фонды потребления, за счет которых осуществляются бесплатное образование, медицинская помощь, материальное обеспечение по старости и другие социальные мероприятия, составили в нашей стране в 1983 г. 134,5 млрд. руб. Много это или мало? Разделим эту сумму на среднюю численность населения за этот год и получим, что в среднем на каждого жителя СССР приходилось выплата из этих фондов в размере 494 руб. Теперь мы уже имеем более глубокое представление о размере общественных фондов потребления в 1983 г. Сравнивая полученный результат с данными 1980 г. (438 руб.), найдем, что за 3 года одиннадцатой пятилетки выплаты из общественных фондов потребления на душу населения возросли более чем в 1,13 раза (494:438). Эти цифры наглядно характеризуют успехи нашей страны в решении социально-экономических задач. Относительные величины широко используются в статистическом исследовании, так как они позволяют провести сравнение различных показателей и делают такое сравнение наглядным. Относительные величины служат орудием контроля за ходом выполнения государственных планов, они помогают выявлять резервы производства и являются тем самым необходимым инструментом в процессе управления на любом уровне.

Относительные величины исчисляются как отношение двух чисел. При этом числитель называется сравниваемой величиной, а знаменатель — базой относительного сравнения. В зависимости от характера изучаемого явления и задач исследования базисная величина может принимать различные значения, что приводит к различным формам выражения относительных величин. Если база сравнения принята за 1, то относительная величина выражается числом (целым или дробным), показывающим, во сколько раз одна величина больше другой или какую часть ее составляет. На практике ее часто называют коэффициентом. Например, в одиннадцатой пятилетке необходимо увеличивать прибыль в промышленности в 1,4 раза; выплавка стали в США в 1982 г. составила менее  $\frac{1}{2}$  объема выплавки стали в СССР.

Часто базисное число принимается за 100, тогда относительные величины выражаются в процентах. Например, машиностроительному заводу запланирован рост производительности труда на 18 %, за счет чего должно быть получено 80—90 % общего прироста продукции.

При расчете относительных величин, у которых числитель значительно меньше знаменателя, целесообразно для наглядности принимать базисную величину за 1000 или за 10000.

В первом случае результат оценивается в *промилле*, во втором — в *процесимилле*. Так, на начало 1983 г. на 1000 человек населения в возрасте 10 лет и старше в СССР приходилось 678 человек, имеющих высшее и среднее образование, или 678 промилле (678 ‰); на 10 000 человек населения в 1982 г. в нашей стране приходилось 39,5 врачей, или 39,5 процесимилле (39,5 ‰).

В каждом конкретном случае выбор той или иной формы относительной величины определяется задачами исследования и социально-экономической сущностью явления, мерой которого выступает искомый относительный показатель. Относительные величины подразделяются по своему содержанию на следующие виды: динамики, планового задания, выполнения плана, структуры, координации, интенсивности и сравнения.

Относительными величинами динамики — *тепмами роста* — называются показатели, характеризующие изменение величины общественных явлений во времени. Для расчета относительных величин динамики необходимо вычислить отношение двух объемов или уровней, имевших место в разные периоды времени. Так, если в 1984 г. завод выпустил 700 станков, а в 1985 г. — 770 станков, то относительная величина динамики составит 1,1, или 110 % (770:700). Она характеризует увеличение производства станков за год по сравнению с 1984 г. на 0,1, или на 10 %. Относительные величины динамики могут исчисляться не только по сравнению с предшествующим периодом в целом, но и по отношению к соответствующей его части. Например, по данным ЦСУ СССР в первом полугодии 1984 г. добыто 288 млрд. м<sup>3</sup> газа, что составляет 109 % соответствующего периода 1983 г., т. е. означает увеличение добычи на 9 %.

В качестве базы сравнения при определении относительных величин динамики часто рассматривается не предшествующий период (или определенная часть его), а период, удаленный на длительный промежуток времени. В качестве таких отдаленных периодов берутся годы, расположенные на исторически обусловленной границе отдельных периодов времени. Например, в докладах и выступлениях руководителей КПСС и Советского правительства, в ежегодниках ЦСУ СССР часто приводятся сравнения отдельных показателей социально-экономического развития нашей страны с 1922 г. — годом образования СССР, 1940 г. — последним предвоенным годом, 1975 и 1980 гг. — последними годами девятой и десятой пятилеток.

Выбору базы сравнения при расчете относительных величин динамики, как и других относительных показателей, следует уделять особое внимание, так как от этого в существенной мере зависит практическая ценность полученного результата.

В СССР процесс общественного развития централизованно планируется. На всех уровнях управления от народного хо-

зяйства в целом до отдельного предприятия устанавливаются плановые задания по росту объема производства, производительности труда, снижению себестоимости продукции и т. д.

Относительная величина планового задания представляет собой отношение предусмотренного планом уровня или объема к соответствующему показателю, фактически достигнутому в отчетном периоде, принятому за базу сравнения:

Плановое задание на предстоящий период

Фактическое выполнение в базисном периоде

Исчисленное отношение может быть выражено в виде числа (целого или дробного) или в процентах. Так, если в 1984 г. завод выпустил 700 станков, а в 1985 г. запланирован выпуск 735 станков, то относительная величина планового задания составляет 1,05, или 105 % (735:700). Следовательно, планируется увеличить выпуск станков в 1,05 раза, или на 5 %.

На XXVI съезде КПСС, апрельском (1985 г.) Пленуме ЦК КПСС подчеркивалась необходимость безусловного выполнения установленных плановых заданий во всех звеньях хозяйственного механизма. В свете этих положений серьезно возрастает роль государственной статистики по контролю за выполнением государственных планов, выявлению резервов, имеющихся в социалистическом производстве.

Относительные величины выполнения плана отражают степень выполнения планового задания и представляют собой отношение фактического выполнения плана к установленному плановому заданию:

Фактическое выполнение плана

Плановое задание

При этом необходимо, чтобы числитель и знаменатель исходного отношения соответствовали одному и тому же периоду или моменту времени. Если, например, в 1985 г. завод выпустил 770 станков при плановом задании 735 станков, то, следовательно, степень выполнения плана составит 104,8 % (770:735) · 100. Таким образом, в 1985 г. завод перевыполнил план производства станков на 4,8 %.

Если плановые задания установлены в относительных величинах, то при расчете относительной величины выполнения плана следует соотносить процентное отношение фактически достигнутого уровня к базисному уровню и показатель планового задания. Так, если в 1983 г. производительность труда в промышленности СССР была увеличена на 3,5 % при пла-



новом росте 2,9 %, то относительная величина выполнения плана составит:

$$\frac{100 + 3,5}{100 + 2,9} \cdot 100 = \frac{103,5}{102,9} \cdot 100 = 100,6 \%$$

Следовательно, в данном случае план по росту производительности труда перевыполнен на 0,6 %.

В статистическом анализе встречаются ситуации, когда приходится оценивать степень выполнения плана по показателю, величину которого планировалось снизить (трудоемкость либо себестоимость единицы продукции). В таких ситуациях приходится отойти от привычного подхода («больше — лучше»), так как увеличение трудоемкости, себестоимости и других подобных показателей свидетельствует об ухудшении работы. Проиллюстрируем это положение следующим примером. Производственному объединению был установлен план снижения трудоемкости единицы продукции на 7 %, фактическое снижение составило лишь 3 %. Относительная величина выполнения плана составляет:

$$\frac{100 - 3}{100 - 7} \cdot 100 = \frac{97}{93} \cdot 100 = 104,3 \%$$

В этом случае, однако, производственное объединение не услышит похвал в свой адрес за перевыполнение плана, ведь фактическая трудоемкость оказалась выше плановой на 4,3 % ( $104,3 - 100$ ), т. е. план по снижению трудоемкости недовыполнен на 4,3 %.

На современном этапе развития экономики нашей страны в целях усиления контроля за ходом выполнения государственных планов как внутри отдельного года, так и по годам пятилетки целесообразно шире использовать расчет относительной величины выполнения плана методом нарастающего итога.

Допустим, в 1984 г. заводу был установлен план выпуска электромоторов в количестве 1000 шт., в том числе по кварталам: 220, 240, 270 и 270 шт. Фактическое выполнение составило за год 980 шт. в том числе по кварталам: 220, 200, 270 и 290 шт.

Относительная величина выполнения плана в нарастающих итогах составит: I кв.  $(220 : 1000) 100 = 22 \%$ ; II кв.  $(220 + 200) : 1000 \cdot 100 = 44 \%$ ; III кв.  $(220 + 200 + 270) : 1000 \cdot 100 = 69 \%$ ; за год  $(220 + 200 + 270 + 290) : 1000 \cdot 100 = 98 \%$ .

Очевидно, что полученные результаты нагляднее характеризуют ход выполнения плана в течение года по сравнению с относительными величинами выполнения плана за каждый квартал и поэтому более пригодны для оперативного руководства. Расчеты показывают, что завод недовыполнил план выпуска электромоторов на 2 %, недопоставив народному хозяйству 20 единиц продукции. Анализ показывает, что причиной

такого положения явилось существенное невыполнение плана во II квартале. Аналогично можно определить степень выполнения плана в каждом году пятилетки и нарастающую долю выполнения плана за пятилетку в целом.

Между относительными величинами динамики, планового задания и выполнения плана существует взаимосвязь. Для того чтобы выразить ее в общем виде, введем следующие обозначения: базисный уровень —  $y_0$ , плановый уровень —  $y_{пл}$ , фактический уровень —  $y_1$ . В принятых обозначениях относительные величины могут быть выражены следующими соотношениями:

$$\text{относительная величина динамики} = \frac{y_1}{y_0}; \quad (8.1)$$

$$\text{относительная величина планового задания} = \frac{y_{пл}}{y_0}; \quad (8.2)$$

$$\text{относительная величина выполнения плана} = \frac{y_1}{y_{пл}}. \quad (8.3)$$

Тогда очевидно, что  $\frac{y_1}{y_0} = \frac{y_{пл}}{y_0} \cdot \frac{y_1}{y_{пл}}$ , т. е. относительная величина динамики равна произведению относительных величин планового задания и выполнения плана, разумеется, при условии использования одной и той же базы сравнения. В приводимом ранее примере о выпуске станков определим относительную величину динамики умножения относительных величин планового задания и выполнения плана:  $1,05 \times 1,048 = 1,1$ .

Полученное соотношение имеет большое практическое значение, так как позволяет определить любую из трех величин при двух известных.

Если, например, районное агропромышленное объединение увеличило в 1985 г. выпуск консервов по сравнению с 1984 г. на 7,8 % при плановом задании 5,2 %, то относительная величина выполнения плана может быть найдена из следующего соотношения:

$$\frac{y_{ф}}{y_0} = \frac{y_{пл, ф}}{y_0} \cdot \frac{y_{ф}}{y_{пл, ф}}$$

$$\text{Откуда } \frac{y_{ф}}{y_{пл, ф}} = \frac{y_{ф}}{y_0} : \frac{y_{пл, ф}}{y_0} = \frac{100 + 7,8}{100 + 5,2} \cdot 100 = 102,5 \%$$

Следовательно, план 1985 г. был перевыполнен на 2,5 %. Относительные величины структуры характеризуют долю отдельных частей в общей совокупности (например, долю городского населения во всем населении страны; долю продукции, удостоенной государственного Знака качества; долю студентов-отличников в академической группе и т. п.).

Относительные величины структуры, обычно называемые удельными весами, рассчитывают делением определенной

части целого на общий итог, принимаемый за 100 %. Рассмотрим в качестве примера структуру валового общественного продукта СССР (табл. 8.1).

В приведенной таблице показатели, расположенные в графах 2 и 4, представляют собой относительные величины структуры, ибо они характеризуют удельные веса продукции отдельных отраслей народного хозяйства в валовом общественном продукте. Однако таблица показывает не только структуру валового об-

Таблица 8.1. Валовой общественный продукт по отраслям народного хозяйства

	1975		1982		Темп роста ( $\frac{y_{82}}{y_{75}} \cdot 100$ ), %
	Сумма млрд. руб.	% к итогу	Сумма млрд. руб.	% к итогу	
А	1	2	3	4	5
Валовой общественный продукт в том числе по отраслям народного хозяйства:	862,6	100,0	1236,0	100,0	143,3
промышленность	558,3	64,7	792,7	64,1	142,0
сельское хозяйство	122,3	14,2	170,3	13,8	139,2
транспорт и связь	36,7	4,3	55,2	4,5	150,4
строительство	91,7	10,6	115,1	9,3	125,5
торговля, заготовки, материально-техническое снабжение и другие отрасли	53,6	6,2	102,7	8,3	191,6

щественного продукта в 1975 и 1982 гг. Она позволяет проследить за структурными сдвигами, т. е. за изменениями удельных весов продукции в общем объеме валового общественного продукта отдельных отраслей народного хозяйства. Так, в 1982 г. по сравнению с 1975 г. снизился удельный вес в валовом общественном продукте таких отраслей, как промышленность, сельское хозяйство, строительство. Обратим внимание на то, что при росте общей суммы валового общественного продукта на 43,3 % указанные отрасли характеризуются более низкими относительными величинами динамики (темпами роста). Другие же отрасли, динамика которых опережала темп роста всего валового общественного продукта (общего итога), увеличили за исследуемый период свои удельные веса. Наблюдаемые изменения удельных весов иллюстрируют закономерность, состоящую в том, что при рассмотрении структуры одной и той же совокупности за ряд периодов увеличиваются удельные веса тех составных частей, которые растут быстрее целого и соответственно они снижаются у тех частей, динамика которых ниже динамики целого. Для доказательства этой закономерности в общем виде обозначим базисный уровень определенной части целого через  $y_{i0}$ , а фактический —  $y_{it}$ , соответственно целое:  $\Sigma y_{i0}$  и  $\Sigma y_{it}$ .

Тогда удельные веса  $i$ -й части целого в двух периодах выражаются отношениями  $d_{i0} = \frac{y_{i0}}{\Sigma y_{i0}}$  и  $d_{it} = \frac{y_{it}}{\Sigma y_{it}}$ . Рассмотрим относительную величину динамики (темпа роста —  $T$ ) удельных весов:

$$T_d = \frac{y_{it}}{\Sigma y_{it}} : \frac{y_{i0}}{\Sigma y_{i0}} = \frac{y_{it}}{y_{i0}} : \frac{\Sigma y_{it}}{\Sigma y_{i0}} = T_{yi} : T_{\Sigma y}$$

Таким образом, темп роста удельного веса определяется отношением темпов роста части и целого. Следовательно, если  $T_{yi} > T_{\Sigma y}$  (часть изменяется во времени быстрее целого), то  $T_d > 1$  (удельный вес этой части увеличивается), в противном случае ( $T_{yi} < T_{\Sigma y}$ ) удельный вес рассматриваемой части по сравнению с базисным периодом снижается.

Эту закономерность можно проследить и на таком примере. За период с 1950 по 1981 г. мировое производство промышленной продукции возросло в 6,4 раза, в том числе в странах социализма в 15 раз, а в остальных странах — в 4,3 раза. Это привело к тому, что за рассматриваемые годы доля продукции социалистических стран в мировой промышленной продукции увеличилась с 20 до 40 %, а удельный вес остальных стран соответственно снизился с 80 до 60 %. Приведенные цифры наглядно показывают победоносное шествие социализма.

Важным моментом расчета относительной величины структуры, как и других относительных величин, является правильный выбор базы сравнения, т. е. того целого, часть которого мы определяем. В противном случае можно получить ошибочные результаты. Например, завод, имеющий 200 станков, получил в начале года 10 новых высокопроизводительных станков, из которых 4 не были установлены до конца года по вине завода и, следовательно, не давали продукции. Каков удельный вес неиспользованных высокопроизводительных станков? Если в качестве базы сравнения взять весь имеющийся станочный парк — 210 (200+10), то относительная величина структуры составит всего 1,9 % ( $4:210 \cdot 100$ ) и особой тревоги это, вроде бы, вызвать не должно. Но такой подход, разумеется, не является верным, ведь на заводе есть всего 10 новых высокопроизводительных станков. Следовательно, действительная доля неиспользованного нового эффективного оборудования равна 40 % ( $4:10 \cdot 100$ ). Таким образом, завод явно неэффективно использует новые высокопроизводительные станки, и следует принять срочные меры по исправлению сложившегося положения.

Относительные величины координации характеризуют отношение отдельных частей совокупности с одной из них, принятой за базу сравнения. Показателями координации является, например, число городских жителей, приходящихся на 100 сельских; число женщин, приходящихся на 100 мужчин; количество гектаров пастбищ на 100 га пашни и т. п. Характеризуя соотношение между отдельными частями целого, относительные величины

координации придают им наглядность и позволяют, если это возможно, контролировать соблюдение оптимальных пропорций. Так как числитель и знаменатель относительных величин координации имеют одинаковую единицу измерения, то эти величины выражаются не в именованных числах, а в процентах, промилле или кратных отношениях. Например, на начало 1983 г. в нашей стране на 100 мужчин приходилось 114 женщин. Такое соотношение сложилось за счет старших возрастов и вызвано главным образом последствиями второй мировой войны.

Характерной особенностью относительных величин координации является то, что из двух сравниваемых частей за базу может браться любая из них. Так, по данным табл. 8.1 можно определить, что в 1982 г. на 1 млрд. руб. валового общественного продукта, созданного в сельском хозяйстве, приходится 4,65 млрд. руб. в промышленности (792,7 : 170,3). С другой стороны, можно утверждать, что на 1 млрд. руб. в промышленности, в сельском хозяйстве приходится 0,21 млрд. (170,3 : 792,7) валового общественного продукта. На практике, в целях достижения большей наглядности, при расчете относительных величин координации чаще используется деление большего числа на меньшее.

**Относительными величинами интенсивности** называются показатели, характеризующие меру распространения или развития данного явления в определенной среде. Они рассчитываются как отношение абсолютной величины данного явления к размеру среды, в которой оно развивается. Относительные величины интенсивности находят широкое применение в практике статистики. Это показатели сельскохозяйственного производства на 100 га земли; фондоотдача (выпуск продукции в расчете на 1 руб. стоимости основных производственных фондов); обеспеченность населения врачебной помощью (численность врачей на 10 000 человек населения); уровень производительности труда (выпуск продукции на одного работника или в единицу рабочего времени); рентабельность производства (прибыль в расчете на 1 руб. стоимости основных производственных фондов и нормируемых оборотных средств) и т. п. Таким образом, относительные величины интенсивности характеризуют эффективность использования различного рода ресурсов (материальных, финансовых, трудовых), социальный и культурный уровни жизни народа, многие другие аспекты общественной жизни.

Относительные величины интенсивности в отличие от других видов относительных величин являются обычно именованными числами и имеют размерность тех абсолютных величин, соотношение которых они отражают. Вместе с тем в ряде случаев, когда полученные при расчете результаты достаточно малы, их умножают для наглядности на 1000 или 10 000, получая характеристики в промилле или проценцимилле. Например, в 1982 г. в СССР на каждую 1000 жителей родилось 18,9 ребенка (коэффициент рождаемости — 18,9 ‰), а на 10 000 человек населения приходилось 39,5 врача (обеспеченность населения врачебной помощью — 39,5 ‰/1000).

Особый интерес представляет разновидность относительных величин интенсивности — **производство продукции на душу населения**. Эти показатели представляют собой отношение количества определенной продукции к средней численности населения за соответствующий год. Если общее производство того или иного вида продукции, выраженное абсолютной величиной, характеризует уровень технического развития страны в определенном периоде, то производство этой продукции на душу населения, выраженное относительной величиной интенсивности, отражает уровень ее экономического развития. В табл. 8.2 приведены данные о производстве некоторых видов продукции на душу населения в СССР за ряд лет и в США.

Из таблицы видно, что в СССР производство некоторых видов продукции на душу населения возросло по сравнению

Таблица 8.2. Производство некоторых видов продукции на душу населения СССР и США

Наименование продукции	СССР			США
	1913	1940	1982	1982
Электроэнергия, кВт·ч	13	249	5063	10 690
Сталь, кг	27	94	545	294
Нефть, кг	65	159	2268	1832
Минеральные удобрения, кг усл. ед.	0,6	17	408 <sup>1</sup>	491 <sup>1</sup>
Сахар-песок (из отечественного сырья), кг	8,6	11	25	21
Молоко, кг	185	172	337	265

<sup>1</sup> 1981 г.

с 1913 г. в десятки и сотни раз. По производству на душу населения стали, нефти, молока и других видов продукции Советский Союз обогнал США.

**Относительными величинами сравнения** называются относительные показатели, получающиеся в результате сравнения одноименных уровней, относящихся к различным объектам или территориям, взятым за один и тот же период или на один момент времени. Обычно они исчисляются в процентах или в кратных отношениях, показывающих, во сколько раз, одна сравниваемая величина больше или меньше другой. В качестве примера приведем соотношение производства некоторых видов промышленной продукции в двух странах — СССР и США, приняв в качестве базы сравнения уровень производства США. Для наглядности не ограничиваясь расчетом относительных величин сравнения за какой-либо один год, а рассмотрим их изменение за ряд лет (табл. 8.3).

Данные таблицы характеризуют гигантские успехи нашей страны. Если в 1940 г. по отдельным видам промышленной продукции Советский Союз производил, как видно из табл. 8.3, от



17 до 33 % объема США, то в 1982 г. мы уверенно опережаем уровень американского производства по всем позициям рассматриваемой совокупности (исключая производство электроэнергии), занимая по ним первое место в мире.

Относительные величины сравнения находят широкое применение при сравнительной оценке различных показателей работы отдельных предприятий, городов, союзных республик. При

Таблица 8.3. Соотношение производства важнейших видов промышленной продукции СССР и США (СССР в процентах к США)

Виды промышленной продукция	1940	1950	1960	1965	1970	1982
Сталь	29	30	71	75	95	216
Нефть	17	14	42	63	74	144
Электроэнергия	26	22	33	41	43	55
Шерстяные ткани	33	92	109	124	252	496

этом, например, результаты работы конкретного предприятия могут последовательно сравниваться с аналогичными предприятиями других городов, со среднеотраслевыми показателями и, наконец, с лучшими отечественными и зарубежными достижениями, что особенно важно для определения путей повышения эффективности производства.

### 8.3. ОБЩИЕ ПРИНЦИПЫ ПОСТРОЕНИЯ И ИСПОЛЬЗОВАНИЯ АБСОЛЮТНЫХ И ОТНОСИТЕЛЬНЫХ СТАТИСТИЧЕСКИХ ВЕЛИЧИН

Абсолютные и относительные величины, являясь важной разновидностью обобщающих показателей, призваны давать числовые характеристики массовых общественных явлений и процессов. Успешное, научно обоснованное применение этих величин в статистическом исследовании предполагает соблюдение некоторых общих принципов.

Важнейшей предпосылкой правильной количественной характеристики общественных явлений и процессов является глубокий качественный анализ, позволяющий выявить их сущность, специфические особенности, взаимосвязи с другими явлениями и процессами. При отрыве количественной характеристики от качественной стороны изучаемого объекта получение правильных, объективных выводов невозможно.

При сравнении абсолютных и рассчитанных на их основе относительных величин важным является обеспечение их сопоставимости. В этом плане прежде всего необходимо, чтобы сравниваемые величины были рассчитаны по одинаковой методологии. Нельзя, например, сравнивать уровни производительности труда на двух предприятиях, если на одном выпущенная продукция была отнесена к численности промышленного произ-

водственного персонала, а на втором — к числу рабочих. Этому важному моменту обеспечения сопоставимости данных большое внимание уделял в своих работах В. И. Ленин. В работе «О статистике стачек в России», рассмотрев данные о стачечном движении с 1905 по 1906 г. и с 1906 по 1907 г., В. И. Ленин указывал, что официальная статистика приводит по этому поводу не вполне сравнимые показатели, что смешанные стачки, вызванные и политическими, и экономическими требованиями, в 1905 г. считались политическими, а позднее — экономическими. Сопоставимость данных с позиций единой методологии расчета особенно важна в международных сравнениях. Так, показатели выработки электроэнергии, публикуемые официальной статистикой США, непосредственно несопоставимы с данными СССР, так как в США производство показывается без внутреннего потребления электростанциями, а в СССР — вместе с потреблением электростанций на собственные нужды. Поэтому при составлении статистических ежегодников, справочников эти и другие случаи несопоставимости данных разных стран учитываются и приводится цифровой материал, рассчитанный по единой методологии.

Сопоставимость предполагает также, что сравниваться должны одинаковый круг объектов. Нарушение этого требования возникает обычно в результате каких-либо организационных изменений. Например, в начале нового года производственному объединению было передано два новых ранее самостоятельных предприятия. В данном случае несопоставимость данных в динамике очевидна, ведь рост объема продукции по объединению будет результатом не только вскрытия неиспользованных резервов, повышения эффективности производства, но и следствием увеличения мощности объединения за счет новых предприятий. Этот момент, конечно же, следует учитывать при исчислении различных абсолютных и относительных величин, характеризующих работу производственного объединения, начиная с момента увеличения числа входящих в него предприятий.

Важно также, чтобы сравниваемые величины измерялись в одних и тех же единицах. Нельзя, например, сравнивать продукцию двух тракторных заводов, если по одному есть данные о производстве в штуках, а по другому — информация о суммарной мощности в лошадиных силах. При сравнении показателей, имеющих стоимостную оценку, необходимо, чтобы в их расчетах были использованы сопоставимые (неизменные) цены.

При пространственном сравнении показателей, относящихся к различным объектам, должны быть взяты за одно и то же время или на одну дату. Неверно, например, давать сравнительную оценку двух соревнующихся областей, предприятий, колхозов, если при этом рассматриваются различные периоды их работы.

Статистические данные должны быть сопоставимы и по продолжительности периода, к которому они относятся. Так, при

подведении итогов выполнения Государственного плана экономического и социального развития СССР в первом полугодии 1984 г. наряду с определением степени выполнения плана достигнутые результаты сравнивают с соответствующим периодом 1983 г. По данным ЦСУ СССР общий объем промышленного производства, например, увеличился по сравнению с первым полугодием 1983 г. на 4,5 %, или на 16,6 млрд. руб. Из этих данных можно заключить, что абсолютная величина одного процента прироста, т. е. его «весомость», составила в рассматриваемом периоде 3,69 млрд. руб. (16,6:4,5). Особенность найденного показателя состоит в том, что он позволяет увязать абсолютную и относительную величины и тем самым с большей наглядностью охарактеризовать изучаемое явление. Действительно, взятые отдельно друг от друга величины 4,5 % и 16,6 млрд. руб. сообщают нам значительно меньше информации о действительных размерах роста продукции промышленности, чем при совместном их рассмотрении.

Таким образом, мы вплотную подошли к важному принципу правильного использования абсолютных и относительных величин — необходимости их комплексного применения. Это условие вытекает прежде всего из сущности относительных величин, получаемых в результате отношения двух абсолютных величин.

Относительные величины не заменяют абсолютных показателей, которые с их помощью сравниваются. Если, например, два предприятия выполнили план на 100 %, то, опираясь только на эту информацию, мы ничего не можем сказать об их действительных объемах производства. Следовательно, при использовании относительных величин необходимо постоянно держать в поле зрения абсолютные величины, послужившие основой для их расчета. В. И. Ленин в работе «К вопросу о политике министерства народного просвещения» отмечал, что «...до смешного маленькие цифры в процентном исчислении их возрастания растут всегда с „громдным“ быстротой. Если нищему, имеющему три копейки, вы дадите пятачок, увеличение его „имущества“ сразу будет „громдное“: на целых 267 %»<sup>1</sup>.

Таким образом, лишь комплексное применение абсолютных и относительных величин позволяет наглядно и всесторонне охарактеризовать количественную сторону массовых общественных явлений и процессов в неразрывной связи с качественной стороной.

#### 8.4. ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ ОТНОСИТЕЛЬНЫХ ВЕЛИЧИН

В практике экономической работы для контроля за ходом выполнения плана часто используются *линейные графики*. Одним из видов их является *график нарастающих итогов с сопереженными шкалами* (рис. 8.1). Такой график дает наглядное

представление о состоянии выполнения плана на определенный момент времени.

Для графического изображения структуры явлений употребляются *столбиковые, полосовые* и чаще всего *секторные диаграммы*. Выбор определенного вида диаграмм связан с особен-

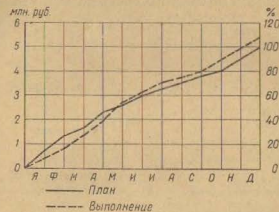


Рис. 8.1. График выполнения годового плана

ностями изображаемых явлений, а также с целью, которую преследуют при построении графика.

Чаще всего структура явления изображается в сравнении со структурой аналогичного явления в другом месте или в другое время. При таких сравнениях используются *структурные диаграммы* двух видов — *структурно-относительные* и *структурно-абсолютные*. Первые служат только для сопоставления относительных величин структуры или удельных весов (рис. 8.2 и 8.4), вторые — для сопоставления абсолютных размеров сравниваемых явлений (рис. 8.3). Построение полосовых диаграмм относительных величин (рис. 8.2 и 8.3) производится так же, как и столбиковых диаграмм сравнения.

При построении секторных диаграмм (рис. 8.4) изображаемые показатели пропорциональны площадям соответствующих секторов или, что то же самое, — величинам центральных углов этих секторов.

На рис. 8.4, например, сопоставлены удельные веса бурого угля в общем количестве добытого в 1980 г. товарного угля в СССР и ЧССР.

Таблица 8.4. Добыча угля в СССР и ЧССР в 1980 г.

	СССР	ЧССР
Товарный уголь, млн. т	653	123
в том числе:		
каменный	493	28
бурый	160	95

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 23, с. 125.

Исходные данные таковы. (См. табл. 8.4).

Удельный вес бурого угля, добытого в 1980 г. в СССР, составляет 24,5 %, (160 : 653 = 0,245), а в ЧССР — 77,2 % (95 : 123 = 0,772).

Центральные углы определим из тех соображений, что они должны быть пропорциональны изображаемому удельным весам. Так, для СССР удельный вес бурого угля 24,5 % соответствует центральному углу  $88^\circ$  ( $(360 : 100) \cdot 24,5$ ), а для ЧССР соответствующий центральный угол составит  $272^\circ$  (см. рис. 8.4).

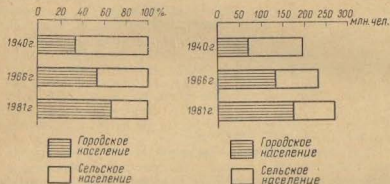


Рис. 8.2. Удельный вес городского населения в общей численности населения СССР (относительная структура)

Рис. 8.3. Население СССР в 1940—81 гг. (абсолютная структура)

Если кроме сопоставления удельных весов целесообразно дать представление и об абсолютном изменении целого и его частей, то площади кругов и их секторов должны быть в этом случае пропорциональны размерам изображаемых величин.

Размещение (распределение) явлений по территориям изображается графически с помощью картодиаграмм и картограмм. Картодиаграмма представляет собой контурную карту или план территории, на отдельные районы которой нанесены диаграммы (столбиковые, круговые и т. д.). На картодиаграммах обычно изображаются суммарные (итоговые) значения абсолютных статистических показателей: численность населения, объем выпускаемой продукции, площадь пахотной земли и т. п. Нередко при этом показывается и структура явления (в виде частей столбиков или секторов кругов).

Картограмма также представляет собой контурную карту или план территории, на которую с помощью точек, фоновых знаков либо изолиний наносится значения показателей, характе-

ризующих отдельные части изображаемой территории. Соответственно различают точечные, фоновые и изолинейные картограммы. В точечных картограммах значения статистических показателей наносятся с помощью точек, количество которых соответствует размерам изображаемых явлений.

На фоновых картограммах статистические величины обозначаются различного вида штриховкой или раскраской территориальных единиц (рис. 8.5). Эти картограммы в основном предназначены для изображения относительных и средних показателей (плотность населения, урожайность, выработка электроэнергии на душу населения и т. п.). При построении таких картограмм вначале выбираются территориальные единицы, характеризующиеся отдельным показателем или пределами его изменения (например, 6—10 чел./км<sup>2</sup>). При большом числе территориальных единиц (превышающем 6—8) их следует группировать по величине изображаемого показателя на несколько групп, для каждой из которых выбирается определенный вид раскраски или штриховки. Число групп при этом не должно превышать 6—8, так как в противном случае картограмма потеряет выразительность и наглядность. При построении фоновых картограмм следует придерживаться правила: большим величинам изображаемого показателя должна соответствовать более густая штриховка или более интенсивная окраска.

Изолинейная диаграмма строится путем проведения на контурной карте изолиний, т. е. линий, соединяющих места с одинаковой величиной показателя. Такие картограммы получили гораздо меньшее распространение, чем точечные и фоновые.

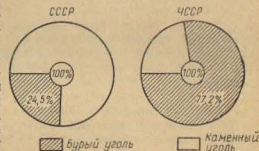


Рис. 8.4. Удельный вес бурого угля в общем количестве товарного угля, добытого в 1980 г. в СССР и ЧССР

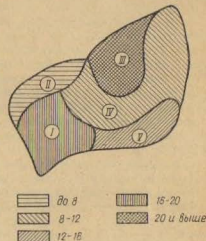


Рис. 8.5. Картограмма плотности населения районов области, чел./км<sup>2</sup>



### 9.1. СУЩНОСТЬ СРЕДНИХ И ИХ ЗНАЧЕНИЕ В СТАТИСТИЧЕСКОМ АНАЛИЗЕ

Средние величины играют в статистике очень важную роль, поскольку они являются обобщенной характеристикой большого количества индивидуальных значений варирующего признака. В экономическом анализе их можно считать наиболее употребительными обобщающими показателями.

Известно, например, что заработная плата работника определяется его квалификацией, характером труда, возрастом, стажем работы и целым рядом других факторов, поэтому изменяется в весьма широких пределах. Тем не менее можно говорить о среднемесячной заработной плате рабочих и служащих всех отраслей народного хозяйства, составившей в 1983 г. 182 руб. Здесь мы оперируем типичным, характерным значением варирующего признака, отнесенным к единице многочисленной совокупности.

Средняя величина отражает то общее, что характерно для всех единиц изучаемой совокупности. В то же время она игнорирует те различия, которые наблюдаются у отдельных единиц совокупности, как бы взаимно погашая их. Уровень (или размер) любого общественного явления обусловлен действием двух групп факторов. Одни из них являются главными, постоянно действующими, тесно связанными с природой изучаемого явления или процесса, и формируют то типичное для всех единиц изучаемой совокупности, которое и отражается в средней величине. Другие являются второстепенными, их действие выражено слабее и носит эпизодический, случайный характер. Они обуславливают различия между количественными признаками отдельных единиц совокупности, погашаемые в средней величине. В этом проявляется в общем виде действие известного из математической статистики фундаментального закона больших чисел. По этому поводу К. Маркс писал: «Труд, овеществленный в стоимости, есть труд среднего общественного качества, т. е. проявление средней рабочей силы. Но средняя величина есть всегда средняя многих различных индивидуальных величин одного и того же вида. В каждой отрасли промышленности индивидуальный рабочий, Петр или Павел, более или менее отклоняется от среднего рабочего. Такие индивидуальные отклонения, называемые на языке математиков «погрешностями», взаимно погашаются и уничтожаются, раз мы берем значительное число рабочих»<sup>1</sup>.

Таким образом, средняя отражает общее, характерное и типичное для всей совокупности благодаря взаимопогашению в ней случайных, нетипичных различий между признаками отдельных

ее единиц. Однако для того, чтобы средняя величина была действительно типизирующим показателем, она должна определяться не для любых совокупностей, а только для совокупностей, состоящих из качественно однородных единиц. Это является основным условием научно обоснованного использования средних величин и предполагает тесную связь метода средних и метода группировок в анализе социально-экономических явлений.

Средние, полученные для неоднородных совокупностей, не только не будут иметь научной ценности, но даже могут принести вред, искажая истинный характер изучаемого общественного явления. В. И. Ленин, широко использовавший в своих работах статистические методы, отмечал огульный, фиктивный характер подобных средних<sup>1</sup>.

Буржуазная статистика применяла и применяет фиктивные «средние» показатели, затуманивающие подлинную картину реальных процессов и явлений общественной жизни и дающие искаженное представление о них.

Итак, под средней величиной в статистике понимается обобщающий показатель, характеризующий типичный уровень варирующего признака в расчете на единицу однородной совокупности.

В экономическом анализе использование средних величин является действенным инструментом для оценки результатов научно-технического прогресса, социальных мероприятий, изыскания скрытых и неиспользованных резервов производства. В то же время следует помнить о том, что чрезмерное увлечение средними показателями может привести к необъективным выводам при проведении статистико-экономического анализа. Это связано с тем, что средние величины, будучи обобщающими показателями, погашают, игнорируют те различия в количественных признаках отдельных единиц совокупности, которые реально существуют и могут представлять самостоятельный интерес. Так, например, среднее выполнение норм выработки всеми рабочими завода на 103,6 % может маскировать тот факт, что наряду с рабочими, значительно перевыполняющими норму выработки, есть определенная часть рабочих, не выполняющих эту норму. Поэтому использовать средние показатели следует осторожно, всегда помня о том, что за средней величиной скрывается большое количество индивидуальных значений изучаемого признака у отдельных единиц совокупности.

### 9.2. СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ, ЕЕ ВЫЧИСЛЕНИЕ И ОБЛАСТЬ ПРИМЕНЕНИЯ

В статистике используются различные виды средних величин. Самый распространенный вид средней — средняя арифметическая. Поэтому, когда речь идет о средней величине и не

<sup>1</sup> Маркс К., Энгельс Ф. Соч., т. 23, с. 334.

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 2, с. 374.

указывается ее вид, чаще всего подразумевается именно средняя арифметическая. В общем случае ее вычисление сводится к суммированию всех значений варьирующего признака и делению полученной суммы на общее количество единиц совокупности. Пусть, например, из пяти рабочих-токарей, выполняющих одинаковую работу, один за смену изготовил 12 деталей, второй—9, третий—11, четвертый—13 и пятый—15. Тогда средняя выработка одного токаря за смену определится как общее количество изготовленных ими деталей  $(12+9+11+13+15=60)$ , отнесенное к общему количеству токарей:  $\bar{x}=60:5=12$  штук.

В данном случае расчет проводился по формуле средней арифметической простой:

$$\bar{x} = \frac{\sum x}{n}, \quad (9.1)$$

поскольку в исходных данных значение каждого варианта встречалось только один раз. В тех же случаях, когда значения вариантов встречаются по несколько раз, следует пользоваться средней арифметической взвешенной:

$$\bar{x} = \frac{\sum xf}{\sum f}, \quad (9.2)$$

где  $x$  — варианты осредняемого признака,  $f$  — веса этих вариантов.

В качестве весов чаще всего используются либо частоты, т. е. числа, показывающие, сколько раз встречаются в совокупности соответствующие варианты, либо частоты (отношение частот к их сумме). Однако в ряде случаев в качестве весов в формуле (9.2) используются и другие показатели, о чем будет сказано ниже.

Допустим, на строительной площадке работает 10 подъемных кранов. Один из них имеет грузоподъемность 40 т, 2—по 25 т, 3—по 10 т и 4—по 5 т. Тогда средняя грузоподъемность одного крана составит:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{40 \cdot 1 + 25 \cdot 2 + 10 \cdot 3 + 5 \cdot 4}{10} = 14 \text{ т.}$$

Нетрудно видеть, что между средними арифметическими простой и взвешенной нет принципиальной разницы, поскольку многократное суммирование значений одного варианта всегда может рассматриваться как произведение его на частоту (или вес).

Средняя арифметическая применяется в тех случаях, когда общий объем варьирующего признака для всей совокупности образуется как сумма значений признаков у отдельных ее единиц. Этим, собственно, определяется область применения средней арифметической и объясняется ее распространенность как обобщающего показателя, поскольку для общественных явлений характерна аддитивность объемов варьирующих признаков. Так, например, общий фонд заработной платы — это сумма

заработных плат всех работников, общий сбор урожая — сумма урожаев с отдельных участков, общее число рабочих отрасли — сумма их численностей по отдельным предприятиям и объединениям.

Определение средней арифметической в ряде случаев сопряжено с большими затратами времени и труда. Однако процедуру расчета средней можно упростить и облегчить, если воспользоваться некоторыми ее свойствами. Приведем (без доказательства) основные свойства средней арифметической.

1. Средняя от постоянной величины равна ей самой:

$$\bar{A} = A.$$

2. Произведение средней на сумму частот равно сумме произведений вариантов на частоты:

$$\bar{x} \sum f = \sum xf. \quad (9.3)$$

3. Изменение каждого варианта на одну и ту же величину изменяет среднюю на ту же величину

$$\frac{\sum (x \pm A)f}{\sum f} = \bar{x} \pm A. \quad (9.4)$$

4. Изменение каждого варианта в одно и то же число раз изменяет среднюю во столько же раз

$$\frac{\sum Ax f}{\sum f} = A \bar{x}. \quad (9.5)$$

5. Изменение каждого из весов в одно и то же количество раз не изменяет величины средней

$$\frac{\sum x(Af)}{\sum Af} = \frac{A \sum xf}{A \sum f} = \frac{\sum xf}{\sum f} = \bar{x}. \quad (9.6)$$

6. Алгебраическая сумма отклонений всех вариантов от средней равна нулю:

$$\sum (x - \bar{x})f = 0. \quad (9.7)$$

7. Средняя суммы равна сумме средних:

$$\overline{x+y} = \bar{x} + \bar{y}. \quad (9.8)$$

8. Сумма квадратов отклонений вариантов от средней арифметической меньше, чем от любой другой величины

$$\sum (x - \bar{x})^2 = \min. \quad (9.9)$$

Рассмотрим теперь основные случаи практического вычисления средней арифметической при различной исходной информации.

1. Имеются полученные из наблюдения значения варьирующего признака. Вычисления средней арифметической производится путем суммирования этих значений и последующего деления полученной суммы на их количество. В этом случае пользуются формулой средней арифметической простой либо,

если варианты обладают повторяемостью, формулой средней арифметической взвешенной (см. приведенные выше примеры).

2. Имеются не отдельные значения варьирующего признака, а сумма их (общий объем признака) и количество единиц совокупности (подобные данные часто приводятся в статистической отчетности). В этом случае расчет средней сводится к делению общего объема признака на количество единиц совокупности. Например, если известен месячный фонд заработной платы по отдельному цеху, равный 4087,5 руб., и численность рабочих этого цеха, равная 25 человек, то средняя зарплата одного рабочего составит

Таблица 9.1. Распределение рабочих по квалификации

Классификационный разряд (варианты) $x$	Количество рабочих (частоты) $f$	Произведение вариантов на частоты $xf$
1	23	23
2	49	98
3	63	189
4	78	312
5	57	285
6	18	108
Всего	288	1015

25 человек, то средняя зарплата одного рабочего составит

$$\bar{x} = \frac{4087,5}{25} = 163,5 \text{ руб. (9.10)}$$

3. Весьма часто при проведении статистических исследований приходится вычислять средние величины по данным вариационных рядов. Если ряд является дискретным, то для вычисления средней нужно значения вариантов умножить на соответствующие частоты и сумму этих произведений разделить на сумму частот.

Пусть, например, результаты обследования показали следующее распределение рабочих промышленного предприятия по их квалификации (табл. 9.1).

Средний разряд рабочих на этом предприятии будет равен:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{1015}{288} = 3,52. \quad (9.11)$$

Рассмотрим теперь вычисление средней арифметической по данным интервального вариационного ряда. В подобных рядах значения вариантов приводятся в виде интервала «от ... до ...». В этом случае для каждой группы находится среднее значение интервала как полусуммы его верхней и нижней границ. Эти средние значения интервалов и будут новыми значениями вариантов, подлежащими усреднению.

Обратимся для примера к условным данным о распределении 1000 рабочих цеха по стажу работы и определим их средний стаж (табл. 9.2).

Средний стаж работы составит:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{9270}{1000} = 9,27 \text{ года.}$$

Если в рассматриваемом ряду имеются интервалы с так называемыми открытыми границами (в приведенном примере — последний интервал с неизвестным верхним пределом), то для нахождения их середины ориентируются на ширину смежного интервала (в данном случае — предыдущего). Следует заметить, что такое преобразование интервального ряда в дискретный приводит к некоторой неточности в определении средней. Это

Таблица 9.2. Распределение рабочих цеха по стажу работы

Группы рабочих по стажу работы, лет	Среднее значение интервалов $x$	Количество рабочих $f$	Произведение вариантов на частоты $xf$
0—5	2,5	146	365,0
5—10	7,5	495	3712,5
10—15	12,5	237	2962,5
15—20	17,5	103	1802,5
20 и более	22,5	19	427,5
Всего	—	1000	9270,0

связано с использованием допущения о равномерном распределении единиц совокупности внутри каждого интервала. При этом чем больше объем изучаемой совокупности, тем указанная неточность будет меньше.

Как уже указывалось, использование некоторых из основных свойств средней арифметической позволяет облегчить ее вычисление. В частности, на свойствах 3, 4 и 5 основан способ расчета средней в интервальных рядах с равными интервалами, называемый «способом моментов» или «способом отсчета от условного нуля». Суть его заключается в следующем. Из всех вариантов вычитается постоянное число  $x_0$  (лучше всего в качестве такового выбрать середину центрального интервала или интервала, обладающего наибольшей частотой, хотя в принципе выбор этого числа произволен). Затем полученные разности следует разделить на другое постоянное число — ширину интервала. Таким образом мы получаем значения преобразованных вариантов  $x'$ , среднее значение из которых называется условным моментом первого порядка:

$$m_1 = \frac{\sum x'f}{\sum f}. \quad (9.12)$$

Читатель, знакомый с математической статистикой, должен иметь представление о начальных, центральных и условных моментах распределения, в общем виде выражающихся как

$$m_k = \frac{\sum (x - A)^k f}{\sum f},$$

где  $k$  — порядок момента,  $A$  — константа, равная соответственно нулю, средней арифметической и любому произвольно выбираемому числу  $x_0$ .



При этом в качестве весов  $f$  целесообразно использовать значения частот, выраженные в долях от общего объема совокупности или в процентах к итогу. Рассмотрим этот расчет применительно к определению среднего стажа рабочих цеха (табл. 9.3).

Момент первого порядка по полученным данным равен:

$$m_1 = \frac{\sum x'f}{\sum f} = -\frac{64,6}{100} = -0,646.$$

Теперь для определения средней арифметической достаточно момент первого порядка умножить на величину интервала  $h$ , на которую раньше были разделены разности  $x - x_0$ , и к получен-

Таблица 9.3. Определение среднего стажа рабочих цеха способом моментов

Группы рабочих по стажу, лет	Среднее значение интервала $x$	$x - x_0$ ( $x_0 = 12,5$ )	$x' = \frac{x - x_0}{h}$ ( $h = 5$ )	$f$ , % к итогу	$x'f$
0—5	2,5	—10	—2	14,6	—29,2
5—10	7,5	—5	—1	49,5	—49,5
10—15	12,5	0	0	23,7	0
15—20	17,5	5	1	10,3	10,3
20 и более	22,5	10	2	1,9	3,8
Всего	—	—	—	100	—64,6

ному произведению прибавить величину ранее вычитаемого из всех вариантов значения  $x_0$ :

$$\bar{x} = m_1 h + x_0. \quad (9.13)$$

В примере  $\bar{x} = -0,646 \cdot 5 + 12,5 = 9,27$  года.

Как мы видим из этого примера, использование способа моментов значительно облегчает работу, что имеет значение при использовании простой (клавишной) вычислительной техники.

4. В статистической практике нередко возникает необходимость определения средней для всей совокупности (общей средней) по данным о средних для отдельных частей этой совокупности (групповых или частных средних). Общая средняя равна средней из групповых средних, взвешенных по численности соответствующих частей общей совокупности, т. е.

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}, \quad (9.14)$$

где  $\bar{x}$  — общая средняя,  $\bar{x}_i$  — групповые средние,  $f_i$  — численности частей совокупности. Довказательство этого правила весьма несложно и предоставляется нами читателю.

Пример. Территория региона площадью 1200 тыс. км<sup>2</sup> разбита на 3 зоны, имеющие площади 450, 390 и 360 тыс. км<sup>2</sup>. Средняя плотность населения этих зон составляет соответственно 16,0; 13,5 и 15,1 чел./км<sup>2</sup>. Определить среднюю плотность населения по району в целом.

В соответствии с (9.14)

$$\bar{x} = \frac{16 \cdot 450 + 13,5 \cdot 390 + 15,1 \cdot 360}{1200} = 15,7 \text{ чел./км}^2.$$

### 9.3. СРЕДНЯЯ ГАРМОНИЧЕСКАЯ, ЕЕ ВЫЧИСЛЕНИЕ И ОБЛАСТЬ ПРИМЕНЕНИЯ

В некоторых случаях характер исходных статистических данных таков, что расчет средней арифметической теряет смысл и единственным обобщающим показателем может служить только другой вид средней — *средняя гармоническая*. Поясним это примерами.

1. Пусть известно, что автомобиль прошел первые 210 км со скоростью 70 км/ч, а оставшиеся 150 км со скоростью 75 км/ч. Определить среднюю скорость автомобиля на протяжении всего пути в 360 км. Можно ли использовать для решения этой задачи среднюю арифметическую?

Очевидно, нельзя, так как если вариантами являются скорости на отдельных участках  $x_1 = 70$  км/ч и  $x_2 = 75$  км/ч, а весами ( $f_i$ ) считать соответствующие отрезки пути, то произведения вариантов на веса не будут иметь ни физического, ни экономического смысла. В данном случае смысл приобретают частные деления отрезков пути на соответствующие скорости (варианты  $x_i$ ), т. е. затраты времени на прохождение отдельных участков пути ( $f_i/x_i$ ). Если отрезки пути обозначить через  $f_h$ , то затраты времени на этих участках выразятся как  $f_h/x$ , а общие затраты времени как  $\sum \frac{f_h}{x}$ . Тогда средняя скорость может быть найдена как частное от деления всего пути ( $\sum f_h$ ) на общие затраты времени:

$$\bar{x}_h = \frac{\sum f_h}{\sum \frac{f_h}{x}}.$$

В нашем примере получим:

$$\bar{x}_h = \frac{210 + 150}{\frac{210}{70} + \frac{150}{75}} = 72 \text{ км/ч}.$$

Двигаясь с этой средней скоростью, автомобиль затратил бы на весь путь (360 км) столько же времени, сколько было затрачено фактически ( $210 : 70 + 150 : 75 = 5$  ч).

2. Известны следующие данные о работе двух рабочих. (Табл. 9.4).

Определим, сколько в среднем затрачивалось времени на одно изделие, т. е. среднюю трудоемкость изделия.

Поскольку средние затраты времени на одно изделие равны общим его затратам, деленным на количество изделий, в данном случае именно так проще всего и найти среднюю:

Таблица 9.4. Затраты времени и выработка продукции двумя рабочими

Табельный номер рабочего	Затраты времени, мин		Выработано изделий, шт.
	всего $f_h$	на одно изделие $x$	$f = \frac{f_h}{x}$
1	1200	12	100
2	300	15	20
Итого	1500	(12,3)	120

$$\bar{x} = 1500 : 120 = 12,5 \text{ мин.}$$

Если бы общие затраты времени не были известны, их можно было бы найти, умножив затраты времени на одно изделие ( $x$ ) на количество изделий (веса, т. е. частоты  $f$ ). Тогда была бы использована формула средней арифметической взвешенной:

$$\bar{x}_a = \frac{\sum x f}{\sum f} = \frac{12 \cdot 100 + 15 \cdot 20}{100 + 20} = 12,5 \text{ мин.}$$

Если же не было бы известно количество изделий ( $f$ ), то его можно было бы найти, разделив общие затраты времени каждым рабочим на трудоемкость изделия (варианты  $x$ ):

$$\bar{x}_h = \frac{\sum f_h}{\sum x} = \frac{1200 + 300}{12 + 15} = 12,5 \text{ мин.}$$

В этом случае для расчета была бы использована формула средней гармонической, где весами были бы общие затраты времени ( $f_h = x f$ ).

В обоих рассмотренных примерах среднюю гармоническую взвешенную приходилось использовать в тех случаях, когда, кроме вариантов осредняемого признака ( $x$ ), были известны показатели, представляющие собой произведения вариантов на веса средней арифметической ( $f$ ). Эти произведения (отрезки пути, пройденные с данной скоростью; общие затраты времени при данной трудоемкости) и выступают в качестве весов средней гармонической ( $f_h = x f$ ).

Если при использовании средней гармонической веса всех вариантов ( $f_h$ ) равны, то вместо взвешенной можно использовать простую (невзвешенную) среднюю гармоническую:

$$\bar{x}_h = \frac{n}{\sum \frac{1}{x}},$$

где  $n$  — число вариантов осредняемого признака.

Так, во 2-м примере простую среднюю гармоническую можно было бы использовать, если бы общие затраты времени каждым рабочим были одинаковы. В примере же со скоростью простую среднюю гармоническую можно было бы применить, если бы были равны отрезки пути, пройденные с разной скоростью.

#### 9.4. ОПРЕДЕЛЯЮЩИЙ ПОКАЗАТЕЛЬ, УРАВНЕНИЕ СРЕДНЕЙ И ЕЕ ФОРМА

Любая средняя величина должна вычисляться так, чтобы при замене ее каждого варианта осредняемого признака не изменялась величина некоторого итогового, обобщающего показателя, который связан с осредняемым показателем. Так, при замене фактических скоростей на отдельных отрезках пути их средней величиной (средней скоростью) не должно измениться общее расстояние. Или если фактические затраты времени на одно изделие отдельными рабочими заменить средними затратами (средней трудоемкостью изделия), то не должны измениться общие затраты времени.

Формула (формула) средней определяется характером (механизмом) взаимосвязи этого итогового показателя с осредняемым. Поэтому итоговый показатель, величина которого не должна изменяться при замене вариантов их средней величиной, называется *определяющим показателем*.

Для вывода формулы средней нужно составить и решить уравнение, используя *взаимосвязь* осредняемого показателя с определяющим. Это уравнение строится путем замены вариантов осредняемого показателя (признака) их средней величиной.

Так, если известны варианты скорости ( $x_i$ ) и время ( $f_i$ ), уравнение средней будет таким:  $x_1 f_1 + x_2 f_2 + \dots + x_n f_n = \bar{x} f_1 + \bar{x} f_2 + \dots + \bar{x} f_n$  или  $\sum x_i f_i = \sum \bar{x} f_i = \bar{x} \sum f_i$ . Здесь  $x_i f_i$  — расстояния, пройденные с данной скоростью. Решая уравнение относительно  $\bar{x}$ , получим формулу средней арифметической взвешенной:  $\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$  или  $\bar{x} = \frac{\sum x f}{\sum f}$ .

Если же непосредственно заданы скорости ( $x$ ) и отрезки пути ( $x f = f_h$ ), то уравнение средней скорости можно построить так:

$$f_{h1} + f_{h2} + \dots + f_{hn} = x_1 \frac{f_{h1}}{x_1} + x_2 \frac{f_{h2}}{x_2} + \dots + x_n \frac{f_{hn}}{x_n},$$

где  $\frac{f_h}{x} = f$  — время. Заменив в правой части этого равенства множители  $x_i$  на  $\bar{x}$  и вынеся  $\bar{x}$  за знак  $\sum$ , получим:  $\sum f_h = \bar{x} \sum \frac{f_h}{x}$ ,

откуда  $\bar{x} = \frac{\sum f_h}{\sum \frac{f_h}{x}}$  (средняя гармоническая взвешенная).

Во многих случаях осредняемый показатель представляет собой отношение двух других показателей, например:

$$\text{скорость} = \frac{\text{путь}}{\text{время}};$$

$$\text{трудоемкость единицы продукции} = \frac{\text{общие затраты времени}}{\text{количество продукции}} \text{ и т. п.}$$

Определяющим показателем здесь является числитель дроби и может быть выражен либо как  $\sum x f$ , либо как  $\sum f_h$  (см. выше). Из этого вытекает следующее практическое правило: если кроме

Таблица 9.5. Затраты времени на выработку изделий двумя рабочими

Табельный номер рабочего	Общие затраты времени, ч	Выработано изделий, шт.	
		всего $f_h$	за 1 ч $x$
1	20	100	5
2	5	20	4
Итого	25	120	(4,8)

В приведенном выше примере с двумя рабочими выразим общие затраты времени в часах и найдем выработку изделий за 1 ч каждым рабочим и среднюю выработку за 1 ч (табл. 9.5). Исходное отношение здесь таково:

$$\text{Выработка изделий за 1 ч} = \frac{\text{Общая выработка изделий}}{\text{Общие затраты времени, ч.}}$$

Определяющим показателем является общая выработка изделий ( $\sum f_h = \sum x f$ ).

Если непосредственно заданы общая выработка изделий и общие затраты времени, то для расчета средней выработки за 1 ч можно использовать исходное отношение:  $\bar{x} = 120 : 25 = 4,8$  шт. Если заданы  $x$  и  $f$  (см. табл. 9.5), т. е. в качестве весов будут использованы показатели, находящиеся в знаменателе исходной дроби, то

$$\bar{x} = \frac{\sum x f}{\sum f} = \frac{5 \cdot 20 + 4 \cdot 5}{25} = 4,8 \text{ шт.}$$

Если же заданы  $x$  и  $f_h$ , то

$$\bar{x} = \frac{\sum f_h}{\sum \frac{f_h}{x}} = \frac{100 + 20}{\frac{100}{5} + \frac{20}{4}} = \frac{120}{25} = 4,8 \text{ шт.}$$

Этот результат полностью согласуется с вычисленными выше средними затратами времени на одно изделие: если на одно изделие затрачивается в среднем 12,5 мин., то за 1 ч в среднем будет сделано 4,8 изделия ( $60 : 12,5 = 4,8$ ).

При использовании взвешенных средних нужно иметь в виду, что понятия *вес* и *частота* в общем случае не совпадают, не являются тождественными. Если, например, дана группировка предприятий по проценту выполнения плана — вариационный ряд, то частотами в этом ряду будут числа предприятий в каждой группе. Однако при вычислении среднего процента выполнения плана, т. е. процента выполнения плана в среднем (в целом) по всем предприятиям, нужно исходить из того, что степень выполнения плана =  $\frac{\text{фактический уровень}}{\text{плановый уровень}}$ . Следовательно,

если по каждой группе известны плановые уровни, то, используя их в качестве весов, нужно применить среднюю арифметическую взвешенную, если же известны и будут использоваться в качестве весов фактические уровни, то расчет следует вести по формуле средней гармонической. В обоих случаях использование в качестве весов частот вариационного ряда привело бы к неверным результатам.

Кроме средней арифметической и средней гармонической, в статистике используются и другие виды (формы) средней. Все они являются частными случаями *степенной средней*:

$$\bar{x} = \sqrt[m]{\frac{\sum x^m f}{\sum f}}; \quad \bar{x} = \sqrt[m]{\frac{\sum x^m}{n}}, \quad (9.14a, б)$$

где  $x$  — варианты осредняемого признака,  $f$  — веса (в частности, частоты),  $m$  — показатель степени, определяющий вид средней.

При  $m=1$  формула 9.14 дает среднюю арифметическую; при  $m=-1$  получим среднюю гармоническую  $\bar{x}_h = \frac{\sum f_h}{\sum \frac{f_h}{x}}$ ;

$\bar{x}_h = \frac{n}{\sum 1/x}$ . Если  $m=0$ , то при раскрытии неопределенности вида  $\frac{0}{0}$  с помощью правила Лопиталля получим *среднюю геометрическую*:  $\bar{x}_g = \sqrt[n]{\prod x}$ , или  $\bar{x}_g = \sqrt[n]{\frac{\sum x^i f}{\sum f}}$ , где  $\prod$  — символ произведения. Использование средней геометрической в статистике будет показано в гл. 14.

При  $m=2$  будем иметь среднюю квадратическую:

$$\bar{x}_q = \sqrt{\frac{\sum x^2 f}{\sum f}} \quad \text{или} \quad \bar{x}_q = \sqrt{\frac{\sum x^2}{n}}. \quad (9.15 а, б)$$

Эта средняя используется при измерении вариации признака (см. гл. 10).



При использовании одних и тех же исходных данных чем больше  $m$  в формуле 9.14, тем больше средняя величина:  $\bar{x}_h < \bar{x}_g < \bar{x}_a < \bar{x}_g$  (мажорантность средних).

#### 9.5. МЕДИАНА, МОДА И ДРУГИЕ ОПИСАТЕЛЬНЫЕ СРЕДНИЕ

Средние величины, описанные выше, дают обобщенное представление об изучаемой совокупности и с этой точки зрения их теоретическое, прикладное и познавательное значение бесспорно. Но бывает, что величина средней не совпадает ни с одним из реально существующих вариантов. Так, например, в одном из приведенных выше примеров полученная средняя грузоподъемность одного подъемного крана составляет 14 т, хотя значения вариантов равны 40, 25, 10 и 5 т. В другом примере вычисленный средний квалификационный разряд оказался равным 3,52, хотя разряд отдельного рабочего не может быть дробным числом. Поэтому, кроме рассмотренных средних, в статистическом анализе целесообразно использовать величины конкретных вариантов, занимающие в упорядоченном (ранжированном) ряду значений признака вполне определенное положение. Среди таких величин наиболее употребительными являются медиана ( $Me$ ) и мода ( $Mo$ ), называемые иногда описательными или структурными средними.

**Медианой** называется вариант, расположенный в центре ранжированного ряда. Если, например, расположить по росту одиннадцать человек, то медианой в таком ряду будет являться значение роста шестого по счету (от начала или конца) человека. Медиана делит ряд на две равные части таким образом, что по обе стороны от нее находится одинаковое количество единиц совокупности. При этом у одной половины единиц совокупности значение варьирующего признака меньше медианы, у другой — больше ее.

Описательный характер медианы проявляется в том, что она характеризует количественную границу значений варьирующего признака, которыми обладает половина единиц совокупности. Так, если мы говорим, что медианное значение возраста вступления в брак мужчин равно 23 годам, это означает, что одна половина всех вступающих в брак мужчин моложе 23 лет, а вторая — старше.

Значительный интерес представляет нахождение медианы в вариационных рядах. Применительно к дискретным рядам эта задача решается сравнительно просто. Если всем единицам ряда придать порядковые номера, то порядковый номер медианного варианта определяется как  $(n+1)/2$  для ряда с нечетным числом членов  $n$ . Если же количество членов ряда является четным числом, то медианой будет являться среднее значение двух вариантов, имеющих порядковые номера  $n/2$  и  $n/2+1$ . Так, в ряду, состоящем из 200 единиц совокупности, медиана будет

определена как среднее значение величин вариантов, имеющих порядковые номера 100 и 101.

Однако при достаточно большом количестве единиц совокупности разница в значениях величин двух смежных вариантов невелика и один из центральных вариантов с порядковым номером  $n/2$  может считаться медианой.

Нахождение медианы в интервальных вариационных рядах требует предварительного определения интервала, в котором она находится (медианного интервала). Этот интервал характерен тем, что его кумулятивная частота (накопленная сумма частот) равна или превышает полусумму всех частот ряда. После этого значение медианы вычисляется путем линейной интерполяции по формуле

$$Me = x_{Me} + h_{Me} \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}}, \quad (9.15)$$

где  $x_{Me}$  — нижняя граница медианного интервала;  $h_{Me}$  — ширина медианного интервала;  $S_{Me-1}$  — сумма частот, накопленная до медианного интервала;  $f_{Me}$  — частота медианного интервала.

Формула (9.15) получена исходя из допущения о равномерном нарастании накопленной частоты внутри интервала и пригодна для любого интервального ряда.

Рассмотрим расчет медианы в интервальном вариационном ряду, характеризующем распределение рабочих по стажу работы (табл. 9.6).

Прежде всего найдем медианный интервал. Таким интервалом, очевидно, будет интервал стажа 5—10 лет, поскольку его кумулятивная частота равна 641, что превышает половину суммы всех частот ( $1000 : 2 = 500$ ). Нижняя граница этого интервала — 10 лет, его частота — 495, частота, накопленная до него, — 146. Подставив эти данные в (9.15), получим:

$$Me = 5 + 5 \cdot \frac{1000 - 146}{495} = 8,6 \text{ года}.$$

Полученный результат говорит о том, что из 1000 рабочих цеха 500 человек имеют стаж работы меньше 8,6 года, а 500 — больше.

Таблица 9.6. К расчету медианного стажа рабочих цеха

Группы рабочих по стажу работы, лет	Количество рабочих (частоты)	Кумулятивные частоты $S$
0—5	146	146
5—10	495	641
10—15	237	878
15—20	103	981
20 и более	19	1000
Всего	1000	—

Наряду с медианой для более полной характеристики структуры изучаемой совокупности применяют и другие значения вариантов, занимающих в ранжированном ряду вполне определенное положение. К ним относятся *квартили* и *децили*. Квартили делят ряд по сумме частот на 4 равные части, а децили — на 10 равных частей. Квартилей насчитывается три, а децилей — девять.

Расчет этих показателей в вариационном ряду аналогичен расчету медианы. Он начинается с нахождения порядкового номера соответствующего варианта и определения по накопленным частотам того интервала, в котором этот вариант находится. Величина искомого варианта внутри интервала определяется путем линейной интерполяции. Формулы для квартилей в интервальном вариационном ряду имеют вид:

$$Q_1 = x_{01} + h_{Q_1} \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}};$$

верхний (или третий) квартиль

$$Q_3 = x_{03} + h_{Q_3} \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}}.$$

В этих формулах  $x_{01}$  и  $x_{03}$  — нижние границы соответствующих квартильных интервалов,  $h_{Q_1}$  — величина соответствующего интервала,  $\sum f$  — сумма частот ряда,  $S_{Q_1-1}$  и  $S_{Q_3-1}$  — накопленные частоты интервалов, предшествующих соответствующим квартильным,  $f_{Q_1}$  и  $f_{Q_3}$  — частоты соответствующих квартильных интервалов. Вторым квартилем является, естественно, медиана. Аналогично могут быть построены формулы децилей.

Рассмотрим на нашем примере (см. табл. 9.6) вычисление квартилей и децилей. Найдем нижний (первый) квартиль. Для этого определим  $\frac{1}{4}$  суммы всех частот. Она составляет 250. Из табл. 9.6 видно, что 250-й вариант находится в интервале 5—10 лет. Следовательно,  $x_{01}=5$ . Сумма накопленных до этого интервала частот  $S_{Q_1-1}=146$ , частота этого интервала — 495. Отсюда

$$Q_1 = 5 + 5 \cdot \frac{\frac{1000}{4} - 146}{495} = 6,1 \text{ года.}$$

Это означает, что четвертая часть рабочих цеха имеют стаж работы меньше, чем 6,1 года, либо что три четверти их имеют стаж, превышающий 6,1 года.

Вычислим верхний (третий) квартиль. Три четверти рабочих цеха составляют 750 человек. Следовательно, искомая величина находится в интервале 10—15 лет. Частота этого интервала  $f_{Q_3}=237$ , частота накопленная до него,  $S_{Q_3-1}=641$ . Следовательно:

$$Q_3 = 10 + 5 \cdot \frac{750 - 641}{237} = 12,3 \text{ года.}$$

Это означает, что три четверти рабочих имеют стаж работы менее 12,3 года, а одна четверть — стаж, превышающий 12,3 года.

*Модой* в статистике называется величина признака, которая чаще всего встречается в данной совокупности. Применительно к вариационному ряду модой является вариант, обладающий наибольшей частотой. Этот показатель широко используется в тех случаях, когда требуется охарактеризовать наиболее часто встречающуюся величину признака (наиболее распространенный размер обуви, головных уборов или одежды, преобладающие цены, спортивные результаты и т. п.).

Нахождение моды в дискретных вариационных рядах несложно. Рассматривая, например распределение рабочих по квалификационному разряду (см. табл. 9.1), мы сразу видим, что наиболее распространенным в данной совокупности рабочих является 4 разряд, так как он обладает наибольшей частотой (78). Следовательно, модальное значение изучаемого признака — 4.

Расчет моды в интервальных вариационных рядах производится следующим образом. Вначале отыскивается модальный интервал, т. е. интервал, обладающий наибольшей частотой. Затем интерполяцией между его границами определяется значение моды. Для интерполяции используется допущение о том, что в модальном и двух смежных с ним интервалах кривая распределения представляет собой квадратичную параболу, абсцисса вершины которой и является модой. Интерполяционная формула для рядов с равными интервалами имеет вид

$$Mo = x_{Mo} + h \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}, \quad (9.16)$$

где  $x_{Mo}$  — нижняя граница модального интервала,  $h$  — ширина интервала,  $f_{Mo}$  — частота модального интервала,  $f_{Mo-1}$  и  $f_{Mo+1}$  — соответственно частоты предмодального и послемодального интервалов.

Определим моду для рассматриваемого распределения рабочих цеха по стажу работы (см. табл. 9.6). Модальным интервалом в этом ряду будет интервал 5—10 лет, так как он имеет наибольшую частоту ( $f_{Mo}=495$ ). Частоты смежных с ним интервалов — предшествующего и последующего — равны соответственно  $f_{Mo-1}=146$  и  $f_{Mo+1}=237$ .

Подставляя эти данные в (9.16), имеем:

$$Mo = 5 + 5 \cdot \frac{495 - 146}{(495 - 146) + (495 - 237)} = 7,9 \text{ года.}$$

Следовательно, наиболее распространенным стажем является 7,9 года, т. е. около 8 лет.

В рассмотренном ряду распределения рабочих цеха по стажу средняя арифметическая равна 9,3, медиана — 8,6 и мода — 7,9 года.

Медиана и мода, в отличие от средней арифметической, не показывают индивидуальных различий в значениях варьирующего признака и поэтому являются дополнительными и очень важными характеристиками статистической совокупности. На прак-

тике они часто используются вместо средней либо наряду с ней. Особенно целесообразно вычислять медиану и моду в тех случаях, когда изучаемая совокупность содержит некоторое количество единиц с очень большим или очень малым значением варьирующего признака. Эти, не очень характерные для совокупности значения вариантов, влияя на величину средней арифметической, не влияют на значения медианы и моды, что делает последние очень ценными для экономико-статистического анализа показателями. Совместное использование средней арифметической, медианы и моды проводится при анализе рядов распределения. Соотношение этих трех показателей позволит достаточно исчерпывающим образом охарактеризовать распределение, о чем подробно будет сказано в гл. 10.

#### 9.6. ПРИНЦИПЫ ИСПОЛЬЗОВАНИЯ СРЕДНИХ СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Средние величины, являясь обобщенной характеристикой совокупности, могут иметь большую научную и познавательную ценность только в том случае, если совокупность состоит из единиц одного и того же рода и является достаточно многочисленной. Основное условие научного использования средних величин в статистике заключается в том, что они, должны вычисляться на основе массового обобщения фактов и применяться к качественно однородным совокупностям. Основоположники марксизма-ленинизма четко сформулировали основы подлинно научного подхода к вычислению и использованию средних статистических показателей. Прекрасно представляя себе природу и свойства этих показателей, они также дали примеры правильного применения их для характеристики изучаемых социальных явлений и процессов. Имея в виду, что средняя, являясь обобщенной характеристикой совокупности в целом, не подменяет конкретных индивидуальных величин, К. Маркс писал: «Отклонения уравниваются для всего общества, но не для отдельного хозяйства»<sup>1</sup>.

В. И. Ленин, широко использовавший статистические методы в своих работах, предостерегал от применения средних величин к качественно неоднородным совокупностям. «Соединение крупных и мелких заведений вместе, — писал он, — дает совершенно фиктивные «средние» цифры, не дающие никакого понятия о действительности, затупевающие кардинальные различия, изображающие однородное нечто совершенно разнородное, разностовное»<sup>2</sup>. В. И. Ленин резко критиковал результаты работ статистиков-пародников, которые, пренебрегая качественными различиями между крупными промышленными предприятиями и мелкими кустарными заведениями, получили

разнородную в статистическом отношении совокупность и характеризовали ее средними показателями. К таким же фиктивным средним они прибегали, анализируя процессы, происходящие в русской деревне. Они рассматривали крестьян как некую однородную массу, игнорируя классовую дифференциацию крестьянских хозяйств и тем самым развитие капиталистических отношений в сельском хозяйстве.

К аналогичным методам прибегают буржуазные статистики и сегодня. Не обращая внимания на принципиальную разницу в законах развития доходов рабочих и капиталистов, они огульным образом усредняют их и получают неверные, но зато пригодные для своих целей показатели развития капиталистического общества.

Советская статистика, основываясь на идейном и теоретическом наследии классиков марксизма-ленинизма, считает метод средних величин неотделимым от метода группировок. Метод группировок дает возможность установить четкие границы качественно однородных совокупностей и тем самым получить научно обоснованные значения средних показателей. Кроме того, он позволяет дополнить общую среднюю групповыми средними, которые делают статистико-экономический анализ глубже и полнее, а его результаты — нагляднее. Групповые средние имеют и самостоятельную ценность, являясь опорными данными для анализа взаимосвязей (см. гл. 12 и 13).

И наконец, советская статистика практикует совместное использование средних показателей и индивидуальных величин. Поскольку средние основываются на массовом обобщении фактов, т. е. на игнорировании значений отдельных единиц совокупности, чрезмерное увлечение ими может принести вред. Иногда за общими, оптимистично выглядящими средними показателями могут скрываться индивидуальные, не находящиеся на должном уровне. Так, например, объективный анализ деятельности отдельной отрасли промышленности предполагают использование не только средних показателей (выполнения плана, экономии электроэнергии, себестоимости продукции, снижения простоев и т. д.), но и индивидуальных показателей, относящихся к передовым и отстающим предприятиям.

Только соблюдение указанных принципов научного использования средних величин может сделать их подлинно объективными характеристиками явлений и процессов общественной жизни.

<sup>1</sup> Маркс К., Энгельс Ф. Соч., т. 23, с. 335.

<sup>2</sup> Ленин В. И. Полн. собр. соч., т. 2, с. 374.



# ГЛАВА 10 ПОКАЗАТЕЛИ ВАРИАЦИИ И АНАЛИЗ ВАРИАЦИОННЫХ РЯДОВ

## 10.1. ОБЩЕЕ ПОНЯТИЕ О ПОКАЗАТЕЛЯХ ВАРИАЦИИ, ИХ НАЗНАЧЕНИЕ И ВИДЫ

Как уже говорилось в предыдущей главе, средняя величина, являясь обобщающим показателем для всех единиц статистической совокупности, не дает представления об индивидуальных значениях варьирующего признака и о различиях между ними. В то же время именно эти различия часто представляют большой интерес для исследователя, так как они позволяют

Таблица 10.1. Распределение зданий по количеству этажей в двух городских районах А и Б

Количество этажей в здании $x$	Количество зданий $f$		Произведения вариантов на частоты $xf$		Отклонения вариантов от средней $x-\bar{x}$		Произведения модуля отклонения на частоту $ x-\bar{x}  \cdot f$	
	А	Б	А	Б	А	Б	А	Б
1	3	9	3	9	-4,23	-4,23	12,69	38,07
2	7	0	14	0	-3,23	-3,23	22,61	0,00
3	14	1	42	3	-2,23	-2,23	31,22	2,23
4	28	4	112	16	-1,23	-1,23	34,44	4,92
5	73	160	365	800	-0,23	-0,23	16,79	36,80
6	46	76	276	456	0,77	0,77	35,42	58,52
7	25	0	175	0	1,77	1,77	44,25	0,00
8	11	3	88	24	2,77	2,77	30,47	8,31
9	2	4	18	36	3,77	3,77	7,54	15,08
Всего	209	257	1093	1344	—	—	235,43	163,93

полнее раскрыть строение изучаемой совокупности и получить дополнительный материал для экономико-статистического анализа.

Недостаточность средней величины для исчерпывающей характеристики совокупности можно проиллюстрировать следующим примером. Рассмотрим распределение зданий по количеству этажей в двух районах города (табл. 10.1). Районы А и Б отличаются друг от друга как общей численностью зданий, так и характером их распределения по количеству этажей. В районе А, например доля пятиэтажных зданий составляет 34,9 %, а в районе Б — 62,3 %. При этом в районе А имеются все виды зданий — от одноэтажных до девятиэтажных, в то время как в районе Б отсутствуют двух- и семизэтажные здания. В то же время среднее количество этажей в одном здании в этих районах одинаково:  $\bar{x}_A = 1093 : 209 = 5,23$ ;  $\bar{x}_B = 1344 : 257 = 5,23$ .

Следовательно, в данном случае средняя величина (средняя этажность) не дает представления об особенностях застройки каждого района. Более того, средние этажности районов А и Б обладают разной степенью типичности, надежности. В этом легко убедиться, сравнив доли пятиэтажных зданий (вариантов, наиболее близких к средней величине) в этих районах. По-видимому, для района Б средняя этажность 5,23 является более характерным показателем, чем для района А. Отсюда вытекает необходимость дополнять средние величины показателями, позволяющими оценить типичность этих средних путем измерения колеблемости (вариации) изучаемого признака. Использование этих показателей вариации дает возможность сделать статистический анализ более полным и содержательным и тем самым глубже понять сущность изучаемых общественных явлений.

Простейшим из показателей вариации является *размах вариации*. Он представляет собой разность между наибольшим и наименьшим значениями варьирующего признака:

$$R = x_{\max} - x_{\min}. \quad (10.1)$$

Этот показатель дает лишь самое общее представление о колеблемости изучаемого признака, ибо показывает разницу только между предельными значениями вариантов. Он совершенно не связан с частотами в вариационном ряду, т. е. с характером распределения, а его зависимость только от крайних значений признака может придавать ему неустойчивый, случайный характер. В приведенном выше примере эти показатели для обоих районов равны ( $R_A = 9 - 1 = 8$  и  $R_B = 9 - 1 = 8$ ), следовательно, они не дают нам никакой информации об особенностях исследуемых совокупностей и не позволяют оценить степень типичности полученных средних.

Из сказанного напрашивается вывод о том, что для характеристики вариации признака нужно знать не только амплитуду (размах) его значений, но и уметь обобщить отклонения всех этих значений от какой-либо типичной для изучаемой совокупности величины. Естественно, в качестве последней уместно использовать среднюю арифметическую. Тогда характеристику вариации даст показатель, именуемый *средним линейным отклонением* и представляющий собой среднюю арифметическую из абсолютных значений отклонений отдельных вариантов от их средней арифметической:

$$Л = \frac{\sum |x - \bar{x}|}{n}; \quad (10.2a)$$

$$Л = \frac{|x - \bar{x}|}{\sum f}, \quad (10.2b)$$

где  $L$  — среднее линейное отклонение,  $|x - \bar{x}|$  — абсолютное значение (модуль) отклонения варианта от средней арифметической,  $f$  — частота.

Первая формула применяется, если каждый из вариантов встречается в совокупности только один раз, а вторая — в рядах с неравными частотами.

Необходимость использования в (10.2а, б) модулей отклонений вариантов от средней вызвана тем, что алгебраическая сумма этих отклонений равна нулю (см. свойства средней арифметической).

Рассмотрим расчет среднего линейного отклонения в распределениях зданий по количеству этажей (см. табл. 10.1):

$$L_A = \frac{235,43}{209} = 1,13; \quad L_B = 163,93 : 257 = 0,64 \text{ этажа.}$$

Результаты расчета показывают, что вариация зданий по количеству этажей в районе Б значительно ниже, чем в районе А, т. е. район Б представляет собой более однородную совокупность зданий, чем район А. Отсюда, кроме того, вытекает, что средняя этажность одного здания 5,23 более типична для района Б, чем для района А.

Существует и другой способ усреднения отклонений вариантов от средней арифметической, позволяющий обойти трудность, обусловленную равенством нулю их алгебраической суммы. Этот очень распространенный в статистике способ сводится к расчету квадратов отклонений вариантов от средней с их последующим усреднением. При этом мы получаем новый показатель вариации — *дисперсию*, представляющую собой средний квадрат отклонений вариантов от их средней арифметической

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}; \quad (10.3а)$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}. \quad (10.3б)$$

Вторая формула применяется при наличии у вариантов своих весов (или частот вариационного ряда).

В экономико-статистическом анализе вариацию признака принято оценивать чаще всего с помощью *среднего квадратического отклонения*, представляющего собой корень квадратный из дисперсии:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}. \quad (10.4)$$

Среднее квадратическое отклонение, как и среднее линейное отклонение, характеризует абсолютную колеблемость значений варьирующего признака и выражается в тех же единицах измерения, что и варианты.

Рассмотрим вычисление дисперсии и среднего квадратического отклонения для распределения зданий по количеству этажей в двух городских районах А и Б (табл. 10.2).

Таблица 10.2. К расчету дисперсии и среднего квадратического отклонения

Количество этажей в здании $x$	Количество зданий $f$		Произведение вариантов на частоты $xf$		Отклонения вариантов от средней $x - \bar{x}$		Квадраты отклонений $(x - \bar{x})^2$		Произведение квадратов отклонений на частоты $(x - \bar{x})^2 f$	
	А	Б	А	Б	А	Б	А	Б	А	Б
1	3	9	3	9	-4,23	-4,23	17,89	17,89	53,67	161,01
2	7	0	14	0	-3,23	-3,23	10,43	10,43	108,78	0
3	14	1	42	3	-2,23	-2,23	4,97	4,97	69,58	4,97
4	28	4	112	16	-1,23	-1,23	1,51	1,51	42,28	6,04
5	73	160	365	800	-0,23	-0,23	0,053	0,053	3,87	8,48
6	46	76	276	456	0,77	0,77	0,593	0,593	27,28	45,07
7	25	0	175	0	1,77	1,77	3,13	3,13	78,25	0
8	11	3	88	24	2,77	2,77	7,67	7,67	84,37	23,01
9	2	4	18	36	3,77	3,77	14,21	14,21	28,42	56,84
Всего	209	257	1093	1344	—	—	—	—	496,50	305,45

Подставив в (10.3б) соответствующие данные из табл. 10.2, получим значение дисперсии этажности для районов А и Б:

$$\sigma_A^2 = \frac{496,5}{209} = 2,36; \quad \sigma_B^2 = \frac{305,45}{257} = 1,19.$$

Средние квадратические отклонения составят:

$$\sigma_A = \sqrt{2,36} = 1,54 \text{ этажа}; \quad \sigma_B = \sqrt{1,19} = 1,09 \text{ этажа.}$$

Таким образом, количество этажей в зданиях отклоняется от среднего значения (5,23) на 1,54 в районе А и на 1,09 в районе Б. Этот результат подтверждает сделанный ранее вывод о большей однородности застройки района Б по сравнению с районом А.

В статистической практике часто возникает необходимость сравнения вариации различных признаков. Например, большой интерес представляет сравнение вариаций возраста рабочих и их квалификации, стажа работы и размера заработной платы, количества вносимых в почву удобрений и урожайности и т. д. Для подобных сопоставлений показатели абсолютной колеблемости признаков — среднее линейное и среднее квадратическое отклонение, конечно, не пригодны. Нельзя, в самом деле, сравнивать колеблемость стажа работы, выражаемую в годах, с ко-

леблемостью заработной платы, выражаемой в рублях и копейках.

Для осуществления такого рода сравнений статистика использует относительный показатель вариации — коэффициент вариации, представляющий собой выраженный в процентах (или в виде доли) отношение среднего квадратического отклонения к средней арифметической:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 (\%). \quad (10.5)$$

Логика подсказывает нам, что такой показатель вполне может использоваться для сравнения колеблемости разнородных признаков. Применительно к рассмотренному выше примеру коэффициенты вариации этажности районов А и Б составляют

$$V_A = \frac{1,54}{5,23} \cdot 100 = 29,4 \%; \quad V_B = \frac{1,09}{5,23} \cdot 100 = 20,8 \%.$$

Наряду с этим коэффициент вариации часто используется для сравнения размеров вариации в совокупностях, отличающихся друг от друга величиной средней (в совокупностях с разными уровнями).

## 10.2. СПОСОБЫ РАСЧЕТА ДИСПЕРСИИ

Вычисление дисперсии по формулам (10.3) является довольно трудоемкой процедурой. Для облегчения расчетных работ часто используются упрощенные способы определения дисперсии (и среднего квадратического отклонения).

Преобразуем числитель выражения (10.3) следующим образом:

$$\sum (x - \bar{x})^2 = \sum (x^2 - 2x\bar{x} + \bar{x}^2) = \sum x^2 - 2\sum x\bar{x} + \sum \bar{x}^2.$$

Разделив обе части этого равенства на  $n$  и имея в виду, что  $\sum x = n\bar{x}$ , получим:

$$\sigma^2 = \frac{\sum x^2}{n} - 2\bar{x}\bar{x} + \bar{x}^2 = \frac{\sum x^2}{n} - \bar{x}^2. \quad (10.6)$$

Следовательно, дисперсия может быть определена как разность среднего квадрата вариантов и квадрата их средней. Обратимся вновь к данным табл. 10.2 и определим дисперсию этажности района Б по формуле (10.6) (табл. 10.3).

$$\frac{\sum x^2}{n} = \bar{x}^2 = \frac{7334}{257} = 28,54; \quad \bar{x}^2 = 5,23^2 = 27,35;$$

$$\sigma^2 = 28,54 - 27,35 = 1,19,$$

что полностью согласуется с расчетом по формуле (10.3).

В вариационных рядах с равными интервалами дисперсия может быть сравнительно легко вычислена способом моментов или способом отсчета от условного нуля. Расчет производится по формуле

$$\sigma^2 = h^2 [(\bar{x}')^2 - (\bar{x}')^2], \quad (10.7)$$

где  $h$  — ширина интервала,  $x' = \frac{x - x_0}{h}$  — преобразование значения вариантов ( $x_0$  — условный нуль, в качестве которого удобно использовать середину интервала, обладающего наибольшей частотой),  $(\bar{x}')^2 = \frac{\sum x'^2 f}{\sum f}$  — так называемый момент второго порядка,  $(\bar{x}')^2 = \left(\frac{\sum x' f}{\sum f}\right)^2$  — квадрат момента первого порядка.

Рассмотрим использование этого способа на примере распределения электродвигателей по мощности (табл. 10.4).

По данным табл. 10.4 имеем:

Таблица 10.3. К расчету дисперсии по формуле [10.6].

$$(\bar{x}')^2 = \frac{119}{100} = 1,19;$$

$$(\bar{x}')^2 = \left(\frac{-37}{100}\right)^2 = 0,137.$$

Отсюда

$$\sigma^2 = 20^2 \cdot (1,19 - 0,137) = 421,2,$$

$$\sigma = \sqrt{421,2} = 20,52 \text{ кВт.}$$

Совершенно очевидно, что расчет дисперсии и среднего квадратического отклонения по формуле (10.7) менее трудоемок.

Особый интерес представляет нахождение дисперсии альтернативного признака, т. е. признака, которым единицы изучаемой совокупности могут либо обладать, либо не обладать. В таких случаях наличие признака обозначается единицей, а его отсутствие — нулем. Доля единиц, обладающих интересующим нас признаком, обозначается через  $p$ , доля остальных единиц, очевидно, составит  $q = 1 - p$ . Определим для этих условий среднюю величину и дисперсию.

Средняя

$$\bar{x}_p = \frac{\sum x f}{\sum f} = \frac{1 \cdot p + 0 \cdot q}{p + q} = p. \quad (10.8)$$



Дисперсия альтернативного признака:

$$\sigma_p^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1-p)^2 p + (0-p)^2 q}{p+q} = q^2 p + p^2 q = pq(q+p) = pq = p(1-p). \quad (10.9)$$

Среднее квадратическое отклонение при подобной альтернативной вариации составит:

$$\sigma_p = \sqrt{\sigma_p^2} = \sqrt{p(1-p)} = \sqrt{pq}. \quad (10.10)$$

Таблица 10.4. К расчету дисперсии и среднего квадратического отклонения

Мощность электро- двигателя, кВт	Количество электродвигателей f, % к итогу	$x'$	$x'f$	$(x')^2$	$(x')^2 f$
		$x - \bar{x}$			
10-30	2	-3	-6	9	18
30-50	8	-2	-16	4	32
50-70	35	-1	-35	1	35
70-90	41	0	0	0	0
90-110	9	1	9	1	9
110-130	4	2	8	4	16
130-150	1	3	3	9	9
Всего	100	—	-37	—	119

Если, например, 73 % студентов занимаются на «хорошо» и «отлично», то 27 % студентов имеют оценки «удовлетворительно» и ниже. Для этого случая дисперсия доли будет равна:

$$\sigma_p^2 = 0,73 \cdot 0,27 = 0,197.$$

Среднее квадратическое отклонение для доли составит:

$$\sigma_p = \sqrt{0,197} = 0,444, \text{ или } 44,4 \text{ \%}.$$

### 10.3. ВНУТРИГРУППОВАЯ И МЕЖГРУППОВАЯ ВАРИАЦИЯ

Если статистическая совокупность разбита на группы по какому-либо признаку, то средняя величина и дисперсия признака могут быть определены не только для всей совокупности, но и для каждой из составляющих ее групп. Групповые средние и дисперсии обозначим соответственно  $\bar{x}_i$  и  $\sigma_i^2$ . Наряду с групповыми дисперсиями может быть найдено и среднее их значение

$$\bar{\sigma}_i^2 = \frac{\sum \sigma_i^2 f_i}{\sum f_i}. \quad (10.11)$$

Кроме того, можно вычислить и межгрупповую дисперсию, характеризующую колеблемость групповых средних ( $\bar{x}_i$ ) около общей средней ( $\bar{x}$ ):

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 f_i}{\sum f_i}. \quad (10.12)$$

В курсе математической статистики доказывается, что между общей дисперсией  $\sigma^2$ , средней из групповых дисперсий  $\bar{\sigma}_i^2$  и межгрупповой дисперсией  $\delta^2$  существует такая связь:

$$\sigma^2 = \bar{\sigma}_i^2 + \delta^2, \quad (10.13)$$

т. е. общая дисперсия равна сумме средней из групповых дисперсий и межгрупповой дисперсии. Это правило сложения дисперсий.

Рассмотрим вычисление этих дисперсий и покажем справедливость соотношения (10.13) на следующем примере. Пусть ко-

Таблица 10.5. К расчету общей и групповой дисперсий

Токари 5-го разряда			Токари 6-го разряда		
№ п/п	Количество деталей, шт. $x_1$	$x_1^2$	№ п/п	Количество деталей, шт. $x_2$	$x_2^2$
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		63	675

личество изготавливаемых за смену одинаковых деталей токарями 5-го и 6-го разрядов характеризуется такими данными (табл. 10.5).

Отсюда групповые средние:

$$\bar{x}_1 = 36 : 4 = 9; \quad \bar{x}_2 = 63 : 6 = 10,5.$$

Общая средняя:  $\bar{x} = (36 + 63) : 10 = 9,9$ .

Групповые дисперсии:

$$\sigma_1^2 = \bar{x}_1^2 - x_1^2 = \frac{330}{4} - 9^2 = 1,5$$

$$\sigma_2^2 = \bar{x}_2^2 - x_2^2 = \frac{675}{6} - 10,5^2 = 2,25.$$

Общая дисперсия:

$$\sigma^2 = \overline{x^2} - \bar{x}^2 = \frac{330 + 675}{10} - 9,9^2 = 2,49.$$

Средняя из групповых дисперсий:

$$\overline{\sigma_i^2} = \frac{\sigma_1^2 f_1 + \sigma_2^2 f_2}{f_1 + f_2} = \frac{1,5 \cdot 4 + 2,25 \cdot 6}{4 + 6} = 1,95.$$

Межгрупповая дисперсия:

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 f_i}{\sum f_i} = \frac{(9 - 9,9)^2 \cdot 4 + (10,5 - 9,9)^2 \cdot 6}{4 + 6} = 0,54.$$

Суммирование средней из групповых дисперсий и межгрупповой дает величину, равную общей дисперсии:  $1,95 + 0,54 = 2,49$ . Каждая из дисперсий, входящих в соотношение (10.13), имеет вполне определенный смысл. Применительно к рассмотренному примеру общая дисперсия характеризует общую вариацию количества деталей, изготавливаемых за смену всеми токарями 5-го и 6-го разрядов. Она отражает суммарное влияние всех возможных факторов: квалификации токарей, технического состояние станков, обеспеченность инструментом и материалами и т. д. Внутригрупповые дисперсии показывают величину вариации, вызванную всеми этими факторами, кроме различий в квалификационном разряде (внутри групп все токари имеют одну квалификацию). Средняя из внутригрупповых дисперсий, очевидно, также отражает вариацию, обусловленную всеми факторами, кроме квалификации токарей, но в среднем по всей совокупности их.

Межгрупповая дисперсия характеризует вариацию групповых средних, обусловленную различиями групп по квалификационному разряду. Очевидно, чем больше вклад межгрупповой дисперсии в общую дисперсию, тем, по-видимому, сильнее влияние группировочного признака (квалификационного разряда) на величину изучаемого признака (количество изготавливаемых деталей). Поэтому в статистическом анализе широко используется показатель, представляющий собой долю межгрупповой дисперсии в общей дисперсии. Он носит название эмпирического коэффициента детерминации:

$$\eta^2 = \frac{\delta^2}{\sigma^2}. \quad (10.14)$$

Этот коэффициент показывает, какая часть общей вариации изучаемого признака обусловлена вариацией группировочного признака.

В нашем примере  $\eta^2 = 0,54/2,49 = 0,217$ . Это означает, что 21,7 % общей дисперсии выработки обусловлены вариацией квалификационных разрядов. Остальные 78,3 % общей дисперсии обусловлены вариацией прочих факторов (техническое состояние станков, ритмичность подачи заготовок, качество резцов, возраст токарей, состояние трудовой дисциплины на предприятии и т. д.).

Корень квадратный из эмпирического коэффициента детерминации носит название эмпирического корреляционного отношения:  $\eta = \sqrt{\eta^2}$ . Оно применяется для оценки тесноты связи между группировочным и изучаемым (результативным) признаками (см. гл. 12).

#### 10.4. ОБЩЕЕ ПРЕДСТАВЛЕНИЕ ОБ АНАЛИЗЕ ВАРИАЦИОННЫХ РЯДОВ

Рассматривая графическое изображение вариационных рядов (полигон распределения, гистограмму), нетрудно прийти к выводу о том, что значения частот (частостей), плотностей распределения определенным образом зависят от значений вариантов. Эти зависимости весьма разнообразны, но для одноступенчатых явлений общественной жизни они носят сходный в качественном отношении, закономерный характер.

Так, например, рассматривая распределение рабочих бригады, цеха или завода по уровню квалификации, мы всегда замечаем, что с ростом квалификационного разряда численности рабочих вначале растут, а по достижении некоторой определенной величины — падают. Такой характер распределения обусловлен в основном существующей спецификой производства и потребностью в рабочих различной квалификации.

Совсем иной характер имеет распределение людей по возрасту: по мере увеличения возраста численности людей уменьшаются, что связано в первую очередь с особенностями человека как биологического вида.

Такая четко прослеживающаяся зависимость распределения частот (частостей) от вариации изучаемого признака выражает закономерность распределения. Конкретный вид распределения частот в вариационных рядах определяется характером рассматриваемого явления, его природой, факторами, влияющими на вариацию признака, и зависит от разных причин. Во-первых, это основные причины, непосредственно формирующие закономерность распределения и органически связанные с природой рассматриваемого явления. Во-вторых, это причины второстепенные, случайные для данного явления, влияние которых проявляется в виде более или менее существенных отклонений от закономерного распределения.

Например, распределение рабочих-сдельщиков промышленного предприятия по размеру месячной заработной платы опре-

делается прежде всего теми принципами, на которых основана сделанная форма оплаты труда. Поэтому закономерность распределения в этом случае обусловлена квалификационным составом рабочих, принятыми нормами выработки и расценками, уровнем организации труда, техническим состоянием технологического оборудования и т. д. Однако нельзя не учитывать и вторую группу причин, вызывающих более или менее заметные отклонения от закономерного распределения: возраст и состояние здоровья отдельных рабочих, их семейные обстоятельства и бытовые условия, межличностные отношения в коллективе и многое другое.

Важнейшей задачей анализа вариационных рядов в статистике является выявление закономерности распределения и определение ее характера. Закономерности распределения, как и любые статистические закономерности, достаточно четко проявляются в больших совокупностях. Следовательно, для успешного проведения статистического анализа вариационного ряда необходимо, чтобы этот ряд содержал достаточно большое количество единиц совокупности. Немалое значение имеет и правильность построения самого вариационного ряда, которая достигается научно обоснованным выбором числа групп и размера интервалов ряда.

Читателю известно, что важнейшим условием обеспечения эффективности статистического анализа является качественная однородность изучаемой совокупности. Применительно к анализу вариационных рядов это означает, что закономерность распределения отчетливо проявляется лишь в рядах, построенных для качественно однородных единиц. Можно, например, выявить отдельно закономерности распределения по размеру заработной платы рабочих-сдельщиков и рабочих-повременщиков в какой-либо отрасли промышленности, поскольку эти закономерности обусловлены соответствующими принципами, лежащими в основе сдельной и повременной оплаты труда. Однако получить отчетливую закономерность распределения по размеру заработной платы для всех рабочих (как сдельно, так и повременно работающих) нам не удастся, так как в этом случае вариационный ряд будет образован из качественно неоднородных единиц. В подобных случаях распределение может вообще не обрисовываться, либо выглядеть как двухвершинное (бимодальное) или даже многовершинное.

Характер распределения отражает общие условия, вытекающие из природы изучаемого явления, из присущих ему закономерностей и особенностей, влияющих на вариацию признака. Определенное представление о закономерностях распределения дают графики вариационных рядов — полигон распределения и гистограмма распределения. Но они характеризуют так называемое эмпирическое распределение — распределение, полученное на основании группировки материалов статистического наблюдения и поэтому отражающее как основные причины, воз-

действующие на характер распределения, так и причины случайные, не характерные для изучаемого явления.

Основная задача анализа вариационных рядов — выявление подлинной закономерности распределения путем исключения влияния второстепенных, случайных для исследуемого явления факторов. Элиминирование влияния случайных для данного распределения факторов достигается увеличением объема исследуемой совокупности при одновременном уменьшении интервалов ряда.

Из курса математической статистики известно, что в результате увеличения объема совокупности и уменьшения интервалов полигон распределения все более и более приближается к некоторой плавной линии, являющейся для него пределом и носящей название *кривой распределения*. Сказанное в полной мере относится и к гистограмме, которая может быть превращена в полигон соединением отрезками прямой средних точек вершин столбиков.

Получение кривой распределения на основе полигона или гистограммы можно представить себе лишь для гипотетического случая, соответствующего бесконечно большому числу единиц совокупности и бесконечно малой ширине интервала ряда. Лишь при этих идеализированных условиях кривая распределения будет выражать функциональную связь между значениями варьирующего признака и соответствующими им частотами (частотами) и представлять так называемое *теоретическое распределение*.

Теоретическое распределение, в отличие от эмпирического, складывается под влиянием только основных, существенных для данного явления причин и имеет поэтому очень большое значение для изучения массовых общественных явлений. Но получение кривой распределения из эмпирических данных (полигон, гистограмма) возможно лишь для описанного идеального случая. К тому же многообразие явлений общественной жизни не позволяет найти универсальное аналитическое выражение функциональной зависимости частот (частостей) от значений варьирующего признака. Поэтому при проведении анализа вариационного ряда целесообразно свести эмпирическое распределение к одному из хорошо исследованных видов теоретического распределения, рассматриваемых математической статистикой. Последнее при этом играет роль некоторой идеализированной модели эмпирического распределения, а сам анализ вариационного ряда сводится к сопоставлению эмпирического и теоретического распределений и определению степени различия между ними.

В статистике широко используются различные виды теоретических распределений — нормальное распределение, биномиальное распределение, распределение Пуассона и др. Каждое из теоретических распределений имеет свою специфику и свою область применения в различных отраслях знания.



Нормальное распределение, предложенное Лапласом, является одним из самых употребительных и выражается уравнением:

$$y_t = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \quad (10.15)$$

где  $y_t$  — ордината кривой нормального распределения;  $t = \frac{x - \bar{x}}{\sigma}$  — так называемое нормированное отклонение.

Кривая нормального распределения (рис. 10.1) симметрична, имеет максимум при  $t=0$  (или при  $x=\bar{x}$ ), обе ее ветви имеют точки перегиба при  $t = \pm 1$  и асимптотически приближаются к оси абсцисс. Кроме того, это распределение характерно тем, что в пределах  $x \pm \sigma$  заключено 68,3 % всех частот, в пределах  $x \pm 2\sigma$  заключено 95,4 % всех частот, в пределах  $x \pm 3\sigma$  — 99,7 % всех частот ряда распределения.

В силу симметричности кривой нормального распределения средняя арифметическая, медиана и мода для

этого распределения равны между собой. В нормальном распределении выражается закономерность, отражающая результат взаимодействия множества различных причин, влияющих на вариацию признака. При достаточно большом объеме вариационного ряда взаимодействие случайных, нетипичных причин приводит к их взаимопоглощению, и нормальное распределение является следствием действия причин коренных, вытекающих из сущности изучаемого явления.

Разумеется, в социально-экономических явлениях нормальное распределение в чистом виде не встречается, но оно служит удачной моделью, с которой сравниваются анализируемые эмпирические распределения.

Для анализа эмпирического распределения и выяснения того, в какой мере оно обусловлено действием внутренних, коренных причин, производится выравнивание фактического распределения признака по кривой нормального распределения. При этом используется то обстоятельство, что нормальное распределение полностью определяется двумя параметрами —  $\bar{x}$  и  $\sigma$ .

Вначале по фактическому распределению вычисляются нормированные отклонения  $t$ . Затем определяются соответствующие

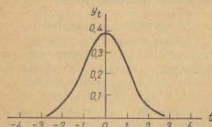


Рис. 10.1. Кривая нормального распределения

им частоты теоретического нормального распределения  $y_t$ . Эти частоты могут быть найдены по уравнению (10.15), но гораздо удобнее находить их по специальным таблицам, обычно приводящимся в учебниках математической статистики<sup>1</sup>.

И наконец, рассчитываются теоретические частоты нормального распределения по формуле

$$f_t = \frac{y_t h \sum f}{\sigma}, \quad (10.16)$$

где  $h$  — ширина интервала ряда, а  $\sum f$  — объем анализируемого ряда.

Расчет теоретических частот проиллюстрируем на примере ранее рассмотренного распределения электродвигателей по мощности (табл. 10.6).

Таблица 10.6. К построению кривой нормального распределения

$x$	$f$	$x - \bar{x}$	$t = \frac{x - \bar{x}}{\sigma}$	$y_t$	$f_t$
20	2	52,6	2,56	0,0151	1,47
40	8	32,6	1,59	0,1127	10,98
60	35	12,6	0,61	0,3312	32,28
80	41	7,4	0,36	0,3739	36,44
100	9	27,4	1,34	0,1626	15,85
120	4	47,4	2,31	0,0277	2,70
140	1	67,4	3,28	0,0040	0,39
Всего	100	—	—	—	100,11

В расчете, сделанном в табл. 10.6, в качестве значений варьирующего признака приняты середины интервалов из табл. 10.4 и использованы значения  $\bar{x} = 72,6$  кВт и  $\sigma = 20,52$  кВт. При вычислениях всем значениям разностей  $x - \bar{x}$  приписывался «+», поскольку функция (10.15) — четная. Следует обратить внимание на определенные расхождения между значениями эмпирических и теоретических частот, обусловленные некоторым несоответствием эмпирического распределения нормальному. Различие же в суммах эмпирических и теоретических частот ( $\sum f_t = 100$ ,  $\sum f = 100,11$ ) объясняется округлениями при расчетах.

График, приведенный на рис. 10.2, наглядно иллюстрирует отличие эмпирического распределения от нормального.

<sup>1</sup> См. например: Венецкий И. Г., Кильдишев Г. С. Теория вероятностей и математическая статистика. — М.: Статистика, 1975, с. 255.

Для различных вариационных рядов могут наблюдаться большие или меньшие различия между значениями эмпирических и теоретических частот. По величине этих различий можно судить о степени близости эмпирического распределения к нормальному. Если расхождения между эмпирическими и теоретическими частотами невелики, их можно объяснить действием случайных для данного явления факторов и считать фактическое распределение близким к нормальному. Если же эти расхождения существенны, то их причиной может быть несоответствие рассматриваемого распределения нормальному.

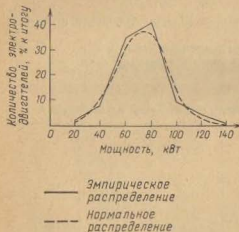


Рис. 10.2. Распределение электродвигателей по мощности (кВт)

Критерий согласия Пирсона ( $\chi^2$ ) определяется по формуле

$$\chi^2 = \sum_{i=1}^k \frac{(t_i - h)^2}{h} \quad (10.17)$$

Для оценки близости эмпирического распределения к теоретическому определяется вероятность  $P(\chi^2)$  достижения этим критерием данной величины. Если эта вероятность превышает 0,05, то отклонения фактических частот от теоретических считаются случайными, несущественными. Если же  $P(\chi^2) < 0,05$ , то отклонения считаются существенными, а эмпирическое распределение — принципиально отличным от теоретического.

Величина критерия  $\chi^2$  зависит не только от расхождений между эмпирическими и теоретическими частотами, но и от числа групп, из которых состоит совокупность. Чем на большее число групп разбита совокупность<sup>1</sup>, тем больше оказывается величина  $\chi^2$ . Значения вероятностей  $P(\chi^2)$  табулированы в зависимости от  $\chi^2$  и числа степеней свободы  $k$ . Последнее пред-

<sup>1</sup> См.: Венецкий И. Г., Кильдишев Г. С. Теория вероятностей и математическая статистика, с. 259.

ставляет собой разность между уменьшенным на единицу числом групп в совокупности и числом общих характеристик теоретического распределения. Для нормального распределения число степеней свободы  $k = n - 3$ , где  $n$  — число групп.

Вычислим критерий  $\chi^2$  для рассматриваемого нами распределения электродвигателей по их мощности (табл. 10.7).

В соответствии с (10.17)  $\chi^2 = 6,336$ ,  $k = 7 - 3 = 4$ . При этих значениях  $\chi^2$  и  $k$  из таблицы вероятностей  $P(\chi^2)$  находим, что  $P(\chi^2) > 0,05$ . Этот результат говорит о том, что отклонения фак-

Таблица 10.7. К расчету  $\chi^2$

$x$	$i_j$	$t_i$	$i_j - t_i$	$(i_j - t_i)^2$	$(i_j - t_i)/t_i$
20	2	1,47	0,53	0,281	0,191
40	8	10,98	-2,98	8,880	0,809
60	35	32,28	2,72	7,398	0,225
80	41	36,44	4,56	20,794	0,571
100	9	15,85	-6,85	46,923	2,960
120	4	2,70	1,30	1,690	0,626
140	1	0,39	0,61	0,372	0,954
Всего	100	100,11	—	—	6,336

тических частот от теоретических незначительны, носят случайный характер, а само изучаемое распределение близко к нормальному.

Величины  $\chi^2$  и  $k$  используются также для вычисления критерия согласия Романовского  $|\chi^2 - k|/\sqrt{2k}$ . Если величина этого критерия меньше 3, расхождение между эмпирическим и теоретическим распределением можно считать несущественным. Если же этот критерий оказывается больше 3, то следует считать, что теоретическое распределение не может служить моделью для изучаемого эмпирического. Для нашего примера этот критерий

$$\frac{6,336 - 4}{\sqrt{2 \cdot 4}} = 0,826,$$

что значительно меньше 3. Это подтверждает близость рассмотренного распределения электродвигателей к нормальному.

Весьма употребительным критерием согласия является также критерий согласия Колмогорова ( $\lambda$ ). Его определение основано на сопоставлении сумм накопленных эмпирических и теоретических частот и производится по формуле

$$\lambda = \frac{D}{\sqrt{\Sigma f_i}}, \quad (10.18)$$

где  $D$  — максимальное значение разности между накопленными эмпирическими и теоретическими частотами;  $\Sigma f_i$  — сумма эмпирических частот.

Определим критерий  $\lambda$  на основании данных табл. 10.8.

Из табл. 10.8 видно, что наибольшее значение разности между накопленными эмпирическими и теоретическими частотами равно 4,83. Отсюда величина критерия согласия  $\lambda$  составит

$$\lambda = \frac{4,83}{\sqrt{100}} = 0,483.$$

Далее с помощью специальных таблиц определяется вероятность близости между эмпирическим и теоретическим распределением при найденном значении  $\lambda$  (табл. 10.9).

Таблица 10.8. К расчету  $\lambda$

x	Частоты ряда		Накопленные частоты		Абсолютные значения разностей накопленных эмпирических и теоретических частот
	$I_n$	$I_t$	$\Sigma I_n$	$\Sigma I_t$	
20	2	1,47	2	1,47	0,53
40	8	10,98	10	12,45	2,45
60	35	32,28	45	44,73	0,27
80	41	36,44	86	81,17	4,83
100	9	15,85	95	97,02	2,02
120	4	2,70	99	99,72	0,72
140	1	0,39	100	100,11	0,11
Всего	100	100,11	—	—	—

Как видно из табл. 10.9, значения вероятностей  $P(\lambda)$  изменяются в пределах от 0 до 1. При  $P(\lambda) = 1$  наблюдается полное совпадение эмпирического и теоретического распределений, при  $P(\lambda) = 0$  имеет место полное расхождение между ними.

В нашем случае ( $\lambda = 0,483$ ) вероятность близости между эмпирическим и теоретическим распределениями составляет при-

Таблица 10.9. Таблица вероятностей для  $\lambda$

$\lambda$	$P(\lambda)$	$\lambda$	$P(\lambda)$	$\lambda$	$P(\lambda)$
0,30	1,000	1,00	0,270	1,70	0,006
0,40	0,999	1,10	0,178	1,80	0,003
0,50	0,964	1,20	0,112	1,90	0,002
0,60	0,864	1,30	0,068	2,00	0,0007
0,70	0,711	1,40	0,040	2,10	0,0003
0,80	0,544	1,50	0,022	2,20	0,0001
0,90	0,393	1,60	0,012	2,31	0,0000

близительно 0,970. Отсюда можно сделать вывод о весьма существенной близости рассматриваемого распределения к нормальному.

В общем случае, однако, следует предварительно удостовериться в том, что анализируемое эмпирическое распределение по природе своей тождественно нормальному. Только при этом условии критерий Колмогорова может дать объективную оценку степени их близости.

Рассмотренные критерии согласия дают общую оценку степени близости эмпирического распределения к нормальному, но не содержат никакой информации о характере расхождений между ними. Поэтому при анализе вариационных рядов применяются также специальные показатели, позволяющие охарактеризовать расхождения между эмпирическим и нормальным распределениями как с качественной, так и с количественной стороны.

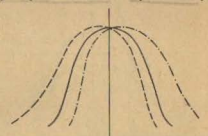
Как уже отмечалось, нормальное распределение строго симметрично, что обусловлено четностью функции (10.15). Фактически же распределения, построенные по эмпирическим данным, как правило, асимметричны, т. е. смещены по отношению к оси симметрии нормального распределения влево или вправо. Для определения направления и величины этого смещения (скошенности) употребляется коэффициент асимметрии:

$$k_A = \frac{m_3}{\sigma^3}, \quad (10.19)$$

где  $m_3$  — центральный момент третьего порядка,  $\sigma^3$  — куб среднего квадратического отклонения.

Из курса математической статистики известно, что для нормального распределения все центральные моменты нечетных порядков равны нулю. В эмпирических же распределениях любой центральный момент нечетного порядка (кроме первого) будет отличаться от нуля в зависимости от характера асимметрии: при левосторонней асимметрии он будет меньше нуля, при правосторонней — больше нуля (рис. 10.3).

Деление центрального момента третьего порядка на куб среднего квадратического отклонения придает коэффициенту асимметрии безразмерный характер и тем самым позволяет использовать его для сравнения между собой различных распределений.



----- Левосторонняя ( $k_A < 0$ )  
 ————— Нулевая ( $k_A = 0$ )  
 - - - - - Правосторонняя ( $k_A > 0$ )

Рис. 10.3. Виды асимметрии распределения



Известно, что нормальное распределение характеризуется равенством средней арифметической, моды и медианы. В реальных же распределениях, построенных на фактических данных, эти величины обычно разнятся между собой. При левосторонней асимметрии  $Mo > Me > \bar{x}$ , при правосторонней асимметрии  $Mo < Me < \bar{x}$ . Это обстоятельство позволяет для характеристики асимметрии использовать более простой показатель:

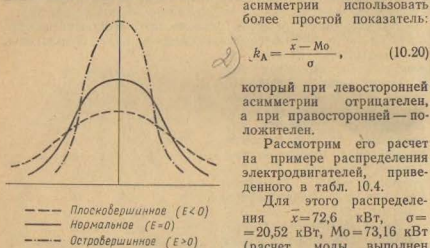


Рис. 10.4. Виды распределений по степени островершинности

Коэффициент асимметрии

$$k_A = \frac{72,6 - 73,16}{20,52} = -0,027,$$

что говорит о наличии очень слабой отрицательной асимметрии.

Анализ вариационного ряда предусматривает также определение эксцесса распределения, т. е. его островершинности по отношению к нормальному распределению. В качестве характеристики эксцесса используется нормированный центральный момент четвертого порядка  $m_4/\sigma^4$ , который, как известно из курса математической статистики, для нормального распределения равен 3. Коэффициент эксцесса для эмпирического распределения представляет собой величину

$$E = \frac{m_4}{\sigma^4} - 3. \quad (10.21)$$

Этот коэффициент положителен при островершинности изучаемого распределения по отношению к нормальному и отрицателен при плосковершинности (рис. 10.4).

Вычислим коэффициент эксцесса для рассматриваемого распределения электродвигателей по мощности, имея в виду, что

центральный момент четвертого порядка представляет собой величину  $m_4 = \sum (x - \bar{x})^4 f / \sum f$ .

Отсюда

$$m_4 = \frac{7124,5 \cdot 10^4}{100} = 71,245 \cdot 10^4 \text{ и}$$

$$E = \frac{71,245 \cdot 10^4}{20,52^4} - 3 = \frac{71,805 \cdot 10^4}{17,73 \cdot 10^4} - 3 = 1,02.$$

Полученный результат свидетельствует о заметной островершинности рассматриваемого распределения по отношению к нормальному.

Таблица 10.10. К расчету коэффициента эксцесса

$x$	$f$	$x - \bar{x}$	$(x - \bar{x})^4 \cdot 10^{-4}$	$(x - \bar{x})^4 \cdot 10^{-4} f$
20	2	-52,6	765,5	1531,0
40	8	-32,6	112,9	903,2
60	35	-12,6	2,5	87,5
80	41	7,4	0,3	12,3
100	9	27,4	56,4	507,6
120	4	47,4	504,8	2019,2
140	1	67,4	2063,7	2063,7
Всего	100			7124,5

## ГЛАВА 11 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

### 11.1. ОБЩЕЕ ПОНЯТИЕ О ВЫБОРОЧНОМ НАБЛЮДЕНИИ

Статистическое наблюдение можно организовать как сплошное и несплошное. Сплошное предусматривает обследование всех единиц изучаемой совокупности явления, несплошное — лишь ее части. К несплошному относится и *выборочное наблюдение*. В основе этого наблюдения лежит идея о том, что отобранная в случайном порядке некоторая часть единиц может представлять всю изучаемую совокупность явления по интересующим исследователя признакам. Целью выборочного наблюдения является получение информации, прежде всего, для определения сводных обобщающих характеристик всей изучаемой совокупности. По своей цели выборочное наблюдение совпадает с одной из задач сплошного наблюдения и поэтому речь может идти о том, какое из двух видов наблюдения — сплошное или выборочное — целесообразнее провести.

При решении этого вопроса необходимо исходить из следующих основных требований, предъявляемых к статистическому наблюдению: информация должна быть достоверной, т. е. максимально соответствовать реальной действительности; сведения должны быть достаточно полными для решения задач исследования; отбор информации должен быть проведен в максимально сжатые сроки для обеспечения ее использования в оперативных целях; денежные и трудовые затраты на организацию и проведение должны быть минимальными.

При выборочном наблюдении эти требования обеспечиваются в большей мере, чем при сплошном. Так как обследуется только часть единиц совокупности, то ошибок регистрации будет меньше, следовательно, информация будет более достоверной. Выборочное наблюдение позволяет собрать более полную информацию за более сжатые сроки при меньших трудовых и денежных затратах. Кроме того, при изучении некоторых явлений невозможно провести сплошное наблюдение. Так, при изучении качества многих видов продукции ее разрушают или уничтожают. Например, качество электрических лампочек определяется продолжительностью их горения в часах, качество консервов — путем их дегустации и т. п.

Сравнивая таким образом сплошное и выборочное наблюдение, следует отметить несомненные преимущества последнего. Поэтому проводить статистическое исследование, предназначенное для получения сводных обобщающих характеристик по всей изучаемой совокупности явления, часто целесообразнее выборочным порядком.

Преимущества выборочного наблюдения по сравнению со сплошным возможно обеспечить, если оно организовано и проведено в строгом соответствии с научными принципами теории выборочного метода. Такими принципами являются: обеспечение случайности отбора единиц и достаточного их числа. Соблюдение этих принципов позволяет получить такую совокупность единиц, которая по интересующим исследователя признакам представляет всю изучаемую совокупность, т. е. является репрезентативной (представительной).

Первый принцип отбора — *обеспечение случайности* — заключается в том, что при отборе каждой из единиц изучаемой совокупности обеспечивается равная возможность попасть в выборку. Случайный отбор — это не беспорядочный отбор. Случайный отбор можно обеспечить только при соблюдении определенной методики. Например, осуществляя отбор по жребию, применяя таблицы случайных чисел и т. д.

Второй принцип отбора — *обеспечение достаточного числа отобранных единиц* излагается в разделе 11.3.

Понятие *репрезентативности* отобранной совокупности единиц не следует понимать как ее представительность во всех отношениях, т. е. по всем признакам изучаемой совокупности.

Такого представительства практически обеспечить невозможно, да такая задача и не ставится. Любое выборочное наблюдение проводится с определенной целью и четко сформулированными конкретными задачами, и понятие репрезентативности должно быть связано с целью и задачами исследования. Отобранная из всей изучаемой совокупности часть должна быть репрезентативной, прежде всего, в отношении тех признаков, которые изучаются или оказывают существенное влияние на формирование сводных обобщающих характеристик.

Введем некоторые понятия, характеристики и придадим им соответствующие условные обозначения.

*Генеральной совокупностью* называется вся изучаемая совокупность единиц, подлежащая изучению по интересующим исследователя признакам. *Выборочной совокупностью* называется отобранная в случайном порядке из генеральной совокупности некоторая ее часть. Характеристиками генеральной и выборочной совокупностей могут служить средние значения изучаемых признаков, их дисперсии и средние квадратические отклонения, мода и медиана и др. Может изучаться и характеристика альтернативного признака, т. е. такого, который может иметь одно из двух его значений — наличие или отсутствие.

Исследователя могут интересовать и распределения единиц по изучаемым признакам в генеральной и выборочной совокупностях. В этом случае частоты будем называть соответственно генеральными и выборочными.

Введем такие условные обозначения:

Показатели	Генеральная совокупность	Выборочная совокупность
Объем (численность единиц)	$N$	$n$
Среднее значение варьирующего признака	$\bar{x}$	$\bar{x}$
Общая дисперсия	$\sigma^2$	$\sigma_n^2$
Средняя из групповых дисперсий	$\overline{\sigma^2}$	$\overline{\sigma_n^2}$
Межгрупповая дисперсия	$\delta^2$	$\delta_n^2$
Доля единиц, обладающих данным качеством	$p$	$\omega$
Доля единиц, не обладающих данным качеством	$q = 1 - p$	$1 - \omega$
Дисперсия альтернативного признака	$\sigma_p^2 = pq$	$\sigma_{\omega}^2 = \omega(1 - \omega)$
Частоты	$f$	$f_n$

Информация, получаемая в результате любого статистического наблюдения, имеет расхождение с реальной действительностью. Такое расхождение получило название ошибок статистического наблюдения. При массовом наблюдении ошибки неизбежны, но возникают они в результате действия различных причин (см. гл. 4).

В данной главе рассматривается только ошибка репрезентативности и причины ее возникновения. Под *ошибкой репрезентативности* (представительства) понимают расхождение между выборочной характеристикой и предполагаемой характеристикой генеральной совокупности. Причиной образования этой ошибки является то обстоятельство, что обследуются не все единицы генеральной совокупности, а лишь их некоторая часть, и различия между единицами, попавшими в выборку, не соответствуют различиям единиц не попавшим в выборку. Вследствие этого выборочная совокупность становится непредставительной по отношению к генеральной совокупности. Ошибка репрезентативности может возникнуть по двум причинам: из-за нарушения научных принципов отбора — *систематическая ошибка* — и в результате случайности отбора — *случайная ошибка*. В результате первой причины выборка легко может оказаться смещенной, так как при отборе каждой единицы допускается ошибка, всегда направленная в одну и ту же сторону. Эта ошибка получила название *ошибки смещения*. Ее размер может превышать величину случайной ошибки. Особенность ошибки смещения состоит в том, что, представляя собой постоянную часть ошибки репрезентативности, она увеличивается с увеличением объема выборки. Случайная же ошибка с увеличением объема выборки уменьшается. Кроме того, величину случайной ошибки можно определить (см. ниже), тогда как размер ошибки смещения непосредственно практически определить очень сложно, а иногда — невозможно. Поэтому необходимо знать причины, вызывающие ошибку смещения и меры, способствующие ее устранению.

Ошибки смещения бывают преднамеренные и непреднамеренные. Причиной возникновения *преднамеренной ошибки* является тенденциозный подход к выбору единиц из генеральной совокупности. Мерой устранения этой ошибки может быть только исключение тенденциозности. Выявить эту ошибку можно только путем проведения повторного отбора с обязательным соблюдением принципа случайности.

*Непреднамеренные ошибки* могут возникать на стадии подготовки выборочного наблюдения, формирования выборочной совокупности и анализа ее данных. Чаще всего создаются условия для возникновения ошибок смещения на стадии подготовки выборочного наблюдения. Недостаточно хорошо продуманные и четко сформулированные взаимосвязанные вопросы плана ор-

ганизации и проведения выборочного обследования могут дать информацию, не соответствующую цели исследования или, что еще хуже, вводящую в заблуждение. Если при сплошном наблюдении это возможно только при преднамеренном искажении фактов, то при выборочном это связано с непреднамеренными ошибками смещения. При разработке плана организации и проведения выборочного наблюдения особое внимание следует уделять *единице отбора*, т. е. такой единице изучаемой совокупности, которая является основанием самого процесса отбора. Единицей отбора могут служить естественные единицы изучаемого явления, например предприятие, рабочий, покупатель, семья и т. д. В некоторых случаях необходимо создать искусственные единицы, не соответствующие естественному делению изучаемой совокупности. Удачное установление единицы отбора уменьшает вероятность получить смещенную выборку.

Сокращению опасности возникновения ошибок смещения во многом способствует хорошая *основа выборки*, т. е. та генеральная совокупность, из которой предполагается производить отбор, например список единиц отбора. Поэтому при подготовке выборочного наблюдения необходимо особенно тщательно ознакомиться с тем, какова основа выборки, пригодна ли она для производства отбора, позволит ли она образовать несмещенную выборку. Если готовой основы выборки нет, то ее необходимо построить.

Основа выборки должна быть достоверной, полной и соответствовать цели исследования, а единицы отбора и их характеристики должны соответствовать действительному их состоянию на момент подготовки выборочного наблюдения. Если основа выборки не отвечает перечисленным требованиям, ее необходимо либо существенно улучшить, внести соответствующие изменения, уточнения, дополнения, либо создать заново.

На стадиях формирования выборочной совокупности и производства наблюдения ошибки смещения особенно опасны, так как их трудно заметить и исправить. При формировании выборочной совокупности ошибку смещения чаще всего дает неточное соблюдение установленного порядка отбора, предусматривающего отбор вполне определенных единиц. Иногда может показаться, что выборочная совокупность «не пострадает», если, например, вместо предусмотренной десятой единицы по списку взять одиннадцатую или двенадцатую; в действительности же такое нарушение установленного порядка отбора нередко приводит к смещенной выборке.

Значительную ошибку может вызвать замена единиц во время производства наблюдения. Нередки случаи, когда в отношении некоторых единиц, попавших в выборку, трудно собрать сведения из-за их отсутствия на момент наблюдения, нежелания дать сведения и т. п. В таких случаях эти единицы приходится заменять другими. Замена неравноценными единицами в отно-



шении интересующих исследователя признаков часто создает условия для образования смещенной выборки.

Ошибки смещения при анализе данных могут возникнуть из-за неправильных приемов распространения выборочных характеристик на генеральную совокупность (см. 11.4).

Случайная ошибка выборки возникает в результате случайных различий между единицами, попавшими в выборку, и единицами генеральной совокупности, т. е. она связана со случайным отбором. Теоретическим обоснованием появления случайных ошибок выборки является теория вероятностей и ее предельные теоремы.

В разработке и обосновании случайных ошибок выборки, как одного из разделов теории выборочного метода и предельных теорем, составляющих их основу, большую роль сыграли крупнейшие русские математики П. Л. Чебышев, А. М. Ляпунов, А. А. Марков и др.

Сущность предельных теорем состоит в том, что в массовых явлениях совокупное влияние различных случайных причин на формирование закономерностей и обобщающих характеристик будет сколь угодно малой величиной или практически не зависит от случая. Так как случайная ошибка выборки возникает в результате случайных различий между единицами выборочной и генеральной совокупностей, то при достаточно большом объеме выборки она будет сколь угодно мала. Этот вывод, опирающийся на доказательство предельных теорем, позволяет предполагать, что характеристики выборочного наблюдения могут достаточно хорошо представлять характеристики генеральной совокупности.

Предельные теоремы исходят из закона нормального распределения, согласно которому большая часть выборочных средних сосредоточивается около генеральной средней ( $\bar{x}$ ). Следовательно, закон нормального распределения теоретически позволяет установить, в какой мере изменяется размер случайной ошибки выборки с изменением вероятности ее появления. Так как многие массовые явления подчиняются закону нормального распределения, то он служит основой при оценке вероятности тех или иных результатов выборочного наблюдения.

Предельные теоремы теории вероятностей позволяют определять размер случайных ошибок выборки. Различают среднюю (стандартную) и предельную ошибку выборки. Под *средней (стандартной) ошибкой выборки* понимают такое расхождение между средней выборочной и генеральной совокупностей ( $\bar{x} - \bar{x}$ ), которое не превышает  $\pm \sigma$ . *Предельной ошибкой* выборки принято считать максимально возможное расхождение ( $\bar{x} - \bar{x}$ ), т. е. максимум ошибки при заданной вероятности ее появления.

На основании теоремы, доказанной П. Л. Чебышевым, величину стандартной ошибки так называемого (собственно-слу-

чайного отбора» (см. формулу 11.5) при достаточно большом объеме выборки ( $n$ ) можно определить по формуле

$$\mu = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}, \quad (11.1)$$

где  $\mu$  — стандартная ошибка.

Величина стандартной ошибки прямо пропорциональна колеблемости признака в генеральной совокупности и обратно пропорциональна корню квадратному из объема выборки. Величина  $\mu$  зависит также от способа и вида отбора (см. 11.7).

Академик А. М. Ляпунов, продолжив исследования П. Л. Чебышева в этой области, доказал, что вероятность появления случайной ошибки выборки при достаточно большом ее объеме подчиняется закону нормального распределения. Эта вероятность определяется по формуле

$$F(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{t^2}{2}} dt. \quad (11.2)$$

Значения функции  $F(t)$  табулированы при разных значениях  $t$ .

Предельная ошибка выборки определяется по формуле

$$\Delta = t\mu, \quad (11.3)$$

где  $\Delta$  — предельная ошибка,  $t$  — заданный коэффициент доверия.

Так при  $t=1$  величина предельной ошибки составит  $\Delta = \mu$ , которая гарантируется с вероятностью 0,683. Это означает, что в 683 выборках из тысячи подобных максимальная ошибка выборки (предельная) не превысит  $\pm \mu$ . При  $t=2$  с вероятностью 0,954 она не выйдет за пределы  $\pm 2\mu$  и т. д. В практике выборочных наблюдений массовых общественных явлений максимальный предел ошибок, как правило, вполне достаточен в пределах  $\pm 3\mu$ .

Однако приведенные формулы (11.1) и (11.3) практически непригодны для расчета ошибок, так как в них  $\sigma$  — это показатель колеблемости признака в генеральной совокупности, который неизвестен, как неизвестна и генеральная средняя. Но в теории вероятностей доказано, что

$$\sigma^2 = \sigma_0^2 \frac{n}{n-1}. \quad (11.4)$$

Так как  $\frac{n}{n-1}$  при достаточно больших  $n$  — величина

близкая к единице, то можно принять, что  $\sigma^2 \approx \sigma_0^2$ . На основании этого утверждения в формулах (11.1) и (11.3) вместо генеральной дисперсии принимают значение выборочной дис-

персии. Тогда формулы для расчета стандартной и предельной ошибок выборки соответственно будут иметь вид:

$$\mu = \sqrt{\frac{\sigma_b^2}{n}} = \frac{\sigma_b}{\sqrt{n}}; \quad + 24 (11.5) \quad (11.5)$$

$$\Delta = t\mu = t\sqrt{\frac{\sigma_b^2}{n}}, \quad (11.6)$$

где  $\sigma_b^2$  — дисперсия выборочной совокупности.

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

$$\bar{x} - \Delta \leq \bar{x} \leq \bar{x} + \Delta. \quad (11.7)$$

Это означает: с заданной вероятностью можно утверждать, что значение генеральной средней можно ожидать в пределах от  $\bar{x} - \Delta$  до  $\bar{x} + \Delta$ , т. е. что доверительный интервал ( $\bar{x} - \Delta$ ,  $\bar{x} + \Delta$ ) с заданной вероятностью заключает в себе генеральную среднюю.

Наряду с абсолютной величиной предельной ошибки выборки рассчитывается и относительная ошибка выборки, которая определяется как процентное отношение предельной ошибки выборки к соответствующей характеристике выборочной совокупности:

$$\Delta_{\%} = \frac{\Delta}{\bar{x}} \cdot 100, \quad \Delta_{\%} = \frac{\Delta}{\bar{w}} \cdot 100. \quad (11.8a, б)$$

Пусть, например, по данным 2 %-ного выборочного обследования жирности молока, реализуемого на колхозном рынке города, средняя жирность 36 проб молока составила 3,8 % при среднем квадратическом отклонении 0,8 %. Найдем с вероятностью 0,954 доверительный интервал средней жирности молока. Здесь  $n=36$ ;  $\bar{x}=3,8$  %;  $\sigma_b=0,8$  %;  $F(t)=0,954$ ;  $t=2$ . Следовательно, стандартная ошибка выборки  $\mu=0,8/\sqrt{36}=0,13$  %; предельная ошибка выборки  $\Delta=2 \cdot 0,13=0,26$  %; относительная ошибка  $\Delta_{\%}=(0,26/3,8) \cdot 100=6,8$  %. Доверительный интервал  $3,8-0,26 \leq \bar{x} \leq 3,8+0,26$ . Итак, с вероятностью 0,954 можно утверждать, что средняя жирность реализуемого на колхозном рынке молока в день обследования находится в пределах от 3,54 до 4,06 %.

Если при выборочном наблюдении изучению подлежит альтернативный признак, то случайная ошибка выборки для доли определяется с теоремой Я. Бернулли. Так как вероятность расхождения между частотой и долей тоже подчиняется закону нормального распределения, то стандартная

ошибка выборки альтернативного признака определяется по формуле

$$\mu = \sqrt{\frac{pq}{n}},$$

где  $pq$  — дисперсия альтернативного признака в генеральной совокупности.

Так как  $pq$  — неизвестно, то в практике ее заменяют дисперсией выборочной совокупности  $w(1-w)$  и формула принимает вид

$$\mu = \sqrt{\frac{w(1-w)}{n}}.$$

Предельная ошибка выборки и доверительный интервал для альтернативного признака определяются аналогично формулам (11.6) и (11.7).

Например, если известно, что при выборочном обследовании 100 рабочих (1 %-ный случайный отбор) у 40 % из них производственный стаж оказался более пяти лет, то стандартная ошибка выборки, рассчитанная по формуле (11.9), составит

$$\mu = \sqrt{\frac{0,4 \cdot 0,6}{100}} = 0,049, \text{ или } 4,9 \%, \text{ а предельная ошибка с вероятностью } 0,954 \text{ равна: } \Delta = 2 \cdot 0,049 = 0,098, \text{ или } 9,8 \%.$$

Следовательно, доверительный интервал (в процентах) таков:  $40,0-9,8 \leq p_{\%} \leq 40,0+9,8$ , т. е. с вероятностью 0,954 можно утверждать, что в генеральной совокупности долю рабочих с производственным стажем более пяти лет можно ожидать в пределах от 30,2 до 49,8 %.

### 11.3. ОПРЕДЕЛЕНИЕ НЕОБХОДИМОЙ ЧИСЛЕННОСТИ ВЫБОРКИ

Одним из научных принципов в теории выборочного метода является обеспечение достаточного числа отобранных единиц. Теоретически необходимость соблюдения этого принципа представлена в доказательствах предельных теорем теории вероятностей, которые позволяют установить, какой объем единиц следует выбрать из генеральной совокупности, чтобы он был достаточным и обеспечивал репрезентативность выборки.

Расчет необходимого объема выборки строится с помощью формул, выведенных из формул предельных ошибок выборки ( $\Delta$ ), соответствующих тому или иному виду и способу отбора. Так, для собственно-случайного повторного отбора имеем

$$\Delta_{\text{сеп}} = t\mu_{\text{сеп}} = t\sqrt{\frac{\sigma_n^2}{n}}, \quad (11.9)$$

откуда

$$n = \frac{t^2 \sigma_n^2}{\Delta^2}. \quad (11.10)$$

Смысл этой формулы в том, что при случайном повторном отборе объем необходимой численности выборки прямо пропорционален квадрату коэффициента доверия ( $t^2$ ) и дисперсии вариационного признака ( $\sigma^2$ ) и обратно пропорционален квадрату предельной ошибки выборки ( $\Delta^2$ ). В частности, с увеличением предельной ошибки в два раза необходимая численность выборки может быть уменьшена в четыре раза. Из трех параметров два ( $t$  и  $\Delta$ ) задаются исследователем. При этом исследователь исходя из цели и задач выборочного исследования должен решить вопрос: в каком количественном сочетании лучше включить эти параметры для обеспечения оптимального варианта? В одном случае его может устраивать в большой мере надежность полученных результатов ( $t$ ), нежели мера точности ( $\Delta$ ), в другом — наоборот. Сложнее решить вопрос в отношении величины предельной ошибки выборки, так как этим показателем исследователь на стадии проектировки выборочного наблюдения не располагает. Поэтому в практике принято задавать величину предельной ошибки выборки, как правило, в пределах до 10 % предполагаемого среднего уровня признака. К установлению предполагаемого среднего уровня можно подходить по-разному: использовать данные подобных, ранее проведенных исследований или же воспользоваться данными основы выборки и произвести небольшую пробную выборку.

Наиболее сложно установить при проектировке выборочного наблюдения третий параметр в формуле (11.10) — дисперсию выборочной совокупности. В этом случае необходимо использовать всю информацию, имеющуюся в распоряжении исследователя, как-то: ранее проведенные подобные исследования, пробные исследования и др.

При изучении альтернативного признака (доли  $p$ ) объем необходимой численности выборки определяется по следующей формуле:

$$n = \frac{t^2 w (1 - w)}{\Delta^2} \quad (11.11)$$

Смысл этой формулы тот же, что и формулы (11.10). На стадии проектировки выборочного наблюдения, если невозможно, хотя бы приблизительно определить дисперсию альтернативного признака [ $w(1-w)$ ], она принимается равной своему максимуму  $= 0,25$ .

Вопрос об определении необходимой численности выборки усложняется, если выборочное исследование предполагает изучение нескольких признаков единиц отбора. В этом случае средние уровни каждого из признаков и их вариация, как правило, различны и поэтому решить вопрос о том, дисперсии какого из признаков отдать предпочтение, возможно лишь с учетом цели и задач исследования.

#### 11.4. ПРИЕМЫ РАСПРОСТРАНЕНИЯ ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК НА ГЕНЕРАЛЬНУЮ СОВОКУПНОСТЬ

Конечной целью любого выборочного наблюдения является распространение его характеристик на генеральную совокупность. Распространять эти характеристики можно с помощью различных приемов. Применение того или иного приема распространения зависит от цели выборочного исследования.

**Прямой пересчет** данных выборки на всю совокупность применяется в том случае, когда целью исследования является определение объема признака генеральной совокупности, если известна лишь численность ее единиц. Предположим, что в результате проведенного выборочного наблюдения установлена средняя урожайность пшеницы, равная 25 ц/га. Предельная ошибка выборки при вероятности 0,954 составила 0,5 ц/га. Известно, что посевная площадь, занятая под этой культурой, — 2000 га (генеральная совокупность). Необходимо установить предполагаемую величину валового сбора пшеницы с этой площади. Учитывая предельную ошибку, можно утверждать, что с вероятностью 0,954 валовой сбор пшеницы в генеральной совокупности будет не меньше, чем 49 тыс. ц  $[(25 - 0,5) \cdot 2000]$ , и не больше, чем 51 тыс. ц.

В том случае, если выборочное наблюдение проводится с целью уточнения результатов сплошного наблюдения, применяется **метод коэффициентов**. В практике этот прием, например, применяется для уточнения данных ежегодных сплошных переписей скота, находящегося в личной собственности населения. После проведения сплошной переписи, проводится 10 %-ное выборочное наблюдение и устанавливается так называемый «процент недоучета» при сплошном наблюдении. Так, например, если данные сплошного учета скота в хозяйствах попавших в выборку показали 200 голов, а данные выборки в этих же хозяйствах — 205 голов, то, следовательно, «процент недоучета» при сплошном наблюдении составит:  $\frac{5 \cdot 100}{200} = 2,5\%$ . Этот процент и есть тот коэффициент, который

необходимо распространить на всю генеральную совокупность. Так, если по данным сплошного учета установлено, что поголовье скота, находящегося в личной собственности населения, составляет в районе 10 000 голов, то в эту величину следует внести поправку, умножив ее на коэффициент 1,025. Тогда поголовье скота в районе следует принять равным 10 250 голов.

Во всех случаях, когда целью выборочного исследования является определение тех или иных обобщающих характеристик генеральной совокупности, определяются доверительные интервалы: нижняя граница — обобщающая характеристика выборочной совокупности минус предельная ошибка, верхняя граница — обобщающая характеристика выборочной совокупности плюс предельная ошибка.



# ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ





СТАТИСТИКА КАК НАУКА  
ОРГАНИЗАЦИЯ СТАТИСТИКИ

СБОР СТАТИСТИЧЕСКОЙ  
ИНФОРМАЦИИ  
И ОБЕСПЕЧЕНИЕ КАЧЕСТВА  
ЕГО МАТЕРИАЛОВ

СТАТИСТИЧЕСКАЯ СОВОКУПНОСТЬ  
И ЕЕ ХАРАКТЕРИСТИКА

СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ  
ВЗАИМОСВЯЗЕЙ

АНАЛИЗ ДИНАМИКИ

# ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

ПОД РЕДАКЦИЕЙ  
А.М. ГОЛЬДБЕРГА,  
В.С. КОЗЛОВА

Допущено Министерством  
высшего и среднего специ-  
ального образования СССР  
в качестве учебника для сту-  
дентов экономических спе-  
циальностей высших учебных  
заведений

  
199617 ОНБУ

ПРОВЕРКА  
200 9



МОСКВА  
„ФИНАНСЫ И СТАТИСТИКА“  
1985



А. Я. БОЯРСКИЙ, Л. Л. ВИКТОРОВА,  
А. М. ГОЛЬДБЕРГ, В. Р. КАМЕНЕЦКИЙ, В. С. КОЗЛОВ,  
Г. Н. КОЗЛОВА, А. С. КРАВЕЦ, Л. П. ОЛЕСЕВИЧ,  
Я. М. ЭРЛИХ

Рецензенты:

М. Р. ЕФИМОВА, Е. В. ПЕТРОВА, В. С. КНЯЗЕВСКИЙ



Общая теория статистики: Учебник/А. Я. Боярский, Л. Л. Викторова, А. М. Гольдберг и др.; Под ред. А. М. Гольдберга, В. С. Козлова.— М.: Финансы и статистика, 1985.— 367 с. ил.

В пер.: 1 р. 10 к. 20 000 экз.

Рассмотрены вопросы организации статистики, сбора статистической информации, характеристики статистической совокупности, статистического изучения взаимосвязей, анализа рядов динамики и индексного анализа. Приводится методология расчета относительных и средних величин, показателей вариации, регрессии, корреляции и динамики, а также выборочного и индексного методов. Использованы данные ежегодников «Народное хозяйство СССР». Примеры условные.

0 0702000000—129  
010(01)—85 80—85

ББК 60.6

31

© Издательство «Финансы и статистика», 1985

## ПРЕДИСЛОВИЕ

Успешное выполнение задач, поставленных Основными направлениями экономического и социального развития СССР на 1986—1990 годы и на период до 2000 года (проект) требует дальнейшего улучшения качества подготовки экономистов высшей квалификации, повышения эффективности учебно-воспитательного процесса. В связи с этим в последние годы значительно обновлены и пересмотрены учебные планы и программы, в которых отражены решения партии и правительства по актуальным проблемам экономического и социального развития страны, совершенствования хозяйственного механизма, новейшие достижения в производстве, науке и технике, в экономической теории и практике. Естественно, что эта работа предъявляет повышенные требования к учебникам и учебным пособиям — важнейшим средствам обучения и самостоятельной работы будущих специалистов в области экономики. Это относится в первую очередь к фундаментальным и специальным дисциплинам, формирующим общетеоретический кругозор, профессиональные знания и практические навыки экономиста любого профиля. Важное место среди этих дисциплин занимает общая теория статистики.

На основе учета этих требований и обобщения опыта применения методов статистического исследования, а также преподавания статистики авторским коллективом и подготовлен настоящий учебник.

Исходя из задач повышения уровня и действенности экономико-статистического анализа обстоятельно изложены такие вопросы курса, как группировки, анализ регрессий и корреляций, анализ динамики, статистическое моделирование и прогнозирование, индексы.

Подробно рассмотрены вопросы организации, содержания статистического наблюдения и обеспечения качества статистической информации в условиях применения новейшей вычислительной техники и внедрения автоматизированной системы государственной статистики.

Более углубленное изложение некоторых вопросов, представляющее интерес для студентов специальности «статистика», набрано в пособии петитом. Этим же шрифтом набраны также



исторические сведения и некоторые дополнительные разъяснения по изложенному материалу.

Учебник подготовлен межвузовским авторским коллективом, в состав которого вошли преподаватели статистики Московского государственного университета им. М. В. Ломоносова, Киевского института народного хозяйства им. Д. С. Коротченко и Одесского института народного хозяйства.

Авторы учебника: проф. А. Я. Боярский — гл. 1; доц. Л. Л. Викторова — гл. 11; проф. А. М. Гольдберг — гл. 2, 3, 5; доц. В. Р. Каменецкий — гл. 7, 8, 9, 10; доц. В. С. Козлов — гл. 12, 13, 14, 15; доц. Г. Н. Козлова — гл. 2, 3, 4, 6; доц. А. С. Кравец — гл. 8, 12, 13; доц. Л. П. Олесевич — гл. 5, 6, 7, 13; доц. Я. М. Эрлих — гл. 12, 13.

Авторы приносят благодарность доцентам кафедры статистики Московского института управления М. Р. Ефимовой и Е. В. Петровой, а также заведующему кафедрой общей теории статистики Ростовского института народного хозяйства профессору В. С. Князевскому за обстоятельное рецензирование рукописи и ценные замечания по улучшению ее структуры и содержания.

## РАЗДЕЛ I

## СТАТИСТИКА КАК НАУКА. ОРГАНИЗАЦИЯ СТАТИСТИКИ

### ГЛАВА 1

### СТАТИСТИКА КАК НАУКА. ПРЕДМЕТ И МЕТОД СТАТИСТИКИ

#### 1.1. СОВОКУПНОСТЬ И ВАРИАЦИЯ

Статистика в ее современном понимании связана со счетом каких-то объектов, с их количественными характеристиками. Из этого, однако, не следует, что всякий счет — статистика. Некто подсчитал, что в доме, в котором ему предстоит поселиться, 8 этажей, что в отведенной ему квартире 36 м<sup>2</sup>, т. е. на одного члена семьи придется по 9 м<sup>2</sup>. Назвать это статистикой нельзя. Но если подсчитать, что из всех построенных в данном пятилетии и в данном городе домов столько-то 5-этажных, столько-то 6-этажных и т. д., что в них живет столько-то человек и, следовательно, сколько-то квадратных метров приходится на одного жителя, — это статистика.

Или другой пример. С заводского конвейера вчера сошло 110 автомобилей. И это еще не статистика. Наблюдением, которое велось изо дня в день, установлено число автомобилей, сошедших с конвейера в каждый день; из этих чисел составлен ряд, рассмотрение которого показало, что при общей тенденции к некоторому возрастанию числа автомобилей отмечаются периодические колебания этого числа по дням недели. Это — статистика и не совсем простая.

Из этих примеров видно, что отличает статистику от простого счета. Количественные характеристики, устанавливаемые статистикой, не являются зафиксированными раз и навсегда, одинаковыми для всех объектов. Они меняются (варьируют) от объекта к объекту (от дома к дому, от человека к человеку, от дня ко дню) и вообще во времени. Но чтобы такие изменения (вариации) могли наблюдаться, необходимо множество объектов — домов, людей, дней. Таким образом, статистика исследует количественные характеристики, которые при определенных обстоятельствах могут варьировать. Важно заметить, что вариации, о которых идет речь в статистике, это не просто различие. Например, в часах определенной марки шестеренки различаются диаметром, числом зубцов или дни года различаются величиной светового дня. Эти изменения к статистике не имеют отношения: размер данной шестеренки в данном механизме, величина светового дня такого-то числа и месяца (в данном месте) могут быть только такими, а не иными. С другой стороны, длительность службы шестеренки до пол-



- качественные (интенсивные) 322
- объемные (экстенсивные) 322
- отчетные 68
- первичные 69
- плановые 68
- синтетические 69
- сводные 70
- Признак статистический 8, 17, 42, 43, 63
  - альтернативный 155
  - атрибутивный 64
  - варьирующий 63
  - группировочный 75
  - количественный 64
  - несущественный 63
  - результативный 196, 217
  - существенный 63
  - факторный 196, 217
- Прирост абсолютный 269
  - базисный 269
  - средний 282, 284
  - цепной 269
- Прогнозирование 263, 316
- Промилле 118
- Регрессионно-корреляционный анализ (РКА) 214—264
  - построение модели 217
  - применение 259—264
  - требования к информации 216—217
  - этапы 215
  - шаговый анализ 251
- Регрессия множественная (многофакторная модель) 239
  - парная (однофакторная модель) 218
- Ряд вариационный 92
  - распределения 92
- Варианты 92
- Ряды динамики (временные) 265
  - анализ ряда 268
    - нескольких рядов 304—307
  - автокорреляция уровней 307
  - остатков 311
  - авторегрессия 307
  - виды 267
    - интервальные 266
    - моментные 266
  - выбор формы тренда 295, 300
  - выравнивание (аналитическое) 295
    - по параболе 300
    - по показательной кривой 299, 321
    - по прямой 297
  - интерполяция 319
  - корреляция 307—316
  - регрессия 307—316
  - сглаживание с помощью скользящей средней 293
  - случайная компонента уровня (остаток) 295—296, 310
  - смыкание рядов 273
- тренд 291
- укрупнение интервалов 292
- экстраполяция 316
- Сводка
  - машинизированная 94
  - программа 73
  - ручная 94
- Связь криволинейная 198
  - корреляционная 197
  - обратная 198
  - прямая 198
  - прямолинейная 198
  - стохастическая 197
  - функциональная 196
- Совокупность выборочная 171
  - генеральная 171
  - статистическая 7—8, 61
- Способ моментов 137, 155
- Среднее квадратическое отклонение 152
  - линейное отклонение 151
- Средняя величина 132—133
  - арифметическая 133
  - взвешенная 134
  - гармоническая 139
  - геометрическая 143, 285
  - определяющий показатель 141, 284
  - простая (невзвешенная) 134
  - степенная 143
  - хронологическая 281
- Таблица статистическая 99
  - групповая 101
  - комбинационная 102
  - макет 100
  - оформление 105
  - подлежащие 100
  - простая 101
  - сказуемое 100
- Темп прироста 272
  - средний 285
- Темп роста 118, 272
  - базисный 271
  - средний 285, 287, 292
  - цепной 271
- Удельный вес 121
- Уравнение регрессии 214
- Уровень ряда динамики 265
  - интервального 277
  - средний 277
  - моментного 278—281
  - средних величин 281
  - относительных величин 281
- Учет 21
  - бухгалтерский 25
  - оперативно-технический 25
  - статистический 26
- Центральное статистическое управление (ЦСУ СССР) 27
- Частоты 92, 134, 143

## ОГЛАВЛЕНИЕ

	Предисловие . . . . .	5
<b>РАЗДЕЛ I</b>	<b>Статистика как наука. Организация статистики</b>	<b>7</b>
ГЛАВА 1	Статистика как наука. Предмет и метод статисти- стики . . . . .	7
ГЛАВА 2	Организация статистики . . . . .	21
<b>РАЗДЕЛ II</b>	<b>Сбор статистической информации и обеспечение качества его материалов . . . . .</b>	<b>36</b>
ГЛАВА 3	Статистическое наблюдение . . . . .	36
ГЛАВА 4	Обеспечение качества материалов статистического наблюдения . . . . .	51
<b>РАЗДЕЛ III</b>	<b>Статистическая совокупность и ее характеристика</b>	<b>61</b>
ГЛАВА 5	Совокупность как объект статистического изуче- ния . . . . .	61
ГЛАВА 6	Сводка и группировка . . . . .	72
ГЛАВА 7	Способы наглядного представления статистиче- ских данных . . . . .	99
ГЛАВА 8	Абсолютные и относительные величины . . . . .	114
ГЛАВА 9	<u>Средние величины</u> . . . . .	132
ГЛАВА 10	<u>Показатели вариации</u> и анализ вариационных рядов . . . . .	150
ГЛАВА 11	Выборочное наблюдение . . . . .	169
<b>РАЗДЕЛ IV</b>	<b>Статистическое изучение взаимосвязей . . . . .</b>	<b>195</b>
ГЛАВА 12	Взаимосвязи социально-экономических явлений и методы их статистического изучения . . . . .	195
ГЛАВА 13	Анализ регрессий и корреляций . . . . .	214
<b>РАЗДЕЛ V</b>	<b>Анализ динамики . . . . .</b>	<b>265</b>
ГЛАВА 14	<u>Анализ рядов динамики</u> . . . . .	265
ГЛАВА 15	<u>Индексный анализ</u> . . . . .	321
	<u>Предметный указатель</u> . . . . .	364





ходимо за короткий срок получить предварительные итоги выборочных обследований. Взаимопроникающие выборки эффективны для оценки результатов обследования. Если в независимых выборках результаты одинаковы, то это свидетельствует о надежности данных выборочного обследования. Взаимопроникающие выборки иногда можно применять для проверки работы различных исследователей, поручив каждому из них провести обследование разных выборок.

Стандартная ошибка при взаимопроникающих выборках определяется так же, как при типической пропорциональной выборке (формула 11.16). Взаимопроникающие выборки по сравнению с другими видами требуют больших трудовых затрат и денежных расходов. Поэтому исследователь должен учитывать это обстоятельство при проектировании выборочного обследования.

Предельные ошибки при различных способах отбора и видах выборки определяются по формуле  $\Delta = t_{\mu}$ , где  $\mu$  — соответствующая стандартная ошибка.

#### 11.8. ПРАКТИКА ПРИМЕНЕНИЯ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ В СОВЕТСКОЙ СТАТИСТИКЕ

Выборочные наблюдения нашли широкое применение в практике советской статистики. Их используют во всех отраслях народного хозяйства с целью изучения тенденций и закономерностей различных социально-экономических явлений и процессов и получения информации для решения многих народнохозяйственных задач. Особое место среди выборочных наблюдений занимают бюджетные обследования семей рабочих, служащих и колхозников. Широко применяются в нашей практике статистические методы контроля качества продукции, в основе которых лежит выборочный метод. Выборочные обследования проводятся для изучения заработной платы рабочих по отраслям и тарифным разрядам, причин текучести рабочих и т. п.

За последние годы значительное место в выборочных наблюдениях занимают обследования по вопросам социальной статистики, статистики культуры, здравоохранения. Исследуются социальные аспекты трудовой деятельности рабочих, служащих, колхозников, активности молодежи, культурного уровня населения. Изучается общественное мнение населения о работе предприятий бытового обслуживания, учреждений здравоохранения, дошкольных учреждений и др.

В советской статистике выборочные обследования нередко используются в сочетании с проведением сплошных наблюдений. При формировании выборочной совокупности применяются в сочетании различные виды выборочного метода в зависимости от конкретной цели обследования, характера основы

выборки и единицы отбора. Так, в практике статистики труда и заработной платы, торговли и бюджетов для ряда обследований использовался типический отбор в сочетании с механической выборкой. В статистике сельского хозяйства и культуры при организации выборочных обследований применялся гнездовой отбор в сочетании с механическим. В социальной статистике при формировании выборочной совокупности использовался двухступенчатый и двухфазный отбор.

Рассмотрим более подробно некоторые наиболее крупные выборочные обследования.

Органы государственной статистики на протяжении многих лет ведут *выборочные обследования бюджетов семей рабочих, служащих и колхозников*. Основная цель таких обследований состоит в получении и разработке данных о структуре доходов и расходов семей различных групп рабочих (квалифицированных и малоквалифицированных) в разрезе отраслей промышленности, колхозников в разрезе отраслей сельскохозяйственного производства, семей служащих. Ежегодное изучение бюджетов семей позволяет следить за структурными изменениями доходов и расходов, устанавливать сложившиеся закономерности, изучать изменение уровня благосостояния населения. Обследование бюджетов населения проводится с добровольного согласия семьи, попавшей в выборку. Бюджетные записи ведутся в течение года под контролем работников органов государственной статистики. Собранные сведения разрабатываются и анализируются по соответствующей программе.

Выборочные бюджетные обследования организованы в соответствии с научными принципами теории выборочного метода. Эти обследования представляют собой комбинированную выборку. Выборочная совокупность формируется как многоступенчатая выборка. На каждой из ступеней установлена соответствующая единица отбора. На всех ступенях отбор бесповоротный. На различных ступенях выборочная совокупность формируется по-разному. На одних — как типическая пропорциональная выборка, на других — как механическая. На конечной ступени формируется выборочная совокупность семей, бюджеты которых и подлежат обследованию.

Организация формирования выборочной совокупности семей рабочих и служащих проводится в определенной последовательности, по этапам. Сначала определяется общее число бюджетов, подлежащих выборочному обследованию<sup>1</sup>, затем общее число бюджетов распределяется по республикам, краям, областям пропорционально общей численности рабочих тех отраслей промышленности, которые выделены для обследования в каждой республике, крае, области. Выборочная совокупность семей

<sup>1</sup> В 1984 г. выборочному обследованию подверглись бюджеты 62 тыс. семей рабочих, служащих и колхозников.



рабочих и служащих формируется путем двухступенчатого отбора.

На первой ступени единицей отбора является предприятие. Для формирования выборочной совокупности на этой ступени создается основа выборки — списки всех предприятий по каждой из отраслей, попавших в число обследуемых. Списки составляются в порядке убывания размера средней месячной заработной платы промышленно-производственных рабочих. Из этих списков производится механический отбор предприятий. Сформированная на первой ступени выборочная совокупность предприятий проверяется на репрезентативность. Степень репрезентативности определяется путем сопоставления выборочной и генеральной средней месячной заработной платы. Если отклонение  $(\bar{x} - \bar{x})$  оказывается выше или ниже установленного предела, то выборочная совокупность соответствующим образом корректируется.

На второй ступени единицей отбора являются рабочие и служащие предприятий, попавших в выборку. Отбор на этой ступени осуществляется как типическая пропорциональная выборка. Внутри выделенных типических групп отбор рабочих и служащих производится механическим бесповторным способом. Для этого создается соответствующая основа выборки — списки рабочих и служащих каждой типической группы, составленные в порядке убывания размера их среднемесячной заработной платы. После того как выборочная совокупность семей рабочих и служащих сформирована, проверяется ее репрезентативность по тому же принципу, что и на первой ступени отбора.

В основу обследования бюджетов семей колхозников положены те же принципы, что и бюджетов семей рабочих и служащих. Выборочная совокупность семей колхозников формируется как двухступенчатая выборка. Однако при этом имеется ряд особенностей. Так, на первой ступени отбираются колхозы. Для этого создается основа выборки — списки типических групп колхозов, образованные по признакам производственного направления колхозов и размера оплаты человека-дня работы. На второй ступени отбираются семьи колхозников. Основой выборки на этой ступени являются списки семей внутри каждой из типических групп, образованные по признакам количества отработанных человеко-дней всеми членами семьи и уровня доходности личного подсобного хозяйства. Внутри типических групп на каждой из ступеней отбор производится как механическая выборка.

Проверка репрезентативности отобранных колхозов и семей колхозников производится по группе установленных показателей. В случае отклонения, выходящих за установленные пределы, выборочная совокупность корректируется.

Всесоюзная перепись населения 1979 г. была проведена в сочетании сплошного наблюдения с выборочным. Как сплошное наблюдение перепись населения проводилась по программе, состоящей из 11 вопросов. Как выборочное — по программе 16 вопросов, из которых последние пять были дополнительными (см. приложение).

Для организации отбора была построена основа выборки, т. е. составлены списки всех жилых помещений. Единичей отбора являлось жилое помещение. Отбор производился как механический бесповторный. В пределах каждого инструкторского участка отбиралось каждое 4-е жилое помещение, и лица, постоянно проживающие в них, переписывались по программе, состоящей из 16 вопросов. Сформированная таким образом выборочная совокупность составила 25 % населения от его общей численности. Достаточно высокая доля отбора была вызвана необходимостью получить репрезентативные характеристики не только по стране в целом, но и по административно-территориальным подразделениям.

Контроль за качеством продукции осуществляется на практике, как правило, в виде выборочного. При этом применяются различные методы. В основе каждого из них лежат научные принципы теории выборочного метода.

Контроль за качеством продукции осуществляется в двух направлениях. Первое состоит в том, что контролю подлежит качество готовой продукции. Если при установлении качества продукции ее приходится разрушать или уничтожать, то контроль проводится в виде малой выборки. В других случаях организуется обычная выборка.

Второе направление — это профилактический контроль или контроль в процессе производства с целью предупреждения появления брака. Такой контроль называется «статистическим контролем качества продукции». Организовать этот контроль возможно на предприятиях с массовым автоматизированным производством. Это связано с тем, что предупредить появление брака в ходе производства продукции можно лишь в том случае, если брак может появиться в результате действия какой-то постоянной причины. Такой причиной при массовом автоматизированном производстве продукции является разладка станка-автомата. Но так как стандарт качества продукции всегда устанавливается с определенными допусками, то уловив начало разладки станка и одновременно отремонтировав его, можно предупредить появление брака. На этих посылах и строится теория «статистического контроля качества продукции» в ходе производства.

При этом обязательным является построение контрольных карт, при которых графически отображается процесс контроля качества продукции. На контрольной карте вычерчиваются две горизонтальные линии, соответствующие нижней и верхней границам, так называемого предупредительного допуска. Затем выборочно с интервалом в один-два часа непосредственно со станков берется «проба» продукции, т. е. несколько единиц изделия (4,5 и более). Для каждой единицы устанавливаются значения тех или иных признаков, определяющих их качество. Результаты, которыми могут быть индивидуальные значения, средние, размах, медиана, среднее квадратическое отклонение

## ГЛАВА 12

ВЗАИМОСВЯЗИ  
СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ  
ЯВЛЕНИЙ И МЕТОДЫ ИХ  
СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ

## 12.1. ВЗАИМОСВЯЗИ

СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ  
И ЗАДАЧИ ИХ  
СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ

Глубокое познание социально-экономических явлений предполагает всесторонний анализ существующих между ними взаимосвязей. Это важное требование вытекает из марксистско-ленинского положения о всеобщей взаимосвязи и взаимообусловленности явлений и процессов общественной жизни. «Чтобы действительно знать предмет, — писал В. И. Ленин, — надо охватить, изучить все его стороны, все связи и «опосредствования»»<sup>1</sup>.

Иллюстрацией диалектической взаимосвязи между социально-экономическими явлениями может служить, например, взаимообусловленность производства и потребления. Производство является необходимой предпосылкой для потребления. С другой стороны, потребление, являясь следствием производства, оказывает на него обратное действие.

Подобные причинно-следственные связи существуют и изучаются статистикой не только на макроуровне, но и в рамках отдельных отраслей, экономических районов, предприятий. Так, увеличение использования основных производственных фондов, рост фондоотдачи в производственном объединении приводит к увеличению выпуска продукции и прибыли этого объединения. В свою очередь рост прибыли позволяет увеличить фонды экономического стимулирования объединения, в том числе и ту их часть, которая используется для развития производства, приобретения новых основных фондов.

Статистическое изучение различных явлений общественной жизни позволяет количественно выражать сложные взаимосвязи и закономерности их развития лишь в том случае, когда исследование основывается на результатах *качественного* марксистско-ленинского политэкономического анализа. Качественный анализ должен не только предшествовать статистическому

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 42, с. 290.

и др., наносятся на график в виде точек. Если точки, нанесенные на график, начинают концентрироваться возле одной из линий предупредительного допуска, значит, можно ожидать появление брака. Это и служит сигналом для остановки производственного процесса и его налаживания.

В практике применяется несколько вариантов статистического метода контроля качества продукции: метод крайних значений, метод медианы и размаха, метод средних арифметических и размахов, метод оценки качества по баллам, метод оценки качества по отдельным признакам и др. Выбор того или иного варианта зависит от оценки параметров, определяющих качество продукции, и особенностей производственного процесса.

*Социологические выборочные обследования* в практике советской статистики организуются, как правило, в виде анкетного способа сбора информации. Анкетные выборочные обследования имеют ряд особенностей, связанных прежде всего с тем, что получаемая информация представляет собой мнения анкетированных лиц. В связи с этим для обеспечения репрезентативности особенно необходима тщательная подготовка обследования и проведение его со строгим соблюдением установленных правил. Особое внимание при этом должно уделяться разработке программы наблюдений.

При проведении выборочных анкетных обследований особые трудности вызывает отсутствие основы выборки или ее надежность. В этом случае в практике советской статистики при подготовке социологических анкетных обследований нередко используют ранее сформированную выборочную совокупность. Так, для изучения культурного уровня населения была использована выборочная совокупность семей рабочих, служащих и колхозников, сформированная для обследования их бюджетов,

Одним из примеров проведенных за последнее время органами государственной статистики социологических анкетных выборочных обследований может служить обследование профессиональной ориентации молодежи в 1981 г. Выборочная совокупность формировалась как двухступенчатая выборка. На первой ступени были отобраны городские и сельские школы (10 %), на второй ступени — учащиеся выпускных десятых классов (50 %). В программу было включено 19 вопросов с несколькими вариантами ответов. Вот некоторые из них: проходили ли вы в школе трудовое воспитание и какую профессию осваивали; имели ли возможность выбрать профессию в процессе трудового обучения; что намерены делать после окончания средней школы, выбрали ли уже свою будущую профессию и др. В результате анализа собранных сведений были установлены намерения школьников, мотивы выбора профессии и ряд других важных сведений, необходимых для целей совершенствования трудового воспитания школьников и профориентационной работы среди молодежи.

На ближайшую перспективу органами государственной статистики намечается провести ряд выборочных обследований для получения материалов, характеризующих влияние отдельных социальных-экономических факторов на формирование и развитие семьи. Так, в 1985 г. проведено крупное *Всесоюзное 5 %-ное выборочное социально-демографическое обследование населения*.

исследованию в целом, но также является критерием справедливости результатов, получаемых на каждой конкретной стадии изучения взаимосвязи. В противном случае полученные статистические характеристики, отражающие различные аспекты проявления изучаемых причинно-следственных связей, будут представлять собой математические абстракции, лишенные содержательного смысла и практической ценности.

Качественный анализ позволяет, в частности, выделить в цепи взаимосвязанных признаков те, которые в данном конкретном случае изменяются независимо от других, — *факторные (независимые)* признаки, и те, величина которых обусловлена изменением факторных признаков, — *результативные (зависимые)* признаки. Очевидно, что отнесение определенного признака к числу факторных или результативных обусловлено социально-экономической сущностью явления и конкретными задачами исследования. Поэтому один и тот же признак при различной постановке задачи может быть либо факторным, либо результативным. Например, если рассматривать производительность труда как результативный показатель, то факторными признаками можно считать уровень механизации производства, стаж работы рабочего, его квалификацию. В другом случае производительность труда непосредственно выступает как независимая переменная, а в роли результативного признака рассматривается, например, прибыль предприятия.

Различные явления по-разному реагируют на изменение факторных признаков. *Связь* между величинами называется *функциональной*, если каждому возможному значению одной из них соответствует одно, вполне определенное, значение другой. Функциональная связь является строгой, точной, полной зависимостью. Чаще всего она встречается в естественных науках: математике, физике, астрономии и т. д. Например, длина окружности всегда в 6,28 раза больше ее радиуса, высота ртутного столба в термометре строго пропорциональна температуре окружающей среды. Функциональная зависимость может связывать результативный признак с несколькими факторными признаками. Так, величина тока в цепи зависит от напряжения и сопротивления, площадь треугольника — от величины его сторон.

Характерной особенностью функциональных связей является то, что нам обычно известен полный перечень факторов, определяющих величину результативного признака, а также механизм этого влияния, выраженный определенным уравнением. Реже встречаются подобные связи в общественной жизни, в частности в экономических процессах. Если, например, при сдельной оплате труда рабочий получает 30 коп. за одну изготовленную деталь, то его общий заработок будет функционально зависеть от количества изготовленных деталей и т. п.

Для социально-экономических явлений характерно то, что наряду с существенными факторами, определяющими в основ-

ном величину результативного признака, на него оказывают воздействие многие другие, в том числе и случайные факторы. Поэтому существующая зависимость не проявляется здесь в каждом отдельном случае, как при функциональных связях, а лишь в общем и среднем при большом числе наблюдений. В таких случаях говорят о наличии *стохастической (вероятностной) связи*, при которой с изменением факторного признака меняется распределение единиц совокупности по результативному признаку.

Так, если рассмотреть оценки в период экзаменационной сессии по различным дисциплинам, то можно заметить, что у одних и тех же студентов они часто неодинаковы. Иными словами, оценки стохастически зависят от различной сложности отдельных учебных дисциплин.

Частным случаем стохастической связи является *корреляционная связь*, при которой каждому значению факторного признака могут соответствовать несколько значений результативного признака, образующих ряд распределения. При этом изменение независимой величины влечет за собой изменение среднего значения зависимого показателя. Следовательно, корреляционная связь является «неполной» зависимостью, которая проявляется не в каждом отдельном случае, а только в средних величинах при достаточно большом числе случаев. Известно, например, что повышение квалификации работника ведет к росту производительности его труда. Это справедливое положение, подтверждаемое в массе явлений, отнюдь не означает, что у двух или трех рабочих одного разряда, занятых аналогичным процессом, будет одинаковая производительность труда. Скорее всего, уровни их выработки будут отличаться, хотя и незначительно, один от другого. Ведь у этих рабочих могут быть различными стаж работы, техническое состояние станка, состояние здоровья и т. д.

Если рассмотреть ряд предприятий отрасли, выпускающих одноименную продукцию, то, при прочих равных условиях, предприятия с большим объемом выпуска характеризуются более низким уровнем себестоимости единицы продукции. Однако связь эта является не функциональной, а корреляционной: нам не известен весь перечень факторов, влияющих на уровень себестоимости, а также точный механизм зависимости между ней и объемом выпуска продукции. Наличие других факторов, определяющих себестоимость, приводит к тому, что в ряде случаев указанная связь нарушается, так что предприятия, имеющие больший объем выпуска, может иметь такую же или даже более высокую себестоимость, чем предприятие с меньшим объемом продукции. Вместе с тем в целом (в среднем) корреляционная связь проявляет себя: при группировке рассматриваемых предприятий по объему производства группам с большим выпуском соответствует более низкая средняя себестоимость единицы продукции.



Рассмотрим приведенные ранее примеры с точки зрения направлений взаимосвязи. Так, с ростом квалификации рабочих производительность их труда в среднем возрастает, а со снижением — уменьшается. Такое же направление взаимосвязи наблюдаем между длиной окружности и ее радиусом. Если функциональные или корреляционные связи характеризуются совпадением направлений изменения факторного и результативного признаков, то говорят о наличии *прямой связи*.

Рост объема выпуска продукции ведет к снижению себестоимости ее единицы, а уменьшение выпуска — к увеличению себестоимости. Аналогично связаны между собой объем продукции и трудоемкость ее единицы. Если функциональные или корреляционные связи характеризуются несовпадением направлений изменения факторного и результативного признаков, то говорят о наличии *обратной связи* между ними. Следует иметь в виду, что прямое и обратное направление связи может соответствовать изменению факторного признака лишь в определенных пределах. Так, при прочих равных условиях, результаты спортсмена улучшаются с увеличением объема тренировочных нагрузок, т. е. имеется прямая связь. Однако если нагрузки будут постоянно возрастать и превысят определенный оптимальный уровень, то дальнейший их рост приведет к снижению спортивных результатов из-за перетренированности спортсмена.

По своему аналитическому выражению связи, изучаемые статистикой, могут быть *прямолинейными (линейными)* и *криволинейными (нелинейными)*. При линейной связи зависимость между факторными и результативными признаками может быть выражена в виде уравнения прямой, а при криволинейной — в виде уравнения кривой (гиперболы, параболы, показательной функции и т. п.).

Не следует смешивать понятия прямой и прямолинейной (линейной) связи, ибо они характеризуют различные стороны этого явления: в первом случае речь идет о направлении связи, во втором — о форме ее аналитического выражения, в частности прямолинейная связь может характеризовать и обратную зависимость.

Функциональные связи являются полными, строгими и поэтому точно выражаются определенным выражением, например:  $C = 2\pi R$ . Корреляционные же связи, будучи неполными, нестрогими, выражаются определенным уравнением лишь приблизительно.

При классификации взаимосвязей учитывается также число факторных признаков, оказывающих влияние на результативный показатель. Если исследуется зависимость результативного признака от какого-либо одного фактора, то принято говорить об *однофакторной связи*. Например, при сдельной оплате труда заработка рабочего за день функционально зависит от количества изготовленных деталей. По-иному следует рассматри-

вать однофакторную корреляционную связь; ведь известно, что при корреляционных зависимостях на результативный признак влияет множество факторов. Следовательно, если мы все же рассматриваем только один фактор, то влияние других просто не учитываем, абстрагируемся от него. Так, при исследовании влияния на дневную выработку рабочего только уровня механизации мы абстрагируемся от уровня его квалификации, стажа работы и ряда других факторов. Подобная *корреляция* называется *парной*, так как исследуются только два признака. Гораздо чаще в практике статистического исследования анализируется влияние на результативный признак нескольких факторов, т. е. рассматривается *множественная корреляция*, которая позволяет глубже изучить интересующее нас явление и определить пути возможного изменения его в нужном направлении.

Изучая взаимосвязи общественных явлений и процессов, статистика дает числовые выражения проявляющимся в них закономерностям. Числовая характеристика взаимосвязей достигается в результате расчета целого ряда статистических показателей, отражающих различные аспекты зависимости между факторными и результативными признаками. Учитывая, что статистическое исследование чаще всего проводится на сравнительно небольшой, ограниченной по объему совокупности, полученные статистические характеристики взаимосвязей должны быть проверены на *существенность (значимость)* своих числовых выражений. Цель такой проверки состоит в том, чтобы выяснить, не может ли являться величина того или иного показателя связи только результатом воздействия случайных факторов. Иными словами, считая объект исследования определенной частью некоторой гипотетической генеральной совокупности, необходимо выяснить, насколько типичны для нее найденные показатели связи. Это делается путем проверки статистических гипотез на основе использования различных критериев математической статистики.

## 12.2. СТАТИСТИЧЕСКИЕ МЕТОДЫ ИЗУЧЕНИЯ СТОХАСТИЧЕСКИХ ВЗАИМОСВЯЗЕЙ

Для исследования связей между социально-экономическими явлениями, их тенденций и интенсивности статистика использует ряд специфических методов, применение которых должно быть основано на теоретическом анализе явлений.

Для исследования функциональных связей применяются балансовый и индексный методы. Для изучения стохастических связей используются метод параллельных рядов, метод аналитических группировок, дисперсионный анализ и анализ регрессии и корреляции.

Метод сравнения параллельных рядов позволяет установить направление связи между социально-экономическими явлениями

ямы и процессами посредством сопоставления двух или нескольких рядов показателей. Сущность этого метода состоит в том, что сначала показатели, характеризующие факторный признак, располагаются в порядке возрастания или убывания (ранжируются), а затем параллельно им располагаются соответствующие показатели результивного признака. Сравнение построенных таким образом рядов дает возможность не только подтвердить само наличие связи, но и выявить ее направление. Очевидно, что подобным образом можно сравнивать ряды статистических показателей, изменяющихся во времени, в пространстве, а также ряды распределения.

Метод сравнения параллельных рядов широко применял в своих работах В. И. Ленин. Он указывал, что показатели сравниваемых рядов должны рассчитываться по одинаковой методике, рассматриваться за одно и то же время и, разу-

Таблица 121. Уровень механизации труда и трудоемкость единицы продукции по двум заводам

Год	Уровень механизации труда, %		Трудоемкость единицы продукции, мин	
	завод А	завод Б	завод А	завод Б
1981	60	64	13,6	11,2
1982	61	66	13,0	10,7

меется, между рядами должна существовать причинно-следственная связь.

Фиксирование простого параллелизма в проявлении и развитии некоторых социально-экономических явлений не должно отождествляться с существованием между ними причинно-следственной зависимости. Часто встречаются ряды, имеющие близкие тенденции развития, между которыми не существует причинной связи. Например, за период с 1950 по 1982 г. в СССР постоянно увеличивалась выплавка стали и столь же постоянно снижалась заболеваемость скарлатиной и корью. Есть ли основания считать, что уменьшение заболеваемости вызывалось ростом выплавки стали? Конечно же, нет! Причины в данном случае совершенно иные: это успехи советского здравоохранения в профилактике инфекционных заболеваний, повышение уровня жизни народа. Таким образом, нет ни малейших оснований для обоснованного параллельного сравнения этих рядов.

Рассмотрим соотношение между уровнем механизации труда и трудоемкостью единицы продукции. Теоретический анализ позволяет предположить, что между этими признаками существует взаимосвязь: ведь степень механизации наряду с другими факторами определяет величину рабочего времени, необходимого на каждую отдельную операцию по изготовлению

единицы изделия, т. е. формирует его трудоемкость. Пусть, например, имеются следующие данные по двум заводам, производящим одноименную продукцию (см. табл. 121).

Сравнительный анализ показывает, что между исследуемыми признаками существует обратная связь: на заводе Б, на котором уровень механизации труда в течение двух лет был выше, трудоемкость ниже, чем на заводе А.

В случае, когда сравниваемые ряды состоят из большого числа единиц, направление связи для разных единиц может

Таблица 122. К расчету коэффициента Фехнера и коэффициента корреляции рангов

Номер завода	Уровень механизации, %	Трудоемкость единицы продукции, мин	Знак отклонения		Сопоставление (С) или несоответствие (НС) знаков	Ранги		d	d²
			от х	от у		по х	по у		
1	60	14	—	+	Н	10	2,5	7,5	56,25
2	61	12	—	+	Н	9	6,5	2,5	16,25
3	62	13	—	+	Н	8	4,5	3,5	12,25
4	63	15	—	+	Н	7	1	6	36
5	64	14	—	+	Н	6	2,5	3,5	12,25
6	65	12	+	+	С	5	6,5	—1,5	2,25
7	66	10	+	+	Н	4	8	—4	16
8	67	13	+	+	С	3	4,5	—1,5	2,25
9	68	8	+	—	Н	2	9	—7	49
10	69	7	+	—	Н	1	10	—9	81

$$\bar{x} = 64,5 \quad \bar{y} = 11,8$$

$$\sum d^2 = 273,5$$

оказаться различным. Имеются, например, данные по 10 заводам одного министерства, производящим одноименную продукцию (см. табл. 122).

Сравнительный анализ двух параллельных рядов показывает, что, как правило, с ростом уровня механизации труда трудоемкость единицы продукции снижается, однако имеются и исключения (например, завод № 3 по сравнению с № 2, № 4 по сравнению с тремя первыми). Для ориентировочного выявления наличия связи и ее направления воспользуемся следующим методом. Определим средние значения факторного (уровень механизации труда) и результивного (трудоемкость единицы продукции) признаков и составим ряды отклонений значений этих признаков от соответствующей им средней. Таким образом, мы переходим к параллельному сравнению уже не самих рядов, а их отклонений от средней. Если знаки отклонений в одном ряду часто совпадают со знаками соответствующих отклонений другого ряда, то можно говорить о нали-



ции положительной связи между признаками; если же знаки чаще не совпадают, то можно предположить наличие отрицательной связи. В случае, если совпадений и несовпадений знаков отклонений от средней примерно одинаково, связь, очевиднее всего, отсутствует или является крайне слабой. На этом принципе сопоставления знаков отклонений от средней основан коэффициент Фехнера, позволяющий получить представление о направлении связи и приблизительную характеристику ее тесноты:

$$K = \frac{\Sigma C - \Sigma H}{\Sigma C + \Sigma H}, \quad (12.1)$$

где  $\Sigma C$  — число случаев совпадения знаков отклонений;  $\Sigma H$  — число случаев несовпадения этих знаков.

Коэффициент Фехнера может принимать значения от  $-1$  до  $+1$ . Если  $K = -1$  (знаки всех отклонений не совпадают), то можно говорить об обратной связи между признаками, когда же  $K = +1$  (знаки всех отклонений совпадают), говорят о наличии прямой связи. При  $K = 0$  связь отсутствует. В нашем примере:  $K = (2 - 8) : (2 + 8) = -0.6$ . Таким образом, можно считать, что между уровнем механизации труда и трудоемкостью единицы продукции существует довольно тесная обратная связь.

Для приблизительной оценки направления и тесноты связи между признаками, представленными двумя рядами, можно также использовать коэффициент корреляции рангов. При определении коэффициента корреляции рангов значения факторного признака записываются в возрастающем или убывающем порядке (ранжируются), а затем ранжируются соответствующие им значения результативного признака. В результате получаем ранги, т. е. места, номера единиц совокупности в упорядоченном ряду. При этом в случае наличия одинаковых вариантов каждому из них присваивается среднее арифметическое значение их рангов. Полученные ранги величин факторного и результативного признаков сопоставляются между собой. Если ранги соответствующих значений двух рядов в большинстве своем совпадают, то можно говорить о наличии тесной прямой связи между признаками, если же ранги изменяются в противоположных направлениях, то говорят об обратной связи. При отсутствии какой-либо согласованности в изменении рангов связь либо отсутствует, либо крайне незначительна. Коэффициент корреляции рангов исчисляется по формуле

$$\rho = 1 - \frac{6 \Sigma d^2}{n(n^2 - 1)}, \quad (12.2)$$

где  $d$  — разность между рангами соответствующих величин двух признаков;  $n$  — число единиц в ряду.

Коэффициент корреляции рангов может принимать значения от  $-1$  до  $+1$ . Если все ранги строго изменяются в одном и том же порядке, то  $d = 0$ , а  $\rho = 1$ , что указывает на тесную прямую связь. Если же ранги изменяются в строго противоположном направлении, то  $\rho = -1$  (тесная обратная связь). Значение  $\rho = 0$  характеризует отсутствие связи.

В рассматриваемом примере заводы ранжированы по факторному признаку (уровню механизации труда) в порядке возрастания, которому соответствуют ранги с 10 по 1. При ранжировании соответствующих значений результативного признака в ряде случаев, учитывая равенство величин, им присвоены средние ранги. Например, первому и пятому заводам, занимающим по результативному признаку второе и третье место, присвоен средний ранг  $-2.5$ . Значение коэффициента корреляции рангов в нашем примере составит:

$$\rho = 1 - \frac{6 \cdot 273.5}{10 \cdot (100 - 1)} = -0.66.$$

Это подтверждает вывод, полученный ранее: между уровнем механизации труда и трудоемкостью единицы продукции существует довольно тесная обратная связь.

Коэффициент корреляции рангов имеет определенные преимущества перед другими характеристиками направления и тесноты связи: его можно определить при исследовании данных, которые не поддаются измерению, но ранжируются (оттенки цветов, качество парфюмерной продукции при дегустации и т. п.). Кроме того, он удобен и при анализе некоторых измеряемых данных, в случае когда нет необходимости их точного измерения. Например, мы легко можем ранжировать группу студентов по росту, не прибегая к его фактическому измерению, для исследования зависимости между ростом и спортивными результатами.

К числу важнейших приемов исследования взаимосвязей относится метод аналитических группировок. Чтобы выявить зависимость с помощью этого метода нужно произвести группировку единиц совокупности по признаку-фактору и для каждой группы вычислить среднюю или относительную величину результативного признака. Сопоставляя затем изменение результативного признака по мере изменения признака-фактора, можно выявить направление и характер связи между ними.

Замечательные образцы применения статистических группировок и групповых средних для установления связей и взаимозависимостей общественных явлений дают такие ленинские работы, как «Развитие капитализма в России», «Новые данные о законах развития капитализма в земледелии» и многие другие. Так, в статье «Язык цифр», анализируя размеры заработной платы рабочих до и после революции 1905 г., В. И. Ленин с помощью статистических группировок показывает, что средняя заработная плата рабочих на крупных фабриках увеличилась

после революции значительно, чем на мелких (см. табл. 12.3). Объясняется это тем, что «более энергичная и дружная стачечная борьба рабочих на крупных фабриках привела в результате к более высокому увеличению заработной платы»<sup>1</sup>.

Статистические группировки и групповые средние, как важнейший прием исследования взаимосвязи находят широкое применение в работах советских экономистов. Они применяются при изучении таких вопросов, как экономическая эффективность специализации производства, зависимость между производительностью труда и себестоимостью продукции и т. д. Покажем

Таблица 12.3

Разряды фабрик по числу рабочих	Средний годовой заработок 1 рабочего		
	За пятилетие (1901—1905)	За четырехлетие (1906—1909)	Увеличение в рублях
Имеющие свыше 1000 чел.	196	234	+38
» 501—1000 »	186	231	+45
» 101—500 »	211	238	+27
» 51—50 »	215	240	+25
» 20 и менее	193	207	+14
<i>Всего</i>	201	235	+34

использование аналитических группировок и групповых средних для анализа связи на следующем примере (табл. 12.4).

У отдельных рабочих наблюдаются различные соотношения между квалификацией (разрядом) и производительностью труда (дневной выработкой). В одних случаях повышению разряда соответствует более высокая выработка, в других же, наоборот, несмотря на более высокий разряд, выработка оказывается ниже (см., например, 1-го и 9-го рабочих, 2-го и 8-го и т. д.). Различные соотношения между уровнем механизации труда и выработкой наблюдаются также у отдельных рабочих, даже при одинаковой квалификации. В большинстве случаев у рабочих одинаковой квалификации с повышением уровня механизации труда повышается и дневная выработка. Однако есть и исключения. Так, 1-й и 5-й рабочие имеют одинаковый разряд, причем несколько лучше механизирован труд у 5-го рабочего. Выработка же, наоборот, выше у 1-го рабочего. Все это свидетельствует о том, что если между рассматриваемыми признаками (квалификацией и производительностью труда) есть связь, то она является корреляционной.

Для выяснения наличия связи между уровнем механизации труда и его производительностью (дневной выработкой) произведем группировку рабочих по проценту механизации труда и вычислим для каждой группы среднюю дневную выработку. Вы-

числим также и средний разряд рабочего в каждой группе (табл. 12.5).

Сопоставляя изменение средней выработки с изменением уровня механизации труда, мы видим, что по мере повышения уровня механизации увеличивается и средняя дневная выра-

Таблица 12.4. Квалификация рабочих, уровень механизации их труда и величина дневной выработки из формовки бетона по одному из предприятий

№ п.п.	Разряд рабочего	Уровень механизации труда, %	Дневная выработка, м <sup>3</sup>	№ п.п.	Разряд рабочего	Уровень механизации труда, %	Дневная выработка, м <sup>3</sup>
А	1	2	3	А	1	2	3
1	2	35	3,0	26	3	69	5,0
2	3	50	6,5	27	2	48	2,5
3	3	44	4,8	28	4	82	6,8
4	3	55	5,7	29	4	98	6,6
5	2	39	2,8	30	3	63	6,3
6	3	56	4,7	31	4	79	7,9
7	2	78	4,2	32	3	41	4,6
8	4	44	5,3	33	3	45	4,2
9	3	43	2,0	34	2	75	4,8
10	3	76	6,5	35	3	45	5,8
11	3	58	5,1	36	4	51	4,9
12	4	41	5,5	37	3	55	4,3
13	2	49	3,0	38	4	95	6,4
14	2	58	3,6	39	4	90	7,0
15	4	58	4,5	40	4	70	7,1
16	4	61	6,7	41	3	56	4,4
17	3	42	5,6	42	3	57	5,1
18	3	46	5,2	43	2	48	5,0
19	2	35	3,2	44	3	72	6,1
20	4	55	5,4	45	3	52	5,9
21	3	38	4,5	46	2	33	3,8
22	3	35	5,5	47	3	55	4,6
23	3	25	2,5	48	2	30	3,4
24	4	90	6,2	49	2	67	5,5
25	2	47	4,1	50	3	57	5,9

ботка, что является, конечно, вполне закономерным. Из данных табл. 12.5 видно также, что с увеличением уровня механизации труда изменяется и средний разряд рабочих. Это также вполне закономерное явление, так как работу с применением машин и механизмов поручают в первую очередь более квалифицированным рабочим.

Квалификация рабочих, однако, сама по себе влияет на уровень производительности труда, так как более квалифицированные рабочие имеют лучшую подготовку, навыки в работе, больший производственный опыт. Поэтому представляет интерес проследить зависимость дневной выработки и от этого фактора.

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. 23, с. 429.



При сопоставлении данных табл. 12.5 и 12.6 возникает вопрос: можно ли считать, что повышение средней выработки объясняется в каждом случае только увеличением величины группировочного признака? Ведь наряду с этим повышается величина и другого признака, который также влияет на выработку: так, в табл. 12.5 более высокая выработка 13 рабочих с наиболее высоким уровнем механизации труда одновременно сопровождается и более высоким разрядом (3,3), а в табл. 12.6 наиболее высокой выработке 13 рабочих 4-го разряда соответствует и наиболее высокий уровень механизации труда (70 %).

Таблица 12.5. Зависимость производительности труда от уровня его механизации

Уровень механизации труда, %	Число рабочих	Средний разряд	Средняя выработка на одного рабочего, м <sup>3</sup>	
Менее 45	14	2,8	4,0	
45—64	23	3,0	4,9	
65 и выше	13	3,3	6,2	
Итого	50	3,0	5,0	

Таблица 12.6. Зависимость производительности труда от квалификации рабочих

Разряд рабочих	Число рабочих	Средняя производительность труда	Средняя выработка, м <sup>3</sup>	
2-й	13	49	3,8	
3-й	24	52	5,0	
4-й	13	70	6,2	
Итого	50	56	5,0	

Очевидно, что поскольку производительность труда зависит как от уровня механизации труда, так и от квалификации рабочих, постольку изменение средней выработки в обоих случаях объясняется изменением обоих этих факторов. В частности, наиболее высокая выработка 13 рабочих последней группы в табл. 12.5 (6,2 м<sup>3</sup>) объясняется не только тем, что их труд является наиболее механизированным, но и тем, что они имеют в среднем и более высокую квалификацию (средний разряд 3,3). Соответственно наиболее высокая выработка 13 рабочих 4-го разряда в табл. 12.6 (6,2 м<sup>3</sup>) объясняется не только их более высокой квалификацией, но и тем, что у них наиболее высок в среднем уровень механизации труда (70 %).

Чтобы выявить влияние каждого из этих факторов в более или менее «чистом» виде, нужно устранить, элиминировать влияние другого фактора, используя для этого комбинационную группировку по обоим факторам (табл. 12.7).

Каждая из трех последних граф таблицы показывает изменение выработки в зависимости от роста уровня механизации труда рабочих соответствующего разряда, т. е. влияние квалификации на выработку здесь элиминировано. Во всех трех группах рабочих по квалификации по мере повышения уровня ме-

ханизации труда растет и средняя выработка. Исключение представляют лишь подгруппы рабочих 4-го разряда с уровнем механизации труда менее 45 % и 45—64 %, где выработка оказалась одинаковой. Это может объясняться случайными обстоятельствами, так как число рабочих в этих подгруппах невелико (2 и 4 человека).

При последних показателях каждой строки таблицы характеризуют зависимость выработки от квалификации при более или менее полном элиминировании уровня механизации труда. Од-

Таблица 12.7. Зависимость производительности труда от уровня его механизации и от квалификации рабочих

Уровень механизации труда, %	Число рабочих				Средняя дневная выработка, м <sup>3</sup>			
	всего	в том числе по разрядам			группы рабочих	в том числе рабочих по разрядам		
		2-й	3-й	4-й		2-й	3-й	4-й
Менее 45	14	5	7	2	4,0	3,2	4,2	5,4
45—64	23	5	14	4	4,9	3,6	5,3	5,4
65 и выше	13	3	3	7	6,2	4,8	5,9	6,9
Итого	50	13	24	13	5,0	3,8	5,0	6,2

нако элиминирование влияния этого фактора является неполным, так как интервалы группировки довольно широки. Неполнота элиминирования объясняется как взаимосвязью уровня механизации и квалификации, так и небольшим числом рабочих в каждой подгруппе.

Значительная наглядность в анализе связей достигается с помощью графиков. Для графического изображения зависимости результативного показателя ( $y$ ) от одного фактора ( $x$ ) при элиминировании влияния другого фактора ( $z$ ) могут быть использованы линейные диаграммы. В прямоугольной системе координат по оси абсцисс откладывается тот фактор, влияние которого нужно изобразить ( $x$ ), а по оси ординат отсчитываются значения результативного показателя. Для каждой группы по  $x$  и подгруппы по  $z$  (табл. 12.7) исчисляются средние значения результативного показателя ( $y_{\text{ср}}$ ) и наносятся на график в виде точек. Абсциссами этих точек являются середины интервалов фактора  $x$ . Соединив отрезками прямой последовательные значения  $y_{\text{ср}}$  по каждой подгруппе, получим для каждой из них ломаную линию, выражающую зависимость  $y$  от фактора  $x$  при элиминировании фактора  $z$  (рис. 12.1). Аналогично можно построить график зависимости  $y$  от  $z$  при элиминировании  $x$  (рис. 12.2).

Аналитическая группировка позволяет не только подтвердить наличие связи и ее направление, но и измерить ее тесноту. Для числовой характеристики тесноты связи могут быть ис-

пользованы показатели вариации результативного признака: общая его дисперсия  $\sigma^2$  и межгрупповая дисперсия:

$$\delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 m_i}{\sum m_i}, \quad (12.3)$$

где  $\bar{y}_i$  — групповые средние;  $m_i$  — число единиц в группах.

Допустим, что группировка произведена так, что признак-фактор в пределах каждой группы совсем не варьирует (например, группировка рабочих по тарифным разрядам и подобные группировки по дискретному признаку без объединения отдельных его значений в одну группу, т. е. без образования интервалов). Если при этом число единиц в каждой группе до-

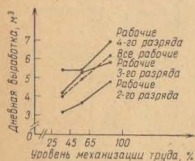


Рис. 12.1. Зависимость средней дневной выработки от уровня механизации



Рис. 12.2. Зависимость средней дневной выработки от квалификации (тарифных разрядов рабочих)

статочно велико, то в силу закона больших чисел влияние на  $y$  прочих факторов будет усреднено, так что вариация  $y$  будет определяться только вариацией  $x$ . Если результативный признак  $y$  совсем не связан с  $x$ , то групповые средние  $\bar{y}_i$  не будут изменяться с изменением  $x$ , т. е. будут равны одна другой и равны общей средней  $\bar{y}$ , а межгрупповая дисперсия  $\delta^2$  будет равна нулю. Наоборот, если результативный признак  $y$  функционально связан с признаком-фактором  $x$ , то в каждой группе внутригрупповая дисперсия  $\sigma_i^2$  будет равна нулю, так как признак  $x$  внутри группы не варьирует. Средняя из групповых дисперсий  $\sigma_i^2$  тоже будет равна нулю, и согласно правилу сложения дисперсий межгрупповая дисперсия  $\delta^2$  совпадает с общей дисперсией  $\sigma^2$ .

В промежуточных случаях, т. е. при корреляционной связи между  $y$  и  $x$ , межгрупповая дисперсия будет находиться между 0 и  $\sigma^2$ , характеризую вариацию  $y$ , объясняемую вариацией  $x$ . Поэтому в качестве показателя тесноты связи между  $y$  и  $x$  при аналитической группировке может быть использовано от-

ношение межгрупповой дисперсии к общей дисперсии результативного признака, называемое эмпирическим коэффициентом детерминации:

$$\eta^2 = \frac{\delta^2}{\sigma^2}. \quad (12.4)$$

Коэффициент детерминации ( $\eta^2$ ) показывает, какая часть общей вариации (дисперсии) результативного признака  $y$  объясняется вариацией группировочного признака-фактора  $x$ . Остальная же часть общей вариации  $y$  объясняется вариацией прочих факторов. При отсутствии связи коэффициент детерминации равен нулю, а при функциональной связи — единице.

Фактически, однако, в большинстве случаев дело осложняется двумя обстоятельствами: 1) при группировке образуются интервалы, так что внутри групп остается вариация признака-фактора и связанная с ней вариация результативного признака; 2) признак-фактор  $x$  может быть корреляционно связан с одним или несколькими прочими факторами, связанными в свою очередь с  $y$ . Вследствие первого обстоятельства межгрупповая дисперсия в таких случаях отражает не всю вариацию, объясняемую фактором  $x$ , а только часть ее. Вследствие же 2-го обстоятельства с изменением от группы к группе признака  $x$  будут изменяться и корреляционно связанные с ним другие факторы, так что межгрупповая дисперсия будет объясняться не только вариацией  $x$ , но и вариацией связанных с ним факторов.

Обычно в качестве показателя тесноты связи используется также эмпирическое корреляционное отношение — корень квадратный из коэффициента детерминации ( $\eta = \sqrt{\frac{\delta^2}{\sigma^2}}$ ). Как и  $\eta^2$ ,

$\eta$  изменяется от 0 до 1: чем ближе оно к 1, тем теснее связь. Вычислим эмпирический коэффициент детерминации по данным табл. 12.4, используя в качестве группировочного признака-фактора разряд рабочих (см. табл. 12.6). Общую дисперсию выработки  $y$  определим по формуле (10.6), т. е. как разность между средним квадратом и квадратом средней (расчет  $\sum y^2$  приведен ниже в табл. 13.1):

$$\sigma_y^2 = \bar{y}^2 - (\bar{y})^2 = \frac{1337,96}{50} - 5^2 = 1,7592.$$

Расчеты, необходимые для вычисления межгрупповой дисперсии, показаны в табл. 12.8.

$$\delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 m_i}{\sum m_i} = \frac{38,1116}{50} = 0,7622.$$

Следовательно, эмпирический коэффициент детерминации и корреляционное отношение равны:

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{0,7622}{1,7592} = 0,4333;$$

$$\eta = \sqrt{0,4333} = 0,658.$$



Таким образом, более 43 % общей вариации выработки объясняется вариацией тарифных разрядов. Связь между этими признаками достаточно тесная. Аналогично может быть измерена теснота связи для отдельных групп комбинационной таблицы, т. е. при более или менее полном элиминировании одного из группировочных признаков.

Отметим, что при группировке с образованием интервалов величин эмпирического коэффициента детерминации и корреляционного отношения зависят от числа групп и их границ. Так, по данным табл. 12.5 межгрупповая дисперсия  $\delta^2 = 0,610$ ;  $\eta^2 = 0,347$ . Если же при группировке рабочих по

Таблица 12.8. К расчету межгрупповой дисперсии

Разряд рабочих	Число рабочих $m_i$	Средняя дневная выработка, $\bar{y}_i$	$\bar{y}_i - \bar{y}$	$(\bar{y}_i - \bar{y})^2$	$(\bar{y}_i - \bar{y})^2 m_i$
2-й	13	3,76	-1,24	1,5376	19,9888
3-й	24	5,03	0,03	0,0009	0,0216
4-й	13	6,18	1,18	1,3924	18,1012
Итого	50	5,00	—	—	38,1116

уровню механизации труда образовать 4 группы (меее 40, 40—60, 60—80, 80—100 %), то  $\delta^2 = 0,812$ ,  $\eta^2 = 0,462$ . Расхождение между этими двумя значениями  $\eta^2$  объясняется тем, что в первом случае вариация уровня механизации внутри групп больше, а соответственно больше и объясняемая ею часть вариации  $y$ , которая не отражается в  $\delta^2$ .

Рассматривая исходные данные как выборку из гипотетической генеральной совокупности, можно оценить значимость полученного коэффициента детерминации и корреляционного отношения с помощью дисперсионного  $F$ -критерия Фишера, оценивающего соотношение межгрупповой ( $\delta^2$ ) и внутригрупповой ( $\sigma^2$ ) дисперсий и рассчитываемого по формуле  $F = \frac{\delta^2}{\sigma^2} \cdot \frac{n-m}{m-1}$ , где  $n$  — число единиц совокупности,  $m$  — количество групп.

Расчетное значение  $F$  сравнивается с табличными (критическими) для принятого уровня значимости  $\alpha$  и чисел степеней свободы  $k_1 = m - 1$  и  $k_2 = n - m$ . В нашем примере зависимости между выработкой рабочих и их квалификацией расчетное значение составит:  $F_{\text{расч}} = \frac{0,7622}{0,997} \cdot \frac{50-3}{3-1} = 17,98$  ( $\sigma_1^2 = \sigma^2 = 0,7622$ ;  $\sigma_2^2 = 0,997$ ;  $n = 50$ ;  $m = 3$ ). Критическое значение Фишера  $\alpha = 0,05$  ( $k_1 = 2$ ;  $k_2 = 47$ ) равно 3,20. Следовательно, связь между выработкой и квалификацией рабочих является существенной ( $F_{\text{расч}} > F_{\text{табл}}$ ).

Основная задача аналитической группировки заключается в том, чтобы обеспечить отчетливое проявление связи (если она есть) путем выравнивания прочих условий в группах. Для этого нужно образовать такое число групп и так установить их границы, чтобы в изменении групповых средних от группы к группе проявилась определенная закономерность, т. е. чтобы эти средние изменялись не хаотично и беспорядочно, а обнаруживали некоторую тенденцию: систематическое возрастание или убыва-

ние с изменением величины группировочного признака. Для этого необходимо, чтобы каждая группа была достаточно однородна и достаточно велика по числу единиц (объему). В то же время, как отмечал А. А. Чупров, число групп должно быть по возможности большим<sup>1</sup>.

При выборе оптимального варианта группировки определенную помощь могут оказать критерии математической статистики. По нашему мнению, при выбранном уровне значимости  $\alpha$  оптимальным можно считать тот вариант, при котором больше разность  $F_{\text{об}} - F_{\text{кр}}$  (при условии, что в изменении групповых средних нет отклонений от определенной тенденции).

Для исследования стохастической связи между качественными (атрибутивными) признаками может быть использована аналитическая группировка. Подобный метод выявления взаимосвязей социально-экономических явлений часто применял в своих работах В. И. Ленин. Анализируя, например, структуру сборов на «правдивские» (марксистские) и «ликвидаторские» газеты, поступивших в годы подъема рабочего движения перед первой мировой войной, В. И. Ленин составил следующую таблицу<sup>2</sup>:

Таблица 12.9

Приходится сборов	Из каждого рубля сборов в газеты:	
	правдивстов	ликвидаторов
от рабочих	87 копеек	44 копейки
не от рабочих	13 копеек	56 копеек
Всего	1 р. 00 копеек	1 р. 00 копеек

Анализ данных таблицы отчетливо показывает наличие связи между социальной принадлежностью лиц, делающих взносы, и направленностью газет: пролетарият России активнее поддерживал марксистские газеты, печать ликвидаторов опиралась в основном на непролетарские слои общества.

Комбинационная группировка по атрибутивным признакам позволяет определить не только направление, но и тесноту связи. Рассмотрим, например, следующие данные о зависимости между степенью механизации труда и годностью 35 изделий в период освоения их выпуска (см. табл. 12.10).

Для сравнения степени механизации труда при изготовлении изделия (факторный признак) и его годности (результативный признак) определим частоты распределения изделий по факторному признаку для каждой группы результативного признака (см. табл. 12.11).

<sup>1</sup> Чупров А. А. Основные проблемы теории корреляции. — М.: Госстатиздат, 1960, с. 11.

<sup>2</sup> Ленин В. И. Полн. собр. соч., т. 25, с. 232.

Частоты распределения изделий в группах с разной степенью годности отличаются. Например, стандартная продукция чаще изготавливалась в условиях автоматизированного труда,

Таблица 12.10

Степень годности изделий	Степень механизации труда			Всего
	частично механизирован	механизирован	автоматизирован	
Соответствует стандартам	1	4	8	13
Исправный брак	3	7	2	12
Неисправный брак	6	3	1	10
Всего	10	14	11	35

а наибольшая доля неисправного брака относится к изделиям, выполненным при частично механизированном труде. Если признаки были независимы, то частоты распределений совпали бы, следовательно, можно предположить наличие связи между

Таблица 12.11

Степень годности изделий	Доля изделий, изготовленных трудом			Всего
	частично механизированным	механизированным	автоматизированным	
Соответствует стандартам	0,077	0,308	0,615	1,000
Исправный брак	0,250	0,583	0,167	1,000
Неисправный брак	0,600	0,300	0,100	1,000
Всего	0,286	0,400	0,314	1,000

признаками. Для измерения ее тесноты воспользуемся коэффициентом взаимной сопряженности А. А. Чупрова:

$$C = \sqrt{\frac{\chi^2}{(k_1 - 1)(k_2 - 1) \sum m_i}}, \quad (12.5)$$

$$\chi^2 = \sum_{i=1}^{k_1} m_i \sum_{j=1}^{k_2} \frac{(w_{ij} - w_{.j})^2}{w_{.j}},$$

где  $i$  — номер группы по годности изделий;  $j$  — номер группы по степени механизации труда;  $k_1$  — число групп по годности изделий;  $k_2$  — число групп по степени механизации труда;  $w_{ij}$  —

частоты подгруппы  $j$  в группе  $i$ ;  $w_{.j}$  — частоты группы  $j$  во всей совокупности;  $m_i$  — число единиц в группе  $i$ .

При независимости признаков  $w_{ij} = w_{.j}$ ;  $\chi^2 = 0$ ,  $C = 0$ . В случае функциональной связи  $C = 1$ . В нашем примере

$$\begin{aligned} \chi^2 &= 13 \left[ \frac{(0,077 - 0,286)^2}{0,286} + \frac{(0,308 - 0,400)^2}{0,400} + \frac{(0,615 - 0,314)^2}{0,314} \right] + \\ &+ 12 \left[ \frac{(0,250 - 0,286)^2}{0,286} + \frac{(0,583 - 0,400)^2}{0,400} + \frac{(0,167 - 0,314)^2}{0,314} \right] + \\ &+ 10 \left[ \frac{(0,600 - 0,286)^2}{0,286} + \frac{(0,300 - 0,400)^2}{0,400} + \frac{(0,100 - 0,314)^2}{0,314} \right] = \\ &= 13,052. \text{ Откуда } C = \sqrt{\frac{13,052}{35(3-1)(3-1)}} = 0,305. \end{aligned}$$

Таким образом, между степенью механизации труда и качеством изготавливаемых изделий существует определенная зависимость, но теснота ее довольно умеренна.

В практике статистических исследований нередко приходится анализировать альтернативные распределения, когда совокупность распределяется по каждому признаку на две группы с противоположными характеристиками. Например, можно анализировать успеваемость студентов, расценивая контингент по полу и в каждой группе выделять две подгруппы: сдавшие и не сдавшие экзамен. В этом случае строится четырехклеточная таблица, имеющая следующий вид:

Таблица 12.12. Зависимость успеваемости студентов от пола

	Контингент студентов		Всего
	сдавших экзамен	не сдавших экзамен	
Женщины	$a = 25$	$b = 2$	$a + b = 27$
Мужчины	$c = 20$	$d = 3$	$c + d = 23$
Итого	$a + c = 45$	$b + d = 5$	50

Тесноту связи в этом случае можно оценить, вычислив коэффициент ассоциации (ингода его называют коэффициентом четырехклеточной корреляции), равный коэффициенту  $C$ :

$$A = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{25 \cdot 3 - 2 \cdot 20}{\sqrt{27 \cdot 23 \cdot 45 \cdot 5}} = 0,09.$$

Следовательно, между полом студента и его успеваемостью связь крайне незначительна, практически она отсутствует. Полученный вывод является, несомненно, справедливым, ведь



истинными факторами успеваемости выступают, конечно же, не пол студентов, а посещаемость и успешная работа на всех видах аудиторных занятий, эффективность самостоятельной работы во внеаудиторное время.

## ГЛАВА 13 АНАЛИЗ РЕГРЕССИИ И КОРРЕЛЯЦИИ

### 13.1. РЕГРЕССИОННО-КОРРЕЛЯЦИОННЫЙ АНАЛИЗ И ЕГО ЭТАПЫ

Для более глубокого исследования взаимосвязи социально-экономических явлений рассмотренные статистические методы часто оказываются недостаточными, ибо они не позволяют выразить имеющуюся связь в виде определенного математического уравнения, характеризующего механизм взаимодействия факторных и результативных признаков. Кроме того, методы параллельных рядов и аналитических группировок эффективны лишь при малом числе факторных признаков, в то время как социально-экономические явления складываются обычно под воздействием множества причин. Эти и другие ограничения рассмотренных ранее статистических методов анализа взаимосвязей устраняет *метод анализа регрессий и корреляций* — *регрессионо-корреляционный анализ* (РКА), являющийся логическим продолжением, углублением более элементарных методов.

РКА заключается в построении и анализе *экономико-математической модели* в виде *уравнения регрессии* (корреляционной связи), выражающего зависимость явления от определяющих его факторов:  $\hat{y}_x = f(x_1, x_2, \dots, x_n)$ . Если, например, зависимость производительности труда ( $y$  — в рублях) от уровня механизации производства ( $x_1$  — в процентах) и квалификации рабочего ( $x_2$  — разрядов) может быть представлена в виде уравнения  $\hat{y}_x = 23,12 + 2,06x_1 + 1,41x_2$ , то, при прочих равных условиях, можно утверждать, что с увеличением уровня механизации на один процент производительность возрастет в среднем на 2,06 рубля, а с ростом квалификации на один разряд — увеличится в среднем на 1,41 рубля. Следовательно, между уровнем механизации ( $x_1$ ) и квалификацией рабочего ( $x_2$ ), с одной стороны, и производительностью труда как результативным признаком, с другой — существует прямая связь.

Успешное проведение регрессионо-корреляционного анализа, как и других статистических методов изучения взаимосвязей, в решающей степени определяется глубиной теоретического качественного исследования сущности рассматриваемого явления, позволяющего установить наличие или отсутствие при-

чинно-следственных отношений между признаками. Математическая статистика и теория вероятностей являются тем аппаратом, который лишь позволяет исходя из информации о данном явлении построить его математическую модель, т. е. математико-статистическое описание, дающее возможность оценить величину явления при определенных величинах факторных признаков. Однако определение вида моделей и факторов, которые следует в нее включить, должно опосредоваться на строгом экономическом анализе, без которого регрессионо-корреляционный анализ будет лишь формальным математическим упражнением.

РКА состоит из следующих этапов (стадий): 1) предварительный (априорный) анализ; 2) сбор информации и ее первичная обработка; 3) построение модели (уравнения регрессии); 4) оценка и анализ модели. Подобное деление на этапы весьма условно, так как отдельные стадии тесно связаны между собой и нередко, как мы позднее увидим, результат, полученный на одном этапе, позволяет дополнить, скорректировать выводы более ранних стадий РКА.

1. Под *предварительным (априорным) анализом* понимается весь процесс исследования рассматриваемого явления до сбора исходной информации. Важность этого этапа состоит в том, что здесь формируются основные направления всего регрессионо-корреляционного анализа.

Прежде всего необходимо в самом общем виде сформулировать задачу исследования. Например, изучение влияния различных факторов на уровень производительности труда. Далее следует определить методику измерения результативного показателя, т. е. выбрать измеритель, который наилучшим образом характеризует этот признак с точки зрения задачи исследования. Так, если результативным показателем является производительность труда, то необходимо выяснить, каким методом (натуральным, трудовым или стоимостным) целесообразнее всего измерять ее в конкретных условиях данного производства.

На стадии априорного анализа решается также вопрос о числе факторов, оказывающих наиболее существенное влияние на формирование результативного показателя и о способах их измерения. Учитывая важность этого момента целесообразно проводить двухстадийный отбор факторов: на первой стадии отбираются все факторы, связанные с результативным показателем, а затем, в процессе количественного анализа, этот перечень конкретизируется и уточняется. Значительную помощь в предварительном формировании перечня факторов, наиболее существенно влияющих на результативный признак, оказывает метод их экспертной оценки, основанный на привлечении специалистов в исследуемой области.

При формировании предварительного перечня факторных признаков необходимо учитывать следующие обстоятельства.

Во-первых, факторные признаки должны быть количественными, лучше всего — непрерывными. Если же в данном исследовании трудно обойтись без атрибутивных признаков (профессия, пол, образование), то целесообразно, как показывает практика, искусственно привести их к количественному виду путем ранжирования.

Во-вторых, факторные признаки должны рассчитываться по отношению к одной базе. Если, например, в качестве результативного показателя исследуется производительность труда одного рабочего, то такие факторы, как машиностроительность и электровооруженность труда, заработная плата и др. следует исчислять в расчете на одного рабочего. В-третьих, не рекомендуется включать в модель факторы, которые взаимно дублируют друг друга, т. е. отражают одну и ту же сторону изучаемого явления. Нарушение этого правила приводит к тому, что истинное влияние каждого из этих факторов на результативный показатель искажается.

2. Результаты регрессионно-корреляционного анализа, как и любого статистического исследования, в значительной мере зависят от качества исходной информации. Поэтому необходимо, чтобы информационная база отвечала определенным требованиям.

Прежде всего совокупность должна быть достаточно большой по объему (числу единиц или наблюдений), чтобы в силу действия закона больших чисел определяемые в процессе РКА статистические характеристики были достаточно типичными и надежными.

Исходные данные должны быть качественно и количественно однородны. Качественная однородность предполагает близость условий формирования результативных и факторных признаков. Так, при изучении производительности труда в промышленности необходима однородность предприятий по виду продукции, применяемой технологии, уровню специализации и кооперированию. Количественная однородность заключается в отсутствии единиц наблюдения, которые по своим числовым характеристикам существенно отличаются от основной массы данных. Подобные аномальные наблюдения следует исключить из совокупности.

Основным методом обеспечения однородности является группировка. Однако в некоторых случаях группировка может привести к тому, что число единиц в каждой группе окажется слишком малым, т. е. нарушится требование о достаточно большом объеме данных для РКА. В подобных случаях для увеличения числа наблюдений нередко используется метод объектов-периодов. Сущность его заключается в том, что данные по каждому объекту совокупности (заводу, колхозу), взятые за ряд лет, рассматриваются как самостоятельные наблюдения. Использование этого метода позволяет довести совокупность до необходимого числа единиц, которое, как показывает прак-

тика, должно быть в 6—8 раз больше количества включаемых в модель факторов.

Кроме этих условий, характерных для любого статистического исследования, РКА предъявляет к качеству исходной информации ряд специфических требований. Во-первых, наблюдения ( $y, x_1, x_2, \dots, x_n$ ) должны быть статистически независимыми, т. е. значения признаков у определенной единицы совокупности не должны зависеть от их значений у других единиц данной совокупности. Во-вторых, переменные  $x_1, x_2, \dots, x_n$  должны быть линейно независимы, т. е. включаемые в модель факторы не должны находиться между собой в тесной взаимной связи. Нарушение этого положения приводит к тому, что коэффициенты регрессии и другие статистические характеристики являются неустойчивыми, что затрудняет их экономическую интерпретацию. Поэтому целесообразно, как показывает практика, исключить из модели один или несколько тесно связанных с другими факторов. В-третьих, каждому значению факторного признака ( $x$ ) должно соответствовать нормальное распределение результативного признака ( $y$ ) с одинаковой дисперсией.

В социально-экономических исследованиях, в отличие от естественнонаучных, для которых первоначально и разрабатывался РКА, эти специфические требования к исходной информации не всегда строго выполняются. В частности, крайне редко соблюдается условие нормального распределения результативного признака, факторы-аргументы обычно связаны между собой в силу взаимобусловленности экономических показателей, а использование метода объектов-периодов нередко нарушает условие статистической независимости: ведь данные по конкретному объекту, конечно же, зависят от их уровня в предшествующие годы. Это вызывает необходимость тщательной проверки выполнения рассмотренных условий, которая производится с помощью критериев математической статистики ( $F$ -критерий Фишера,  $t$ -критерий Стьюдента и др.). Практика показывает, что незначительные отклонения от этих условий не являются препятствием к проведению РКА. Существенные же отклонения указывают на необходимость пересмотра информационной базы исследования.

3. При построении регрессионно-корреляционной модели (уравнения регрессии) прежде всего возникает вопрос о типе аналитической функции, характеризующей механизм взаимосвязи между результативным признаком и одним или несколькими признаками-факторами. Так, применительно к одному факторному признаку речь идет о том, будет ли геометрическим изображением этой аналитической функции прямая линия ( $\hat{y}_x = a_0 + a_1x$ ), парабола второго порядка ( $\hat{y}_x = a_0 + a_1x + a_2x^2$ ), двучленная гиперболола ( $\hat{y}_x = a_0 + \frac{a_1}{x}$ ) или какая-либо иная кривая. Иными словами, возникает вопрос о выборе (установлении) формы связи. Выбор той или иной формы связи означает выдвижение и принятие некоторой теоретической обоснованной или практически приемлемой рабочей гипотезы о механизме взаимодействия изучаемых признаков. Поэтому этот этап РКА является очень важным и ответственным.



После того как выбрана форма связи и уравнение регрессии поставлено в общем виде, нужно на основе исходной информации найти числовые значения параметров уравнения регрессии ( $a_0$ ,  $a_1$  и т. д.).

### 13.2. ПОСТРОЕНИЕ И ОЦЕНКА ОДНОФАКТОРНЫХ МОДЕЛЕЙ

Выбор формы связи между результативным признаком и признаком-фактором должен быть основан на качественном, теоретическом и логическом анализе природы изучаемых явлений, их социально-экономической сущности. Поэтому прежде всего нужно попытаться установить форму связи путем *экономического и логического анализа*. При построении однофакторных моделей (уравнений парной регрессии) такой анализ в ряде случаев позволяет вывести уравнение регрессии, приближенно характеризующей зависимость результативного признака от признака-фактора. Если бы зависимость  $y$  от  $x$  была функциональной, то теоретический анализ дал бы возможность представить  $y$  как строгую и точную функцию от  $x$ . Однако и при корреляционной связи можно нередко представить зависимость  $y$  от  $x$  аналитически в виде приближенной формулы, если абстрагироваться от того обстоятельства, что, кроме данного фактора, на  $y$  влияют и многие другие факторы. Это абстрагирование приведет, конечно, к некоторому упрощению действительного механизма взаимосвязи, однако оно позволит сконцентрировать внимание на зависимости  $y$  от  $x$ .

Рассмотрим, например, вопрос о форме связи между дневной выработкой рабочего на формовке бетона ( $y$  — куб. м) и уровнем механизации труда ( $x$  — доля времени, отработанного в течение смены на механизированной формовке). Доля времени, отработанного в течение смены вручную, выразится тогда как  $(1-x)$ . Обозначим дневную выработку рабочего при выполнении работы механизированным способом через  $y_m$ , а при формовке вручную — через  $y_p$ . Тогда уровень дневной выработки бетонщика, проработавшего часть смены  $x$  на механизированной формовке, а часть  $(1-x)$  — на ручной, может быть выражен как средняя арифметическая взвешенная из дневной выработки при механизированной работе ( $y_m$ ) и работе вручную ( $y_p$ ) с весами  $x$  и  $(1-x)$ :

$$y = y_m x + y_p (1-x) = y_m x + y_p - y_p x = y_p + (y_m - y_p) x.$$

Считая  $y_p$  и  $y_m$  величинами постоянными (средними за смену), мы видим, что зависимость дневной выработки бетонщика от уровня механизации его труда может быть выражена уравнением прямой:  $y = A_0 + A_1 x$ , где  $A_0 = y_p$ ,  $A_1 = y_m - y_p$ . Так как  $y_m > y_p$ , то  $A_1 > 0$ . Это означает, что с возрастанием уровня механизации труда  $x$  дневная выработка бетонщика увеличивается равномерно (по прямой линии).

Естественно, что у каждого бетонщика могут быть различ-

ные уровни выработки как на механизированной  $y_m$ , так и на ручной формовке  $y_p$ , что связано с влиянием других факторов (квалификация, производственный стаж и др.). Поэтому коэффициенты  $A_0$  и  $A_1$  могут быть различными для отдельных рабочих, и при этом и том же уровне механизации труда уровень дневной выработки у отдельных бетонщиков неодинаков. Чтобы поставить дневную выработку бетонщиков ( $y$ ) в зависимость только от уровня механизации труда ( $x$ ) и исключить влияние прочих факторов, необходимо определить типичные значения коэффициентов полученного уравнения, в данном случае средний для всех бетонщиков уровень дневной выработки на механизированной и на ручной формовке бетона. Обозначим  $A_0$  через  $a_0$  и  $A_1$  через  $a_1$ . Тогда получится следующее *уравнение линейной регрессии*:

$$\hat{y}_x = a_0 + a_1 x. \quad (13.1)$$

Нужно иметь в виду, что уравнение регрессии правильно выражает форму связи лишь при условии независимости коэффициентов  $a_0$  и  $a_1$  от факторного признака  $x$  либо такой незначительной зависимости, которой можно пренебречь.

Уравнение похоже и непохоже на уравнения функциональной зависимости. Оно похоже на них своей формой, но отличается по своему смыслу. Объясняется это тем, что уравнение регрессии с осредненными коэффициентами отражает соотношение признаков только в общем итоге по изучаемой совокупности. Следовательно, уравнение регрессии показывает типичное для данной совокупности соотношение между размерами факторного и результативного признаков (тогда как уравнение функциональной зависимости было бы справедливо для каждого отдельного случая).

В ряде случаев теоретический анализ приводит к выводу *уравнения криволинейной регрессии*. Рассмотрим, например, зависимость себестоимости единицы продукции от объема производства этой продукции. Себестоимость единицы продукции рассчитывается путем деления общей суммы затрат на производство данной продукции на ее количество, объем. Поэтому общая сумма затрат на производство (общая себестоимость всей продукции) равна произведению себестоимости единицы продукции ( $y$ ) на объем продукции ( $x$ ), т. е.  $yx$ . В то же время затраты на производство можно условно и приближенно подразделить на две части: 1) затраты, которые возрастают более или менее пропорционально увеличению объема произведенной продукции, — условно-переменные расходы (затраты на сырье и материалы, на топливо и электроэнергию для технологических целей, оплата труда основных производственных рабочих и т. п.); 2) затраты либо совершенно не зависящие от объема продукции, либо зависящие от него в незначительной степени, — условно-постоянные расходы (оплата труда инженер-

но-технических работников и служащих, расходы на содержание зданий и сооружений и другие административно-управленческие и общехозяйственные расходы).

Обозначим переменные расходы в расчете на единицу продукции через  $a_0$ ; тогда их общая сумма составит  $a_0x$ . Общую сумму условно-постоянных расходов обозначим  $a_1$ . Тогда общая себестоимость продукции составит  $y_x = a_0x + a_1$ , откуда себестоимость единицы продукции будет равна:

$$\hat{y}_x = a_0 + \frac{a_1}{x} \quad (13.2)$$

Для каждого предприятия, выпускающего данную продукцию, расходы  $a_0$  и  $a_1$  имеют различную величину и здесь взяты их типичные (средние) значения. Мы получили уравнение двучленной гиперболы. Поскольку  $a_1 > 0$ , то зависимость себестоимости единицы продукции от объема ее выпуска является обратной: с увеличением  $x$  средняя себестоимость снижается. Однако это снижение не является равномерным, так как по мере увеличения  $x$  снижение  $\hat{y}_x$  постепенно замедляется.

В некоторых случаях теоретический и логический анализ показывает, что неравномерное изменение результативного признака должно иметь иной характер. Так, при недостаточном количестве осадков урожайность будет, естественно, очень низкая, а по мере его увеличения она будет повышаться. Однако это повышение не будет беспредельным, так как для каждой культуры в данных конкретных условиях есть какое-то оптимальное количество осадков, при котором достигается наиболее высокая урожайность. По мере того как количество осадков будет приближаться к оптимальной величине, рост урожайности будет постепенно замедляться и прекратится совсем по достижении этого оптимума. Дальнейшее увеличение количества осадков может привести к снижению урожайности. Такого рода зависимость можно приближенно выразить уравнением *параболы второго порядка* ( $\hat{y}_x = a_0 + a_1x + a_2x^2$ ). Аналогичный характер связи можно ожидать и в ряде других случаев, например для зависимости уровня производительности труда рабочего от его возраста. С увеличением возраста и стажа работы накапливается опыт работы, совершенствуются навыки, повышается квалификация и производительность труда. Но для каждого вида работы есть, очевидно, некоторый определенный возраст, при котором выработка рабочего на единицу времени будет наибольшей. В дальнейшем благоприятное влияние производственного опыта и навыков перекрывается неблагоприятным влиянием возрастных особенностей пожилого человека, в результате чего выработка будет снижаться.

В некоторых случаях теоретический анализ дает основания предполагать, что относительное изменение результативного признака должно быть пропорционально относительному при-

росту признака-фактора. Например, увеличение  $x$  на 1 % должно приводить к увеличению или уменьшению  $y$  на  $a$  процентов. В таких случаях в качестве формы связи можно принять *степенную функцию* вида  $y_x = a \cdot x^a$ .

Социально-экономические явления очень сложны. Как, правдо, мы не имеем о них исчерпывающей информации, а внутреннего логика их связей мало изучена. Факторы, влияющие на то или иное явление, взаимно переплетаются и взаимодействуют друг с другом. Поэтому очень часто не удается сделать теоретически обоснованный вывод уравнения регрессии, т. е. формы связи, внутренне присущей изучаемому явлению. В ряде случаев на основе теоретического анализа могут быть высказаны лишь более или менее обоснованные предположения о том, следует ли ожидать линейную или какую-либо нелинейную (криволинейную) связь, имеет ли ожидаемая криволинейная функция экстремальные значения, асимптоты и т. п. Более того, если явление мало изучено, иногда могут быть выдвинуты и различные гипотезы о механизме и форме взаимосвязи.

Для проверки тех или иных предположений и гипотез может быть использован графический метод — построение *графика групповых средних*, полученных в процессе аналитической группировки (см. рис. 12.1 и 12.2). Ломаная линия, изображающая изменение групповых средних результативного признака в зависимости от изменения группировочного признака-фактора, называется *эмпирической линией регрессии* (эмпирической регрессией). Форма эмпирической регрессии дает возможность проверить, соответствует ли фактическое соотношение признаков тому или иному теоретически предполагаемому их соотношению. При этом нужно, однако, иметь в виду, что при относительно небольшом числе единиц совокупности (числе наблюдений) форма эмпирической линии регрессии может изменяться при изменении числа групп и их границ. Поэтому при небольшом числе наблюдений нельзя слишком полагаться на форму эмпирической регрессии, графический метод в таких случаях может оказаться недостаточно надежным.

Если относительно формы связи могут быть выдвинуты различные теоретические гипотезы, а по виду эмпирической регрессии трудно судить о том, какой из этих гипотез наиболее соответствуют фактические данные, то в этих случаях строятся и решаются уравнения регрессии с различными формами связи, а затем с помощью специальных статистико-математических критериев оценивается их адекватность и выбирается та форма связи, которая обеспечивает наилучшую аппроксимацию (приближение) и достаточную статистическую достоверность и надежность.

Выбрав тем или иным путем форму связи и построив уравнение регрессии в общем виде, необходимо далее найти числовые значения его *параметров*. Напомним, что вследствие влияния прочих факторов параметров уравнения для отдельных



единиц совокупности могут быть различными, так что задача заключается в том, чтобы найти их типические, средние для всей совокупности значения. При этом в отличие от обычного порядка нахождения средних величин значения параметров у отдельных единиц совокупности здесь не известны. Кроме того, процесс осреднения параметров уравнения есть в то же время процесс осреднения значений результативного признака, причем эта средняя ( $\bar{y}_x$ ) является не постоянной, а переменной величиной, меняющейся в связи с изменением факторного признака ( $x$ ). Таким образом, процесс осреднения при построении уравнения регрессии носит особый характер.

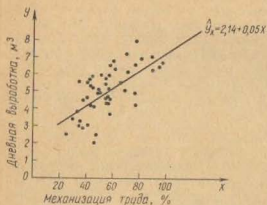


Рис. 13.1. Зависимость дневной выработки от уровня механизации труда (корреляционное поле и уравнение регрессии)

ной регрессии это означает, что минимальной должна быть  $\sum (y - \hat{y}_x)^2$ . Заменяя затем  $\hat{y}_x$  выражением, соответствующим выбранной форме связи [например, при линейной связи  $\hat{y}_x = a_0 + a_1 x$  будем иметь  $\sum (y - a_0 - a_1 x)^2$ ], и рассматривая полученную сумму квадратов отклонений как функцию параметров, можно найти те их значения, при которых эта сумма будет минимальной. Это достигается путем приравнивания нулю первых частных производных по каждому параметру и решения полученной при этом системы так называемых нормальных уравнений. Такой способ расчета параметров называется **методом наименьших квадратов**.

Если исходные данные (значения  $x$  и  $y$  (см. табл. 12.4)) нанести на график в виде точек в прямоугольной системе координат, то получим **поле корреляции** (рис. 13.1). Если бы зависимость  $y$  от  $x$  была функциональной, то все точки были бы расположены на какой-то линии. При корреляционной связи вследствие влияния прочих факторов точки не лежат на одной линии, но все же их расположение обнаруживает определенную тенденцию. Так, на рис. 13.1 видно, что по мере повышения

уровня механизации труда ( $x$ ) дневная выработка рабочего ( $y$ ) в общем и целом повышается. Как показал сделанный ранее теоретический анализ, зависимость здесь должна выражаться уравнением прямой  $y_x = a_0 + a_1 x$ , где параметр  $a_0$  характеризует средний для всех бетонщиков уровень дневной выработки на ручной формовке, а параметр  $a_1$  — превышение средней дневной выработки на механизированной формовке над средней выработкой при работе вручную (при условии, что уровень механизации труда ( $x$ ) выражен в виде доли).

Расчет параметров прямой методом наименьших квадратов дает такие их значения, при которых прямая на графике пройдет наиболее близко к точкам, изображающим исходные фактические данные. Наибольшая близость в данном случае означает, что если в уравнение прямой последовательно подставлять фактические значения  $x$  у каждой единицы совокупности и рассчитывать соответствующие значения результативного признака ( $\hat{y}_x$ ), то сумма квадратов отклонений фактических значений  $y$  от расчетных, т. е.  $\sum (y - \hat{y}_x)^2$ , будет при этом меньше, чем для любой другой прямой, которую можно провести на корреляционном поле.

Аналогично обстоит дело и при других (криволинейных) формах связи: расчет параметров кривой методом наименьших квадратов дает возможность найти ту кривую, которая по сравнению со всеми другими кривыми данного вида проходит наиболее близко к точкам корреляционного поля, т. е. дает наименьшую сумму квадратов отклонений  $[\sum (y - \hat{y}_x)^2]$ .

Рассмотрим расчет параметров для **линейной парной регрессии**  $\hat{y}_x = a_0 + a_1 x$ . Сумма квадратов отклонений  $y$  от  $\hat{y}_x$  составляет  $S = \sum (y - \hat{y}_x)^2 = \sum (y - a_0 - a_1 x)^2$ . В математической статистике доказывается, что система нормальных уравнений для парной линейной регрессии такова:

$$\left. \begin{aligned} a_0 n + a_1 \sum x &= \sum y; \\ a_0 \sum x + a_1 \sum x^2 &= \sum xy. \end{aligned} \right\}$$

Решение этой системы в общем виде дает следующие значения параметров:

$$a_0 = \frac{\sum y \cdot \sum x^2 - \sum x \cdot \sum xy}{n \sum x^2 - \sum x \cdot \sum x}; \quad (13.3a)$$

$$a_1 = \frac{n \sum xy - \sum y \cdot \sum x}{n \sum x^2 - \sum x \cdot \sum x}. \quad (13.3b)$$

Рассчитаем параметры линейного уравнения регрессии, выражающего зависимость дневной выработки рабочего на формовке бетона ( $y$ ) от уровня механизации труда ( $x$ ) по табл. 12.4. Составим расчетную таблицу ( $y^2$ ,  $x^2$  и другие будут использованы в дальнейших расчетах).

ние вычислений в этом случае достигается путем комбинационной группировки по факторному и результированному признакам, однако точность расчета при этом снижается, так как вместо индивидуальных значений признаков используются середины интервалов.

Следует отметить, что расчет параметров по сгруппированным данным как способ упрощения построения модели в настоящее время теряет свое значение. Наличие быстродействующих ЭВМ с пакетами стандартных программ позволяет достаточно быстро оценить параметры модели и на основе первичных данных.

Выше было показано, что зависимость себестоимости единицы продукции от объема ее производства может быть приближенно выражена уравнением двучленной гиперболической регрессии вида  $\hat{y}_x = a_0 + \frac{a_1}{x}$ . Она отличается от уравнений линейной регрессии только тем, что вместо величины  $x$  в нем содержится обратная величина  $\frac{1}{x}$ . Произведя такую же замену в формуле (13.3), получим следующие формулы для расчета параметров такой гиперболы:

$$a_0 = \frac{\sum y \cdot \frac{1}{x^2} - \sum \frac{y}{x} \cdot \sum \frac{1}{x}}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \cdot \sum \frac{1}{x}};$$

$$a_1 = \frac{n \sum \frac{y}{x} - \sum y \cdot \sum \frac{1}{x}}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \cdot \sum \frac{1}{x}}. \quad (13.4 \text{ а, б})$$

Например, по 20 заводам имеются данные о выпуске продукции и себестоимости ее единицы, изображенные на рис. 13.2 в виде точек. На графике построена также эмпирическая регрессия, форма которой подтверждает, что исходные данные не противоречат предположению о гиперболической форме связи.

Таблица 13.2. Расчет сумм для определения параметров гиперболы

№ п/п	Объем производства, тыс. ед. $x$	Себестоимость единицы, руб. $y$	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{y}{x}$	$y^2$
1	5,0	30	0,200	0,0400	6,00	900
2	5,5	28	0,182	0,0331	6,09	784
3	5,6	29	0,179	0,0319	5,18	841
...	...	...	...	...	...	...
19	10,0	23	0,100	0,0100	2,30	529
20	11,0	24	0,091	0,0083	2,18	576
Итого	×	515	2,748	0,3939	71,71	13 341

Для расчета нужных сумм построим таблицу, в которую включим также необходимые в дальнейшем значения  $y^2$  (табл. 13.2).

Подставив в формулу (13.4) вычисленные суммы, найдем, что  $a_0 = 17,76$  и  $a_1 = 58,13$ , откуда искомое уравнение регрессии таково:  $\hat{y} = 17,76 + \frac{58,13}{x}$ . На рис. 13.2 показан график этого уравнения.

Если бы в расчетах объем выпуска ( $x$ ) выражался в единицах, а не в тысячах единиц, параметр  $a_0$  остался бы таким же, а параметр  $a_1$  был бы в 1000 раз больше. Следовательно, в среднем по 20 заводам затраты, пропорциональные объему продукции, составляют 17,76 руб. на каждую единицу продукции, а общая сумма условно-постоянных расходов равна 58 130 руб. Чем больше объем продукции, тем меньшая часть этой суммы падает на каждую единицу продукции, в результате чего зависимость себестоимости единицы продукции от ее объема и оказывается обратной.

Если форма связи выражается параболой второго порядка  $y_x = a_0 + a_1x + a_2x^2$ , то для определения ее параметров, удовлетворяющих требованиям метода наименьших квадратов, нужно решить следующую систему нормальных уравнений:

$$\left. \begin{aligned} a_0n + a_1\sum x + a_2\sum x^2 &= \sum y; \\ a_0\sum x + a_1\sum x^2 + a_2\sum x^3 &= \sum xy; \\ a_0\sum x^2 + a_1\sum x^3 + a_2\sum x^4 &= \sum x^2y. \end{aligned} \right\} \quad (13.5)$$

Допустим, по 30 колхозам, находящимся в примерно одинаковых почвенно-климатических условиях, имеются данные об общем количестве внесенных на 1 га удобрений (в пересчете на действующие вещества) и урожайности зерновых. Эти данные изображены на рис. 13.3 в виде точек корреляционной поля.

Как видно из графика, по мере увеличения количества внесенных на каждый гектар удобрений урожайность имеет тенденцию к повышению. Эта тенденция подтверждается и изменением групповых средних (табл. 13.3).

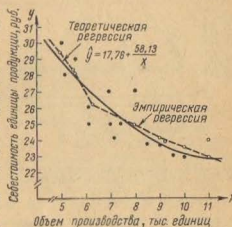


Рис. 13.2. Зависимость себестоимости единицы продукции от объема производства



Однако, как видно из последней графы таблицы, прибавка урожайности не пропорциональна приросту количества удобрений: при относительно невысоких нормах внесения удобрений эта прибавка выше, а при высоких нормах — ниже.

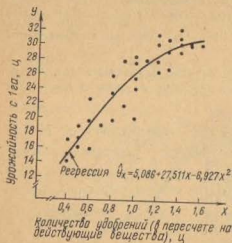


Рис. 13.3. Зависимость урожайности от количества удобрений, внесенных на 1 га

Таблица 13.3. Зависимость урожайности от количества удобрений

Внесено удобрений на 1 га, ц	Число поделков	Средняя урожайность, ц с 1 га	Увеличение урожайности по сравнению с предыдущей группой, ц
Менее 0,65	6	18,0	...
0,65—0,95	6	22,8	4,8
0,95—1,25	10	26,7	3,9
1,25—1,55	8	29,8	3,1
Итого	30	25,0	—

Решение этой системы уравнений дает следующие значения параметров:  $a_0 = 5,086$ ;  $a_1 = 27,511$ ;  $a_2 = -6,927$ . Следовательно, уравнение регрессии имеет вид

$$\hat{y}_x = 5,086 + 27,511x - 6,927x^2.$$

Теоретический и логический анализ также приводит к выводу о том, что по мере увеличения норм внесения удобрений их эффект не будет оставаться постоянным, так как в каждом конкретном случае существует та или иная оптимальная норма удобрений, обеспечивающая максимальную урожайность. Следовательно, в порядке первого приближения можно считать, что уравнение регрессии в нашем случае будет представлять ветвь параболы второго порядка.

Обозначив количество удобрений через  $x$ , определим параметры параболы по исходным данным нашего примера, для чего сначала найдем нужные суммы. (См. табл. 13.4).

Подставив найденные суммы в систему нормальных уравнений (13.5), получим:

$$\left. \begin{aligned} 30a_0 + 30a_1 + 32,9a_2 &= 750; \\ 30a_0 + 32,9a_1 + 38,484a_2 &= 791,1; \\ 32,9a_0 + 38,484a_1 + 47,0762a_2 &= 899,95. \end{aligned} \right\}$$

Подставляя в это уравнение соответствующие значения  $x$ , найдем расчетные значения  $\hat{y}_x$ . Так, при  $x = 0,4$  получим  $\hat{y} = 5,086 + 27,511 \cdot 0,4 - 6,927 \cdot (0,4)^2 = 14,98$ ;  $x = 0,5$  — соответственно 17,11 и т. д. (см. последнюю графу табл. 13.4). См. рис. 13.3.

Значения  $\hat{y}_x$ , полученные по уравнению регрессии, показывают, какую в среднем величину урожайности можно ожидать

Таблица 13.4. Расчет сумм для определения параметров параболы второго порядка

Номера колхозов	$x$	$y$	$x^2$	$x^3$	$x^4$	$yx$	$yx^2$	$\hat{y}_x$
1	0,4	14	0,16	0,064	0,0256	5,6	2,24	14,98
2	0,5	16	0,25	0,125	0,0625	8,0	4,00	17,11
3	0,5	19	0,25	0,125	0,0625	9,5	4,75	17,11
...	...	...	...	...	...	...	...	...
29	1,4	32	1,96	2,744	3,8416	44,8	62,72	30,02
30	1,5	30	2,25	3,375	5,0625	45,0	67,50	30,77

Итого 30,0 750 32,90 38,484 47,0762 791,1 899,95 749,95

при данном количестве удобрений и прочих равных (средних для данной совокупности) условиях. Из графика и приведенных значений  $\hat{y}_x$  видно, что повышение урожайности за счет увеличения количества удобрений происходит неравномерно: равные изменения факторного признака  $x$  вызывают неравные, меняющиеся для разных уровней  $x$  изменения  $\hat{y}_x$ . Так, при увеличении  $x$  с 0,4 до 0,5 (см. табл. 13.4) ожидаемая прибавка урожайности составляет 2,13 ц (17,11—14,98), а при увеличении с 1,4 до 1,5 — только 0,75 ц (30,77—30,02). Скорость изменения урожайности является здесь переменной, она постепенно замедляется при увеличении  $x$ . Размер же этого замедления — величина постоянная: каждое новое увеличение  $x$  на 0,1 дает прибавку урожайности, которая на 0,14 ц меньше предыдущей [ $2a_2 \cdot (0,1)^2$ ].

Для определения параметров уравнения регрессии, выраженного степенной функцией вида  $\hat{y}_x = a_0 x^{a_1}$ , нужно привести эту функцию к линейному виду путем логарифмирования:  $\lg \hat{y}_x = \lg a_0 + a_1 \lg x$ . Полученное уравнение отличается от уравнения обычной линейной регрессии тем, что вместо  $y$ ,  $x$  и  $a_0$  содержит их логарифмы. Поэтому формулы параметров степенной регрессии можно получить, заменив в формуле (13.3) показатели  $y$ ,  $x$  и  $a_0$  их логарифмами (величина  $a_1$  логарифмом не заменяется):

$$\lg a_0 = \frac{\sum \lg y \cdot \sum (\lg x)^2 - \sum \lg y \cdot \lg x \cdot \sum \lg x}{n \sum (\lg x)^2 - \sum \lg x \cdot \sum \lg x}; \quad (13.6a)$$

$$a_1 = \frac{n \sum \lg y \cdot \lg x - \sum \lg y \cdot \sum \lg x}{n \sum (\lg x)^2 - \sum \lg x \cdot \sum \lg x}. \quad (13.6 б)$$

Параметр  $a_1$  здесь показывает, на сколько процентов изменяется в среднем  $y$  при изменении  $x$  на один процент.

После построения модели (уравнения регрессии) ее необходимо оценить и проанализировать. Прежде всего нужно проверить, согласуются ли знаки параметров с теоретическими представлениями и соображениями о направлении влияния признака-фактора на результативный признак (показатель). Иными словами, нужно проверить, как (в каком направлении) изменяется  $\hat{y}_x$  при изменении  $x$ , соответствует ли направление этого изменения тому направлению, которое предполагалось при теоретическом и логическом анализе.

Для однофакторных моделей (уравнений парной регрессии) такое соответствие чаще всего имеет место. Так, в уравнении линейной регрессии  $\hat{y}_x = 2,142 + 0,51x_1$ , характеризующем зависимость дневной выработки рабочего-бетонщика ( $\hat{y}_x$ ) от уровня механизации его труда ( $x_1$ ), параметр  $a_1 > 0$ , т. е. при повышении уровня механизации выработка, как и ожидалось, также увеличивается. В другом уравнении линейной регрессии ( $\hat{y}_x = 1,377 + 1,208x_2$ ), выражающем зависимость дневной выработки от квалификации рабочего (тарифного разряда —  $x_2$ ), параметр  $a_1$  также положительный, что вполне согласуется с экономическими соображениями о том, что повышение квалификации рабочего должно приводить к росту производительности труда.

В уравнении регрессии ( $\hat{y}_x = 17,76 + \frac{58,13}{x}$ ), характеризующем зависимость себестоимости единицы продукции ( $\hat{y}_x$ ) от объема ее производства ( $x$ ), с увеличением  $x$  величина  $\frac{58,13}{x}$  уменьшается и  $\hat{y}_x$  также уменьшается, что тоже полностью соответствует теоретическим представлениям об обратной зависимости себестоимости единицы продукции от объема ее производства.

При построении уравнения параболы регрессии, характеризующего зависимость урожайности ( $y_x$ ) от количества внесенных на 1 га удобрений ( $x$ ), предполагалось, что с увеличением  $x$  урожайность должна повышаться, но ее увеличение будет постепенно замедляться. Как видно из рис. 13.3, знаки параметров  $a_1$  и  $a_2$  оказались таковы ( $a_1 > 0$ ;  $a_2 < 0$ ), что изменение  $\hat{y}_x$  соответствует тому, которое предполагалось.

Иногда может оказаться, что знаки параметров уравнения регрессии не соответствуют теоретическим соображениям о направлении изменения  $\hat{y}_x$ . Например, вместо ожидавшегося положительного параметра  $a_1$  он оказался отрицательным или наоборот. В подобных случаях нужно попытаться выяснить причины выявившегося несоответствия. Ими могут быть либо недоброкачественность исходной информации (ошибки в первич-

ных данных, недостаточное число наблюдений, неоднородность совокупности и т. п.), либо ошибки в расчетах, либо, наконец, не учтенные при теоретическом анализе взаимосвязи рассматриваемых признаков с другими. Последнее обстоятельство может привести к тому, что априорные предположения о направлении или характере изменения  $\hat{y}_x$  в конкретных условиях исследуемой совокупности не подтвердятся, рабочая гипотеза окажется недостаточно обоснованной или ошибочной.

Важное место при оценке модели занимает измерение тесноты связи. Как и при методе аналитических группировок, для измерения тесноты связи используются показатели вариации результативного признака ( $y$ ). Общая дисперсия  $\sigma^2 = \frac{\sum (y - \bar{y})^2}{n}$

характеризует общую вариацию результативного признака  $y$ , объясняемую влиянием всех факторов, от которых он зависит. Отклонения  $y - \bar{y}$  объясняются тем, что сочетание значений факторов, влияющих на  $y$ , у каждой единицы совокупности является индивидуальным, различным.

Кроме общей дисперсии, на основе уравнения регрессии можно вычислить средний квадрат отклонений расчетных значений  $\hat{y}_x$ , вычисленных по уравнению регрессии, от их средней величины (средняя величина  $\hat{y}_x$  равна общей средней  $\bar{y}$ ). Это мера колеблемости расчетных значений признака около их средней величины:

$$\sigma_{\hat{y}_x}^2 = \frac{\sum (\hat{y}_x - \bar{y})^2}{n} \quad (13.7)$$

На основе уравнения регрессии можно также вычислить средний квадрат отклонений фактических значений результативного признака от расчетных его значений, т. е. меру колеблемости  $y$  около  $\hat{y}_x$  (около линии регрессии):

$$\sigma_a^2 = \frac{\sum (y - \hat{y}_x)^2}{n} \quad (13.8)$$

В математической статистике доказывается, что

$$\sigma_y^2 = \sigma_{\hat{y}_x}^2 + \sigma_a^2 \quad (13.9)$$

Рассмотрим смысл дисперсий  $\sigma_{\hat{y}_x}^2$  и  $\sigma_a^2$  на примере зависимости урожайности от количества удобрений (см. рис. 13.3). Если бы урожайность зависела только от количества удобрений, т. е. если бы связь между этими признаками была бы не корреляционной, а функциональной и совершенно точно выражалась бы уравнением построенной нами параболы, то каждому значению  $x$  (количество удобрений) соответствовала бы вполне определенная величина  $y$  (урожайность). Все фактические значения  $y$  совпадали бы с расчетными:  $y = \hat{y}_x$ , т. е. все



точки на рис. 13.3 были бы расположены на параболе (линии регрессии). Отклонения  $y - \hat{y}_x$  были бы тогда равны нулю, дисперсия  $\sigma_y^2$  тоже равнялась бы нулю, отклонения  $\hat{y}_x - y$  совпадали бы с отклонениями  $y - \hat{y}_x$ , а дисперсия  $\sigma_{\hat{y}_x}^2$  — с общей дисперсией  $\sigma_y^2$ . Следовательно, равенство  $\sigma_{\hat{y}_x}^2 = \sigma_y^2$  свидетельствует о том, что связь  $y$  с  $x$  функциональная и вся вариация  $y$  объясняется вариацией фактора  $x$ .

Рассмотрим теперь противоположный случай. Если бы урожайность совершенно не зависела от количества удобрений, а число наблюдений было достаточно велико, то при любом количестве удобрений (при любом значении  $x$ ) урожайность ( $y$ ) была бы одинаковой или примерно одинаковой, равной или близкой к средней величине ( $\bar{y}$ ). Иначе говоря, при отсутствии связи и достаточно большом числе наблюдений все значения  $\hat{y}_x$ , независимо от величины  $x$ , были бы равны  $\bar{y}$ , т. е.  $y_x = \bar{y}$ . Уравнение регрессии выражалось бы тогда прямой, параллельной оси абсцисс, а отклонения  $\hat{y}_x - y$  и дисперсия  $\sigma_{\hat{y}_x}^2$  были бы равны нулю. Отклонения же  $y - \hat{y}_x$  совпадали бы с отклонениями  $y - \bar{y}$ , а дисперсия  $\sigma_y^2$  — с общей дисперсией  $\sigma_y^2$ . Следовательно, равенства  $\sigma_{\hat{y}_x}^2 = 0$  и  $\sigma_y^2 = \sigma_y^2$  свидетельствуют о том, что  $y$  не связан с  $x$ , что вся вариация  $y$  объясняется другими, не входящими в уравнение регрессии факторами.

Фактически в нашем примере, как видно из рис. 13.3, мы имеем дело не с рассмотренными крайними случаями (функциональная связь и полное отсутствие связи), а с некоторым промежуточным случаем, когда связь есть, но она является не функциональной, а корреляционной, так как, кроме количества удобрений ( $x$ ), на урожайность влияют и другие факторы. Поскольку  $y$  зависит от  $x$ , регрессия не совпадает с прямой  $y = \bar{y}$ ; поскольку же, кроме  $x$  на  $y$ , влияют и другие факторы, фактические значения урожайности не совпадают со значениями, вычисленными по уравнению регрессии, и точки на графике не лежат на линии регрессии.

Следовательно, то обстоятельство, что линия регрессии не совпадает с прямой  $y = \bar{y}$ , а отклонения  $\hat{y}_x - y$  не равны нулю, объясняется зависимостью  $y$  от  $x$ , т. е. тем, что вариация  $y$  сопровождается вариацией  $\hat{y}_x$ . Иными словами, вариация расчетных значений  $\hat{y}_x$ , их колеблемость около  $\bar{y}$ , измеряемая дисперсией  $\sigma_{\hat{y}_x}^2$ , объясняется влиянием на  $y$  только признака-фактора  $x$ , входящего в уравнение регрессии. Поэтому дисперсия  $\sigma_{\hat{y}_x}^2$ , которая характеризует вариацию  $y$ , объясняемую только фактором  $x$ , называется *факторной дисперсией*.

С другой стороны, если форма связи выбрана правильно (т. е. зависимость урожайности от количества удобрений, при прочих равных условиях, действительно выражается параболой второго порядка), отклонения фактических значений  $y$  от расчетных  $\hat{y}_x$  объясняются тем, что, кроме  $x$ , на  $y$  влияют и другие факторы, вариация которых сопровождается колебаниями  $y$  около  $\hat{y}_x$ , т. е. около линии регрессии. Иными словами, колеблемость  $y$  около  $\hat{y}_x$ , измеряемая дисперсией  $\sigma_y^2$ , объясняется вариацией остальных факторов, влияющих на  $y$ , т. е. прочих факторов, не входящих в уравнение регрессии. Из равенства (13.9) следует, что эта дисперсия может быть получена как остаток при вычитании факторной дисперсии из общей  $\sigma_y^2 = \sigma_{\hat{y}_x}^2 + \sigma_y^2$ .

Поэтому дисперсия  $\sigma_y^2$  называется *остаточной дисперсией*. Она характеризует остаточную вариацию  $y$ , т. е. вариацию, не объясненную фактором  $x$ , включенным в уравнение регрессии.

Для измерения тесноты связи между  $y$  и  $x$  логично поэтому использовать отношение факторной дисперсии к общей дисперсии результивативного признака. Это отношение называется *индексом детерминации* ( $i^2$ ):

$$i^2 = \frac{\sigma_{\hat{y}_x}^2}{\sigma_y^2}. \quad (13.10)$$

Из смысла дисперсий  $\sigma_{\hat{y}_x}^2$  и  $\sigma_y^2$  вытекает, что индекс детерминации показывает, какая часть общей вариации результивативного признака  $y$  объясняется признаком-фактором  $x$ , входящим в соответствующее уравнение регрессии.

Если форма связи, принятая в уравнении регрессии, не соответствует действительному механизму взаимосвязи, то остаточная дисперсия будет выражать не только вариацию  $y$ , объясняемую прочими (кроме  $x$ ) факторами, но и некоторую погрешность, вызванную неточностью формы связи, ее приближенным характером. Поэтому и факторная дисперсия также не измерит точно вариацию, объясняемую фактором  $x$ . С другой стороны, если фактор  $x$  корреляционно связан в свою очередь с одним или несколькими из прочих факторов, влияющих на  $y$ , то с изменением  $x$  будут изменяться и эти связанные с ним факторы, так что факторная дисперсия будет характеризовать вариацию  $y$ , объясняемую не только  $x$ , но и связанными с ним факторами. Рассматривая пока лишь парную регрессию и корреляцию, мы абстрагируемся от этого обстоятельства. Уточнение смысла индекса детерминации будет сделано при рассмотрении множественной регрессии и корреляции.

Если связь между  $y$  и  $x$  отсутствует и факторная дисперсия равна нулю, индекс детерминации также будет равен нулю. Линия же регрессии совпадает с прямой  $y = \bar{y}$ , параллельной оси абсцисс. При функциональной связи между признаками, когда факторная дисперсия  $\sigma_{\hat{y}_x}^2$  совпадает с общей дисперсией  $\sigma_y^2$ , индекс детерминации будет равен единице. Все точки на графике в этом случае лежат на линии регрессии.

Если же связь между признаками есть, но она является не функциональной, а корреляционной, то по мере ее усиления, т.е. по мере повышения тесноты связи, индекс детерминации увеличивается, приближаясь к единице, а по мере ослабления — уменьшается, приближаясь к нулю. Таким образом, индекс детерминации, измеряя тесноту связи, характеризует степень близости корреляционной связи к строгой функциональной.

В качестве показателя тесноты связи используются также квадратный корень из индекса детерминации, называемый ин-

Таблица 13.5. Расчет  $\Sigma(\hat{y}-\bar{y})^2$  для регрессии  $\hat{y}_x = 2,142 - 0,051x_1$

№ п/п	$x_1$	$\hat{y}_x$	$\hat{y}_x - \bar{y}$	$(\hat{y}_x - \bar{y})^2$
1	35	3,93	-1,07	1,1449
2	59	5,15	0,15	0,0225
3	44	4,39	-0,61	0,3721
...	...	...	...	...
...	...	...	...	...
49	67	5,56	0,56	0,3136
50	57	5,05	0,05	0,0025
Итого	2800	249,42	—	38,2976

дексом корреляции ( $i$ ). Индекс корреляции также может принимать значения от 0 до 1. Индекс детерминации ( $i^2$ ) и индекс корреляции ( $i$ ) могут быть использованы для изучения тесноты связи при любой ее форме — как прямолинейной (линейной), так и криволинейной.

Если, однако,  $y$  нелинейно зависит от параметров уравнения регрессии (например, если уравнение регрессии выражается степенной функцией  $\hat{y}_x = a_0 x^{a_1}$ , где зависимость  $y$  от  $x$  нелинейна), равенство (13.9) может не соблюдаться, так как средняя величина  $\hat{y}_x$  не равна  $\bar{y}$ . В таких случаях для приближенного измерения тесноты связи, как и при построении уравнения регрессии, функция приводится к линейному виду (например, путем логарифмирования  $\lg \hat{y}_x = \lg a_0 + a_1 \lg x$ ) и при расчете  $i^2$  в формулах (13.7) и (13.10) вместо  $y$ ,  $\hat{y}_x$  используются их логарифмы.

Определим тесноту связи между дневной выработкой рабочего ( $y$ ) и уровнем механизации труда ( $x_1$ ), зависимость между которыми выражается линейной регрессией  $\hat{y}_x = 2,142 + 0,051x_1$  (см. выше). Подставляя в это уравнение фактические значения  $x_1$  (см. табл. 12.4), найдем расчетные значения  $\hat{y}_x$ . Расчет суммы квадратов отклонений, необходимой для вычисления факторной дисперсии, показан в табл. 13.5 ( $y = 250; 5 = 5$ ).

Факторная дисперсия  $\sigma_{\hat{y}_x}^2 = \frac{38,2976}{50} = 0,7660$ . Общая дисперсия  $\sigma_y^2 = \bar{y}^2 - (\bar{y})^2 = 1337,96; 50 - \bar{y}^2 = 1,7592$  ( $\Sigma y^2 = 1337,96$ ; см. табл. 13.1).

Отсюда индекс детерминации равен:

$$i^2 = \frac{\sigma_{\hat{y}_x}^2}{\sigma_y^2} = \frac{0,7660}{1,7592} = 0,435.$$

Следовательно, 43,5 % общей вариации выработки ( $y$ ) объясняется вариацией уровня механизации труда ( $x_1$ ).

Расчет факторной дисперсии по формуле (13.7), требующей вычисления всех значений  $\hat{y}_x$ , является очень трудоемким, что заставляет прибегать к округлениям. Более удобно для расчета факторной дисперсии использовать другую формулу, не требующую вычисления  $\hat{y}_x$ .

Для линейной регрессии эта формула такова:

$$\sigma_{\hat{y}_x}^2 = a_0^2 + a_1^2 x = (a_0 \Sigma y + a_1 \Sigma yx) : n - (\Sigma y : n)^2. \quad (13.11)$$

В нашем примере имеет  $\Sigma y = 250; \Sigma yx = 14\,752$ , см. табл. 13.1):  $\sigma_{\hat{y}_x}^2 = (2,142 \cdot 250 + 0,051 \cdot 14\,752) : 50 - (250 : 50)^2 = 0,7570$ . Расхождение этого результата с предыдущим объясняется округлением при расчетах.

Для двучленной гиперболической регрессии формула факторной дисперсии имеет такой вид:

$$\sigma_{\hat{y}_x}^2 = a_0^2 + \frac{a_1}{x} \left( a_0 \Sigma y + a_1 \Sigma \frac{y}{x} \right) : n - (\Sigma y : n)^2. \quad (13.12)$$

Так, для построенного выше уравнения регрессии, характеризующего зависимость себестоимости единицы продукции от объема ее производства ( $\hat{y}_x = 17,76 + \frac{58,13}{x}$ ), факторная дисперсия равна:  $\sigma_{\hat{y}_x}^2 = (17,76 \cdot 515 + 58,13 \cdot 71,71) : 20 - (515 : 20)^2 = 2,6826$ ; общая дисперсия —  $\sigma_y^2 = 13\,341 : 20 - (515 : 20)^2 = 3,9875$  (необходимые суммы взяты из табл. 13.2).

Следовательно, индекс детерминации равен:  $i^2 = 2,6826 : 3,9875 = 0,673$ , т.е. более 67 % общей вариации себестоимости объясняется объемом производства продукции.

Для параболической регрессии второго порядка факторная дисперсия может быть исчислена по такой формуле:

$$\sigma_{\hat{y}_x}^2 = a_0^2 + a_1 x + a_2 x^2 = (a_0 \Sigma y + a_1 \Sigma yx + a_2 \Sigma yx^2) : n - (\Sigma y : n)^2. \quad (13.13)$$

Так, для регрессии, характеризующей зависимость урожайности от количества удобрений ( $y_x = 5,086 + 27,511x - 6,927x^2$ ), факторная дисперсия составляет:  $\sigma_{\hat{y}_x}^2 = (5,086 \cdot 750 + 27,511 \cdot 791,1 - 6,927 \cdot 899,95) : 30 - (750 : 30)^2 = 19,8166$  (суммы взяты из табл. 13.4).



Общая дисперсия урожайности, исчисленная по первичным данным, здесь равна 25,6; следовательно, индекс детерминации равен:  $r^2 = 19,8166 : 25,6 = 0,774$ , т. е. более 77 % общей вариации урожайности объясняется колеблемостью количества удобрений, внесенных на 1 га.

При линейной форме связи показатели ее тесноты называются коэффициентом детерминации ( $r^2$ ) и коэффициентом корреляции ( $r$ ). Коэффициент детерминации имеет тот же смысл, что и индекс детерминации, но используется только в случае линейной связи. Обычно сначала рассчитывается коэффициент корреляции. Он может быть исчислен по одной из следующих формул:

$$r = \frac{\sum (y - \bar{y})(x - \bar{x})}{n \sigma_y \sigma_x} = \frac{\bar{y}x - \bar{y} \cdot \bar{x}}{\sigma_y \sigma_x} = \frac{n \sum yx - \sum y \cdot \sum x}{\sqrt{[n \sum y^2 - (\sum y)^2] \cdot [n \sum x^2 - (\sum x)^2]}} \quad (13.14a, б, в)$$

Коэффициент корреляции может принимать значения от  $-1$  до  $+1$ , включая и 0. Отрицательные значения указывают на обратную связь, положительные — на прямую связь. При  $r=0$  линейная связь отсутствует. При  $r=\pm 1$  связь является функциональной. По абсолютной величине коэффициент корреляции равен индексу корреляции, если связь линейная.

Необходимо отметить, что сама по себе величина коэффициента (индекса) корреляции, даже близкая к единице, не является доказательством наличия причинно-следственной связи между изучаемыми явлениями. Она характеризует лишь меру корреляции между признаками.

Установление причинно-следственной зависимости происходит, как было сказано, в процессе анализа качественной природы явлений, который непременно должен предшествовать моделированию. В противном случае можно получить ложную (бессмысленную) корреляцию. Ранее, в гл. 12, рассматривался пример простого параллелизма в развитии двух явлений за достаточно большой период времени: рост выплавки стали и снижение заболеваемости корью и скарлатиной в нашей стране. Расчеты показывают, что механическое коррелирование уровней этих явлений, дает высокий отрицательный коэффициент корреляции, однако ясно, что в данном случае наличие корреляции не означает, что одно явление выступает причиной другого. Снижение заболеваемости, конечно же, является результатом достижений советского здравоохранения, повышения уровня жизни народа.

Найдем коэффициент корреляции между дневной выработкой рабочего бетонщика и уровнем механизации труда ( $x_1$ ), ис-

пользуя для расчета по последней формуле итоговые данные табл. 13.1.

$$r = \frac{50 \cdot 14\,752 - 250 \cdot 2\,800}{\sqrt{(50 \cdot 1337,96 - 250^2)(50 \cdot 171\,536 - 2\,800^2)}} = 0,6605.$$

Коэффициент детерминации  $r^2 = 0,436$ . Этот результат почти совпадает с исчисленным выше индексом детерминации (расхождение вызвано округлениями при расчетах).

По данным табл. 13.1 можно найти и коэффициент корреляции между дневной выработкой и квалификацией (тарифным разрядом —  $x_2$ ):

$$r = \frac{50 \cdot 781,4 - 250 \cdot 150}{\sqrt{(50 \cdot 1337,96 - 250^2)(50 \cdot 476 - 150^2)}} = 0,6566,$$

откуда  $r^2 = 0,431$ , т. е. 43 % вариации выработки ( $y$ ) объясняется вариацией тарифных разрядов.

При анализе и экономической интерпретации уравнений регрессии используются также коэффициент эластичности результативного признака относительно признака-фактора. Коэффициент эластичности показывает, на сколько процентов изменяется в среднем  $y$  при изменении  $x$  на 1 %. Коэффициент эластичности ( $\mathcal{E}$ ) вычисляется по такой формуле:

$$\mathcal{E} = \frac{\hat{y}_x}{\hat{x}} \cdot \frac{x}{y_x}, \quad (13.15)$$

где  $\frac{\hat{y}_x}{\hat{x}}$  — первая производная  $\hat{y}_x$  по  $x$ .

При большинстве форм связи коэффициент эластичности является переменным, т. е. изменяется с изменением признака-фактора  $x$ . Так, для линейной регрессии  $\mathcal{E} = a_1 x : (a_0 + a_1 x)$ . При степенной форме связи ( $\hat{y}_x = a_0 x^{a_1}$ ) коэффициент эластичности является постоянной величиной, равной  $a_1$ .

Показатели регрессии и корреляции — параметры уравнения регрессии, индексы или коэффициенты детерминации и корреляции, — исчисленные для ограниченной по объему совокупности, могут быть искажены действием случайных факторов. Поэтому нужно проверить, насколько эти показатели характерны для того комплекса условий, в которых находится исследуемая совокупность, не являются ли они результатом стечения случайных обстоятельств. Проверка значимости (существенности) показателей регрессии и корреляции производится с помощью критериев математической статистики: критерия  $t$ -Стюдента, дисперсионного критерия  $F$ -Фишера и др.

При линейной связи для оценки значимости параметров уравнения регрессии используется критерий  $t$ -Стюдента, который для  $a_0$  рассчитывается по формуле  $t_0 = a_0 \frac{\sqrt{n-2}}{\sigma_0}$ , а для  $a_1$  — по формуле  $t_1 = a_1 \frac{\sqrt{n-2}}{\sigma_1}$ .

Вычисления по этим формулам значения  $t$  сравниваются затем с критическими их значениями при принятом уровне значимости (существенности)  $\alpha$  и числе степеней свободы  $k=n-2$ . Критические значения  $t$  находятся по таблице распределения Стьюдента. В социально-экономических исследованиях уровень значимости  $\alpha$  обычно принимается равным 0,05. Если расчетное значение  $t$  больше критического, параметр признается значимым (отличается гипотеза о том, что параметр в действительности равен нулю и лишнее в силу случайных обстоятельств он оказался равным проверяемой величине).

Для уравнения линейной регрессии  $\hat{y}_x = 2,142 + 0,031 x_1$ , характеризующего зависимость дневной выработки от уровня механизации труда (см. выше),  $t_0 = (2,142 \cdot \sqrt{50-2}) : 0,997 = 14,88$  (остаточная дисперсия  $\sigma_a^2 = \sigma_e^2 = -0,006 = 1,7592 - 0,7660 = 0,9932$ , откуда  $\sigma_e = 0,997$ );  $t_1 = (0,031 \cdot \sqrt{50-2} \cdot 17,167) : 0,997 = 6,08$ ;  $[\sigma^2 x = \Sigma x_1^2 : n - (\Sigma x_1 : n)^2 = 171 536 : 50 - (2 800 : 50)^2 = 294,72$  (см. табл. 13.1), откуда  $\sigma_x = 17,167$ ].

По таблице распределения Стьюдента для  $k=50-2=48$  и  $\alpha=0,05$ , находим критическое значение  $t=2,01$ . Следовательно, оба параметра значимы.

Для каждого параметра можно построить доверительные интервалы (границы):  $a_j \pm t_{\alpha/2}$ , где  $t$  — коэффициент доверия по распределению Стьюдента при  $k=n-2$  степенях свободы;  $\mu_j$  — средняя ошибка  $a_j$  для  $a_0$ , равная  $\mu_0 = \sigma_e : \sqrt{n-2}$ , для  $a_1 = \mu_1 = \sigma_e : \sigma_x \sqrt{n-2}$ .

При линейной связи для оценки значимости коэффициента корреляции (и коэффициента детерминации) также можно использовать критерий  $F$ , который в этом случае рассчитывается по формуле

$$t = \frac{r \sigma_y}{\sigma_e} \cdot \sqrt{n-2} = r \sqrt{\frac{n-2}{1-r^2}}.$$

Так, для коэффициента корреляции между дневной выработкой и уровнем механизации получим  $t = 0,6605 \sqrt{\frac{50-2}{1-0,4363}} = 6,10$ . Это значи-

тельно больше критического значения  $t$  для 48 степеней свободы и  $\alpha=0,05$  (2,01), что свидетельствует о значимости коэффициента корреляции и существенности связи между выработкой и уровнем механизации.

При криволинейной связи для оценки значимости индекса корреляции (и индекса детерминации) используется критерий  $F$ , который может быть вычислен по одной из следующих формул:

$$F = \frac{\sigma_x^2}{\sigma_e^2} \cdot \frac{n-m}{m-1} = \frac{r^2}{1-r^2} \cdot \frac{n-m}{m-1},$$

где  $m$  — число параметров в уравнении регрессии. Расчетное значение  $F$  сравнивается с критическим (табличным) для принятого уровня значимости  $\alpha$  и числа степеней свободы  $k_1=m-1$  и  $k_2=n-m$ .

Для зависимости себестоимости единицы продукции от объема ее производства ( $n=20$ ;  $\hat{r}=0,673$ ;  $m=2$ ) получим  $F = \frac{0,673^2}{1-0,673^2} \cdot \frac{20-2}{2-1} = 37,05$ . Критическое значение  $F$  при  $\sigma=0,05$  ( $k_1=1$ ;  $k_2=18$ ) равно 4,41. Связь себестоимости с объемом производства продукции существенна.

Показатели регрессии и корреляции используются не только для оценки и анализа уже построенной модели (уравнения рег-

рессии), но и в процессе построения модели — для выбора оптимального ее варианта, если теоретический анализ и эмпирическая линия регрессии не позволяют отдать предпочтение ни одному из возможных вариантов формы связи. В этом случае можно построить уравнения регрессии с различными формами связи и оценить их сравнительную пригодность на основе сопоставления индексов корреляции или детерминации с учетом их значимости по критерию  $F$ .

Так, целесообразность использования параболы 2-го порядка вместо линейной регрессии можно проверить с помощью критерия  $F_3 = \left( \frac{\sigma_{\text{линейн}}^2}{\sigma_{\text{параб}}^2} - 1 \right) (n-3)$ . Если  $F_3 > F_{\text{кр}}$  при приня-

том  $\alpha$  и числе степеней свободы  $k_1=1$  и  $k_2=n-3$ , то использование параболы целесообразно; в противном случае предпочтительнее линейная регрессия.

Показатели тесноты связи используются также для отбора факторов при построении многофакторных моделей (уравнений множественной регрессии) и при подготовке исходной информации для математического программирования.

### 13.3. ПОСТРОЕНИЕ И ОЦЕНКА МНОГОФАКТОРНЫХ МОДЕЛЕЙ (УРАВНЕНИЙ МНОЖЕСТВЕННОЙ РЕГРЕССИИ). ОТБОР ФАКТОРОВ

Социально-экономические явления отличаются большой сложностью. Их уровень формируется под влиянием целого комплекса переплетающихся и взаимодействующих между собой факторов, действующих с разной силой и в разных направлениях. Поэтому построение однофакторных моделей — парных уравнений регрессий — обычно оказывается недостаточным. Сложные взаимодействия данного фактора с другими могут сделать неточными, искаженными показатели парной регрессии и корреляции. Сама специфика корреляционных связей требует включения в модель наиболее важных и существенных факторов. Отбор таких факторов и является важнейшей проблемой при построении уравнений множественной регрессии.

Как уже говорилось выше, отбор факторов производится на основе качественного, теоретического анализа с одновременным использованием статистико-математических критериев. Наиболее целесообразен трехстадийный отбор. На первой стадии — при априорном анализе — на факторы, включаемые в предварительный их перечень, можно не накладывать никаких особых ограничений (могут включаться различные варианты измерителей одного и того же фактора и т. п.). Далее, на второй стадии производится сравнительная оценка и отсев части факторов. Это делается на основе сочетания качественного анализа с анализом парных коэффициентов и индексов корреляции и оценкой



их существенности (значимости). Для этого составляется *матрица парных коэффициентов корреляции*, измеряющих тесноту линейной связи каждого фактора с результативным признаком и с каждым из остальных признаков-факторов (*корреляционная матрица*): (см. табл. 13.6).

Здесь  $x_0$  — результативный признак ( $y$ );  $x_1, x_2, x_3$  — признаки-факторы;  $r_{ij}$  — парный коэффициент корреляции между  $x_i$  и  $x_j$  ( $r_{11}=r_{22}=\dots=r_{ss}=1$ ;  $r_{01}=r_{10}$  и т. д.).

Таблица 13.6. Матрица парных линейных коэффициентов корреляции

	$x_0$	$x_1$	$x_2$	...	$x_j$	...	$x_s$
$x_0$	1	$r_{01}$	$r_{02}$	...	$r_{0j}$	...	$r_{0s}$
$x_1$	$r_{10}$	1	$r_{12}$	...	$r_{1j}$	...	$r_{1s}$
$x_2$	$r_{20}$	$r_{21}$	1	...	$r_{2j}$	...	$r_{2s}$
...				...		...	
$x_j$	$r_{j0}$	$r_{j1}$	$r_{j2}$	...	1	...	$r_{js}$
...				...		...	
$x_s$	$r_{s0}$	$r_{s1}$	$r_{s2}$	...	$r_{sj}$	...	1

Корреляционная матрица позволяет выявить факторы, которые находятся между собой в тесной линейной корреляционной взаимосвязи, близкой к функциональной.

Если в модель включаются два или более тесно взаимосвязанных фактора, то наряду с уравнением регрессии появляются и другие линейные зависимости. Подобное явление, называемое *мультиколлинеарностью*, искажает величину коэффициентов регрессии, затрудняет их экономическую интерпретацию.

Граница коллинеарности является переменной величиной и определяется в каждом конкретном случае теснотой связи отдельных факторов с результативным показателем. Считается, что между факторами отсутствует коллинеарная связь, если парные коэффициенты корреляции удовлетворяют неравенствам (рассмотрим два фактора):  $r_{01} > r_{12}$  и  $r_{02} > r_{13}$ . Если хотя бы одно из неравенств не выполняется, следует исключить один из факторов на основании качественного анализа и с учетом тесноты его связи с результативным признаком.

На основе корреляционной матрицы производятся также предварительный выбор между различными вариантами измерения того или иного фактора и исключение некоторых из них. Предпочтение отдается тем измерителям (или одному из них), который более тесно, чем другие, связан с результативным признаком, причем желательно, чтобы связь данного измерителя с  $y$  была теснее, чем его связь с другими факторными признаками, т. е. чтобы  $r_{0j} > r_{ij}$ .

Третья, заключительная стадия отбора факторов — их окончательный отбор производится в процессе построения различных вариантов многофакторных моделей — уравнений множественной регрессии и оценки значимости (существенности) их параметров (см. ниже).

Важное значение при построении уравнения множественной регрессии имеет также вопрос о *форме связи*. Выбор формы связи при наличии нескольких факторов значительно усложняется по сравнению с уравнениями парной регрессии. Одна из основных причин этого — взаимосвязь самих факторов, их взаимодействие. Поэтому выбору формы уравнения множественной регрессии должен предшествовать анализ парных связей, в том числе и между факторными признаками.

Вопросы выбора формы уравнения множественной регрессии разработаны еще недостаточно. Согласно одной точке зрения форма связи, оказавшаяся наилучшей при построении парной регрессии (минимальная величина скорректированной остаточной дисперсии), необязательно будет оптимальной для данного фактора в уравнении множественной регрессии, так как влияние фактора с учетом и без учета его взаимосвязи с другими факторами может проявляться различно. Другая точка зрения состоит в том, что если форма парной связи теоретически достаточно обоснована и обеспечивает по сравнению с другими формами связи существенно меньшую остаточную дисперсию, то эта же форма связи является наиболее приемлемой для данного фактора и в множественной регрессии (особенно если парные связи этого фактора с другими факторами менее тесные, чем с результативным признаком).

Построение уравнений множественной регрессии часто производится путем так называемого шагового (многошагового) анализа, в процессе которого производится расчет параметров и с помощью статистико-математических критериев завершается отбор факторов и уточняется форма связи каждого фактора с результативным признаком (шаговый анализ рассматривается ниже).

Определение числовых значений параметров уравнения множественной регрессии, как и парной регрессии, обычно производится методом наименьших квадратов, для чего строится и решается система нормальных уравнений. Для линейной множественной регрессии  $\hat{y}_{12} \dots \hat{y}_s = a_0 + a_1x_1 + a_2x_2 + \dots + a_sx_s$  система нормальных уравнений такова:





регрессий, то при истолковании их смысла фиксирование других факторов на постоянном уровне не предполагается. Это означает, что коэффициенты парных регрессий показывают, как изменяется  $y$  при увеличении данного фактора на единицу и при одновременном соответствующем изменении корреляционно связанных с ним других факторов. Так, в нашем примере увеличение тарифного разряда на единицу приведет к повышению выработки в среднем на 1,208 м<sup>3</sup> лишь при том условии, что одновременно изменится и величина всех факторов, которые корреляционно связаны с разрядом ( $x_2$ ). Но мы видели, что в условиях исследуемой совокупности увеличение разряда на единицу сопровождается повышением уровня механизации труда на 19,46 %. Следовательно, увеличение выработки на 1,208 м<sup>3</sup> объясняется не только повышением квалификации как таковой, но и ростом механизации труда, которым сопровождается повышение квалификации (тарифного разряда).

Таким образом, коэффициент 1,208 в уравнении парной регрессии характеризует не эффект повышения квалификации самой по себе, а условный эффект этого фактора, т. е. эффект, в котором отражается как непосредственное (чистое) влияние квалификации ( $x_2$ ) на выработку, так и влияние других факторов, корреляционно связанных с  $x_2$  и через этот фактор также влияющих на выработку. Иначе обстоит дело с коэффициентами при  $x_1$  и  $x_2$  в уравнении множественной регрессии. Каждый из этих коэффициентов тоже характеризует не только непосредственное влияние соответствующего фактора, не только его чистый эффект, но также и влияние других факторов, корреляционно связанных с данным и не входящих в уравнение регрессии. Однако влияние второго фактора, включенного в уравнение регрессии при этом элиминировано, устранено. Следовательно, хотя параметры 0,0356 и 0,8351 в уравнении множественной регрессии и не отражают только непосредственный, чистый эффект соответствующего фактора, все же эффект, который они характеризуют, «очищен» от влияния второго фактора, входящего в уравнение регрессии. Если бы в уравнении регрессии, кроме  $x_1$  и  $x_2$ , был включен еще один существенный фактор, корреляционно связанный с  $x_1$  и  $x_2$ , то в новом уравнении коэффициенты при  $x_1$  и  $x_2$  вновь изменились бы и стали бы еще точнее отражать непосредственное влияние факторов, их более чистую эффективность. И чем полнее в уравнении множественной регрессии представлены наиболее существенные и важные факторы, от которых зависит результативный признак, тем точнее коэффициенты регрессии выражают чистую эффективность факторов, тем менее условной становится характеристика этой эффективности, если форма связи выбрана правильно.

Для того чтобы отличить друг от друга показатели регрессии и корреляции при различном числе факторов и при различных признаках, принятых в качестве результативного (зависимого) признака, в ряде случаев целесообразно использовать

более гибкую систему обозначений. При этом все признаки обозначаются одной буквой  $x$  с соответствующим подстрочным номером, причем вместо прежнего обозначения  $y$  используется  $x_0$ . Параметры уравнения регрессии, расчетные значения результативного признака, а также дисперсия и коэффициенты и индексы корреляции и детерминации имеют тройную подстрочную нумерацию. Первым указывается номер результативного признака, затем — номер (или номера) факторного признака, влияющего (вариации) которого отражает данный показатель. После этого в скобках указываются номера других признаков-факторов, которые включены в уравнение регрессии, но влияние которых в данном случае элиминировано, устранено. В соответствии с этим уравнение регрессии, характеризующее зависимость уровня механизации ( $x_1$ ) от тарифного разряда ( $x_2$ ), будет записано так:  $\hat{x}_{12} = a_{1(12)} + a_{12}x_2$ ; уравнение же, выражающее зависимость выработки ( $y = x_0$ ) от разряда, — так:  $\hat{x}_{02} = a_{0(2)} + a_{02}x_2$ . Наконец, полученное выше уравнение множественной регрессии будет выглядеть так:

$$\hat{x}_{012} = a_{0(12)} + a_{01(12)}x_1 + a_{02(12)}x_2.$$

При оценке многофакторных моделей следует прежде всего обратить внимание на соответствие знаков параметров теоретическим и логическим представлениям о предполагаемом направлении влияния факторов на результативный показатель. В случаях, когда наблюдается несоответствие знаков апriorных предположениям, необходимо выяснить причины этого явления.

Как уже говорилось, причины такого несоответствия могут быть различными. Недостатки исходной информации (малое число наблюдений, неоднородность совокупности и т. д.) и ошибки при построении модели (при отборе факторов, при выборе формы связи) могут привести к тому, что уравнение регрессии окажется недоброкачественным: знаки параметров будут искажать действительное направление влияния соответствующих факторов или другие особенности изменения результативного признака. Однако ошибками могут оказаться как знаки параметров, так и наоборот, те апriorные теоретические и логические соображения, с которыми сопоставляются параметры уравнения регрессии: эти апriorные соображения могут оказаться недостаточно обоснованными, в них могут быть упущены и не учтены какие-либо взаимодействия факторов и тенденции, не совпадающие с предполагавшейся основной и очевидной тенденцией. Может, например, оказаться, что за факторами, предполагавшееся направление влияния которых не подтвердилось в модели, скрывается действие других факторов, корреляционно связанных с ними, но не включенных в модель.

Как и при парной регрессии, на основе уравнения множественной регрессии общую дисперсию результативного признака

( $\sigma_y^2 = \sigma_0^2$ , так как  $y = x_0$ ) можно разложить на две части: 1) средний квадрат отклонений расчетных значений от их средней величины — факторную дисперсию  $\sigma_{0(12...s)}^2$ ; и 2) средний квадрат отклонений фактических значений результативного признака от расчетных его значений — остаточную дисперсию  $\sigma_{\epsilon}^2 = \sigma_{0(12...s)}^2$ :

$$\sigma_0^2 = \sigma_{0(12...s)}^2 + \sigma_{\epsilon}^2 \quad (13.18a)$$

где

$$\sigma_{0(12...s)}^2 = \frac{\sum (\hat{x}_{0(12...s)} - \bar{x})^2}{n}; \quad (13.18b)$$

$$\sigma_{\epsilon}^2 = \frac{\sum (x_0 - \hat{x}_{0(12...s)})^2}{n} \quad (13.18b)$$

Факторная дисперсия  $\sigma_{0(12...s)}^2$  характеризует вариацию результативного признака ( $x_0$ ), которая при данной форме связи объясняется факторами  $x_1, x_2, \dots, x_s$ , включенными в уравнение регрессии. Остаточная дисперсия  $\sigma_{\epsilon}^2$  характеризует остаточную вариацию  $x_0$ , т. е. вариацию, объясняемую прочими, не включенными в уравнение регрессии факторами.

Факторная и остаточная дисперсии имеют указанный смысл лишь при условии, что связь между факторами, включенными и не включенными в уравнение регрессии, отсутствует (или мы от нее абстрагировались). При наличии такой связи факторная дисперсия, как и параметры уравнения, косвенно отражает и вариацию не включенных в модель факторов.

Для измерения тесноты связи между  $x_0$  и факторами, включенными в уравнение регрессии, используется, как и ранее, отношение факторной дисперсии к общей дисперсии результативного признака. Это отношение называется *совокупным индексом детерминации* (индексом множественной детерминации):

$$I_{0(12...s)}^2 = \frac{\sigma_{0(12...s)}^2}{\sigma_0^2} \quad (13.19)$$

$I_{0(12...s)}^2$  может принимать значения от 0 до 1 и показывает, какая часть общей вариации  $x_0$  объясняется факторами, включенными в уравнение регрессии (при данной форме связи).

В качестве показателя тесноты связи используется также квадратный корень из  $I_{0(12...s)}^2$ , называемый *совокупным индексом корреляции* ( $I$ ).  $I^2$  и  $I$  могут быть использованы для измерения тесноты связи при любой ее форме — как линейной, так и криволинейной.

Факторную дисперсию можно определить, не вычисляя предельно  $\hat{x}_{0(12...s)}$ , по рабочей формуле, аналогичной формуле

(13.11). При линейной связи  $x_0$  со всеми факторами эта формула такова:

$$\sigma_{0(12...s)}^2 = (a_0 \sum x_0 + a_1 \sum x_0 x_1 + \dots + a_s \sum x_0 x_s) : n - (\sum x_0 : n)^2 \quad (13.20)$$

Найдем с помощью этой формулы факторную дисперсию для полученного выше уравнения множественной регрессии, используя данные табл. 13.1:  $\sigma_{0(12...s)}^2 = (0,5002 \cdot 250 + 0,0356 \cdot 14\,752 + 0,8351 \times \times 781,4) : 50 - (250 : 50)^2 = 1,0554$ .

Отсюда совокупный индекс детерминации равен:  $I_{0(12...s)}^2 = 1,0554 : 1,7592 = 0,600$  ( $\sigma_0^2 = 1,7592$ ; см. выше). Таким образом, 60 % общей вариации дневной выработки объясняется вариацией уровня механизации труда и тарифных разрядов бетонщиков.

При линейной связи результативного признака со всеми включенными в уравнение регрессии факторами показатели тесноты связи называются *коэффициентами*. Совокупный коэффициент детерминации  $R_{0(12...s)}^2$  имеет тот же смысл, что и совокупный индекс детерминации. Он может быть исчислен не только на основе факторной и общей дисперсий (формула 13.19), но и с помощью следующих формул:

$$R_{0(12...s)}^2 = (a_1 \sigma_1 r_{01} + a_2 \sigma_2 r_{02} + \dots + a_s \sigma_s r_{0s}) : \sigma_0; \quad (13.21a)$$

$$R_{0(12...s)}^2 = \beta_1 r_{01} + \beta_2 r_{02} + \dots + \beta_s r_{0s}, \quad (13.21b)$$

где  $a_i = a_{0(12...s)(i+1)(i+1...s)}$  — коэффициент уравнения множественной регрессии при факторе  $x_i$ ;  $\sigma_i$  — среднее квадратическое отклонение фактора  $x_i$ ;  $r_{0i}$  — парный коэффициент корреляции;  $\beta_i$  — коэффициент (коэффициент регрессии, выраженный в стандартизованном масштабе):

$$\beta_i = a_i \sigma_i : \sigma_0 \quad (13.22)$$

Определим  $R^2$  для нашего примера, где  $a_1 = 0,0356$ ;  $a_2 = 0,8351$ ;  $\sigma_0 = 1,3263$ ;  $\sigma_1 = 17,1674$ ;  $\sigma_2 = 0,7211$ ;  $r_{01} = 0,6605$ ;  $r_{02} = 0,6566$  (см. выше).  $\beta$ -коэффициенты равны:  $\beta_1 = a_1 \sigma_1 : \sigma_0 = 0,0356 \cdot 17,1674 : 1,3263 = 0,4608$ ;  $\beta_2 = 0,8351 \cdot 0,7211 : 1,3263 = 0,4540$ . Отсюда  $R_{0(12...s)}^2 = 0,4608 \cdot 0,6605 + 0,4540 \cdot 0,6566 = 0,602$  (расхождение с  $I_{0(12...s)}^2$  вызвано округлениями при расчетах).

Совокупный коэффициент корреляции ( $R$ ) в отличие от совокупного индекса корреляции ( $I$ ) может принимать как положительные значения (прямая линейная связь), так и отрицательные (обратная линейная связь), тогда как равенство его нулю говорит об отсутствии линейной связи.

$\beta$ -коэффициенты, используемые в формуле (13.21б), имеют при анализе модели и самостоятельное значение. Дело в том, что коэффициенты множественной регрессии [ $a_{0(12...s)}$  и т.д.] зависят от единиц измерения соответствующих факторов и



поэтому непосредственно сопоставимы между собой. Более того, даже если все факторы измеряются в одинаковых единицах, коэффициенты регрессии при них все же несравнимы, так как не учитывают степени вариации факторов. Чтобы сделать коэффициенты регрессии сопоставимыми, все признаки нужно выразить в среднеквадратических отклонениях (в стандартизованном масштабе):  $t_i = \frac{x_i - \bar{x}_i}{\sigma_i}$ . Параметры уравнения регрес-

сии в этом случае тоже будут выражены в стандартизованном масштабе и будут представлять собой  $\beta$ -коэффициенты, связанные с обычными параметрами указанными выше соотношением (13.22).  $\beta$ -коэффициенты показывают, на сколько средних квадратических отклонений  $\sigma_0$  изменяется  $x_0$  при увеличении  $x_i$  на одно среднее квадратическое отклонение  $\sigma_i$  и при неизменности остальных факторов, входящих в уравнение регрессии. Для парной линейной регрессии  $\beta = r$ .

В рассматриваемом примере уравнения множественной регрессии параметры  $a_{01(2)}$  и  $a_{02(1)}$  значительно отличаются по величине один от другого, тогда как  $\beta$ -коэффициенты почти равны. Это свидетельствует о том, что степень влияния обоих факторов на результативный признак примерно одинакова: изменение любого из факторов на одно среднеквадратическое отклонение сопровождается изменением  $x_0$  примерно на  $0,46\sigma_0$ .

Для множественной регрессии могут также быть найдены частные коэффициенты эластичности  $x_0$  относительно каждого отдельного фактора:

$$\mathcal{E}_i = \frac{\partial \hat{x}_{012} \dots}{\partial x_i} \cdot \frac{x_i}{\hat{x}_{012} \dots}, \quad (13.23)$$

где  $\partial \hat{x}_{012} \dots$  — первая частная производная  $\hat{x}_{012} \dots$  по  $x_i$ .

Частные коэффициенты эластичности показывают, на сколько процентов изменяется в среднем  $x_0$  при увеличении  $x_i$  на 1% и при фиксировании других факторов на том или ином уровне. В нашем примере  $\mathcal{E}_1 = 0,0356 \cdot x_1 : (0,5002 + 0,0356 \cdot x_1 + 0,8351 \cdot x_2)$ . При фиксировании обоих факторов на среднем уровне ( $x_1 = 56$ ;  $x_2 = 3$ ) получим средний коэффициент эластичности:  $\mathcal{E}_1 = 0,40$ . Следовательно, при увеличении  $x_1$  с 56 до 56,56% (в 1,01 раза)  $x_0$  увеличивается на 0,4%, если  $x_2$  зафиксировано на среднем уровне. Аналогичные расчеты для  $x_2$  дают  $\mathcal{E}_2 = 0,50$ .

Если в уравнении множественной регрессии все факторы, кроме одного, зафиксировать на каком-либо (например, на среднем) уровне, то получим уравнение условно-чистой регрессии. Так, в нашем примере, подставив в уравнение множественной регрессии  $x_2 = x_2 = 3$ , получим:  $\hat{x}_{01(2)} = 0,5002 + 0,0356x_1 + 0,8351 \times 3 = 3,0055 + 0,0356 \cdot x_1$ . Уравнения условно-чистой регрессии могут быть изображены графически.

В случае множественной регрессии рассмотренные выше показатели тесноты связи (парные и совокупные индексы детерминации и корреляции) могут быть дополнены индексами частной детерминации и корреляции. Подобно тому как коэффициент уравнения множественной регрессии характеризует условно-чистое влияние соответствующего фактора на  $x_0$  при исключении влияния других факторов, входящих в уравнение регрессии, так и частный индекс детерминации характеризует тесноту связи между  $x_0$  и данным фактором при элиминировании связи  $x_0$  с другими факторами, включенными в уравнение регрессии.

Для парной регрессии, характеризующей зависимость дневной выработки ( $x_0$ ) от уровня механизации труда ( $x_1$ ), выше была вычислена факторная дисперсия  $\sigma_{01}^2 = 0,7660$ . После того как в уравнение регрессии был включен второй фактор — тарифный разряд ( $x_2$ ), факторная дисперсия увеличилась до  $\sigma_{012}^2 = 1,0554$ , т. е. на  $0,2894 (\sigma_{012}^2 - \sigma_{01}^2)$ . Следовательно, это увеличение вариации  $x_0$ , объясняемой входящими в уравнение регрессии факторами, и представляет собой ту часть вариации, которая объясняется новым фактором  $x_2$ , дополнительно включенным в модель. Иными словами, прирост факторной дисперсии ( $\sigma_{012}^2 - \sigma_{01}^2$ ) характеризует вариацию  $x_0$ , которую удалось дополнительно объяснить, включив в модель фактор  $x_2$ . Если этот прирост отнести к остаточной дисперсии для парной регрессии ( $\sigma_{0(1)}^2 = \sigma_0^2 - \sigma_{01}^2$ ), т. е. к вариации, объясняемой прочими, не входящими в парную регрессию факторами, начиная с  $x_2$ , то мы получим частный индекс детерминации между  $x_0$  и  $x_2$  при элиминировании влияния  $x_1$ :

$$i_{02(1)}^2 = \frac{\sigma_{012}^2 - \sigma_{01}^2}{\sigma_{0(1)}^2}; \quad i_{02(1)}^2 = \frac{\sigma_{012}^2 - \sigma_{01}^2}{\sigma_0^2 - \sigma_{01}^2}. \quad (13.24a, б)$$

Этот индекс характеризует тесноту связи между  $x_0$  и  $x_2$  после исключения влияния на эти признаки фактора  $x_1$ , т. е. при условии, что вариация  $x_0$  и  $x_2$  под влиянием  $x_1$  устранена, элиминирована. Если вариацию  $x_0$ , которая объясняется факторами, включенными в уравнение регрессии, называть объясненной вариацией, а остаточную вариацию, которая этими факторами не объясняется, — необъясненной, то смысл частного индекса детерминации можно выразить так: этот индекс показывает, какую часть не объясненной ранее вариации можно объяснить, включив в модель данный фактор. Графически элементы формул (13.24) можно представить на рис. 13.4.

В нашем примере частный индекс детерминации для  $x_2$  при элиминировании  $x_1$  равен:

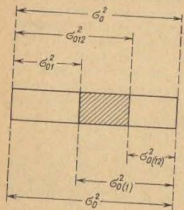
$$i_{02(1)}^2 = \frac{1,0554 - 0,7660}{1,7592 - 0,7660} = \frac{0,2894}{0,9932} = 0,291,$$

т. е. 29,1 % вариации  $x_0$ , не объясненной парной регрессией, объясняется фактором  $x_2$ .

Для фактора  $x_1$  будем иметь:

$$r_{01(2)}^2 = \frac{\sigma_{012}^2 - \sigma_{02}^2}{\sigma_0^2 - \sigma_{02}^2} = \frac{1,0554 - 0,7589}{1,7592 - 0,7589} = 0,296.$$

Выше был вычислен парный индекс детерминации  $r_{01}^2 = 0,435$ . Поскольку фактор  $x_1$  связан с  $x_2$ , парный индекс, кроме влияния  $x_1$ , отражает и косвенное влияние  $x_2$ . Поэтому элиминирование влияния  $x_2$  привело к уменьшению частного индекса детерминации по сравнению с парным.



В общем виде при элиминировании влияния факторов, начиная с  $x_1$  и кончая  $x_{k-1}$ , формула частного индекса детерминации между  $x_0$  и  $x_k$  имеет такой вид:

$$r_{0k(12 \dots s-1)}^2 = \frac{\sigma_{012 \dots s-1}^2 - \sigma_{012 \dots s-1}^2}{\sigma_0^2 - \sigma_{012 \dots s-1}^2} = \frac{\sigma_{012 \dots s-1}^2 - \sigma_{012 \dots s-1}^2}{\sigma_0^2 - \sigma_{012 \dots s-1}^2} \quad (13.25a, б)$$

Рис. 13.4. Общая ( $\sigma_0^2$ ), факторные ( $\sigma_{01}^2$  и  $\sigma_{012}^2$ ) и остаточные ( $\sigma_{0(1)}^2$  и  $\sigma_{0(12)}^2$ ) дисперсии для парной и множественной корреляции

Сравнивая формулу (13.25б) с формулой совокупного индекса детерминации при  $s$  факторах (13.19), нетрудно видеть, что элиминирование влияния факторов достигается путем вычитания из числителя и знаменателя, совокупного индекса детерминации факторов, объясняемых факторами, влияние которых нужно элиминировать.

Квадратный корень из частного индекса детерминации называется *частным индексом корреляции*  $r_{0k(12 \dots s-1)}$ . Частные индексы детерминации и корреляции изменяются от 0 до 1. Частный коэффициент детерминации  $r_{0k(12 \dots s-1)}^2$  имеет такой же смысл, как и частный индекс детерминации, но применительно к линейной связи. Частные коэффициенты корреляции  $r_{0k(12 \dots s-1)}$  имеют те же знаки, что и соответствующие им коэффициенты множественной линейной регрессии  $a_{0k(12 \dots s-1)}$ .

Показатели множественной регрессии и корреляции, как и показатели парной, могут оказаться подверженными действию случайных факторов. Проверка их значимости (существенности), т. е. оценка того, насколько они свободны от случайных воздействий, производится с помощью критерия  $t$ -Стьюдента и критерия  $F$ -Фишера.

При линейной зависимости  $x_k$  от двух факторов критерий  $t$ , используемый для оценки значимости коэффициентов регрессии, рассчитывается так:

$$t_1 = a_{01(2)} \sigma_1 \sqrt{(n-3)(1-r_{12}^2)} : \sigma_{0(12)};$$

$$t_2 = a_{02(1)} \sigma_2 \sqrt{(n-3)(1-r_{12}^2)} : \sigma_{0(12)}.$$

где  $r_{12}^2 = r_{21}^2$  — коэффициент детерминации, характеризующий тесноту связи между  $x_1$  и  $x_2$ .

Вычисленные по этим формулам значения  $t$  сравниваются с критическими (табличными) значениями при принятом уровне значимости  $\alpha$  и числе степеней свободы  $k = n - m$ .

Для построенного выше уравнения множественной регрессии имеем:  $n = 50$ ;  $m = 3$ ;  $a_{0(2)} = 0,0356$ ;  $a_{0(1)} = 0,8351$ ;  $\sigma_1 = 17,1674$ ;  $\sigma_2 = 0,7211$ ;  $r_{12}^2 = r_{21}^2 =$

$$= 0,1931; \sigma_{0(12)} = 0,8389. \text{ Отсюда } t_1 = \frac{0,0356 \cdot 17,1674 \sqrt{50 - 3(1 - 0,1931)}}{0,8389} =$$

$= 4,49$ ;  $t_2 = 4,42$ . По таблице распределения Стьюдента для  $k = 47$  и  $\alpha = 0,05$  находим  $t_{кр(2)} = 2,01$ . Так как вычисленные значения  $t_1$  и  $t_2$  больше критического, оба коэффициента  $[a_{0(2)}]$  и  $[a_{0(1)}]$  значимы.

Значимость совокупных индексов и коэффициентов детерминации (и корреляции) оценивается с помощью критерия  $F$ , который рассчитывается по формуле, приведенной на с. 238, в которой факторная и остаточная дисперсии берутся для множественной регрессии. Так, в нашем примере  $\sigma_{012}^2 =$

$$= 1,0554, \sigma_{0(12)}^2 = 0,7038, \text{ откуда } F = \frac{1,0554}{0,7038} \cdot \frac{50 - 3}{3 - 1} = 35,24. F_{кр} \text{ при}$$

$\alpha = 0,05$  ( $k_1 = 2$ ,  $k_2 = 47$ ) равно 3,20. Связь  $x_k$  с совокупностью двух факторов существенна.

Оценка значимости коэффициентов регрессии с помощью критерия  $t$  часто используется для завершения отбора факторов в процессе *шагового анализа*. Он заключается в том, что после решения модели и оценки значимости всех коэффициентов регрессии из модели исключается тот фактор, коэффициент при котором незначим и имеет наименьший коэффициент доверия  $t$ . После этого модель решается заново и снова производится оценка значимости всех оставшихся коэффициентов регрессии. Если среди них опять окажутся незначимые, то снова исключается фактор с наименьшим коэффициентом  $t$ . Процесс исключения факторов продолжается до тех пор, пока не будет получено уравнение регрессии, все коэффициенты в котором значимы.

Шаговый анализ целесообразен в тех случаях, когда основной интерес представляют сами коэффициенты регрессии как показатели эффекта влияния факторов. Если же задачи исследования предусматривают главным образом использование модели в целом для получения расчетных значений  $y$ , то недостаточный уровень значимости коэффициента регрессии не является еще решающим аргументом в пользу исключения из модели соответствующего фактора, особенно если он важен экономически и его нельзя игнорировать в практическом воздействии на результативный показатель. Поэтому нередко из модели исключаются лишь те факторы, без которых существенно не увеличивается скорректированная остаточная дисперсия (при ее



расчете сумма квадратов отклонений делится не на  $n$ , а на  $(n-m)$ , т. е. на число наблюдений за вычетом числа параметров уравнения регрессии).

### 13.4. ИСПОЛЬЗОВАНИЕ ГРУППИРОВОК ДЛЯ ВЫДЕЛЕНИЯ ОДНОРОДНЫХ ГРУПП ПРИ РЕГРЕССИОННО-КОРРЕЛЯЦИОННОМ АНАЛИЗЕ

В гл. 6 отмечалось, что группировки в статистике выполняют две функции: 1) являются самостоятельным методом статистического исследования, с помощью которого выделяются социальные-экономические типы, изучается структура, выявляются взаимосвязи; 2) являются приемом, определяющим границы и возможности применения других статистических методов. Рассмотрим вопросы применения группировок в роли приема, определяющего границы и возможности применения РКА.

Как уже говорилось, основным и обязательным условием корректности применения регрессионно-корреляционного анализа (РКА) является однородность исходной статистической совокупности. Так, например, если изучается зависимость урожайности определенной сельскохозяйственной культуры от количества внесенных удобрений, очень важно, чтобы совокупность колхозов была однородна по климатическим условиям, почвенным зонам, специализации и т. п., различие которых оказывает влияние на величину урожайности. Уровень производительности труда зависит от многих факторов. При изучении влияния какой-то определенной их группы, например факторов научно-технического прогресса, важно выяснить и по возможности устранить существенное влияние на изучаемую взаимосвязь некоторых условий производства, не имеющих отношения к изучаемым факторам (масштаб производства и др.).

Достижение определенной степени однородности возможно посредством применения статистических группировок. РКМ может быть построена на базе уже имеющихся типологических, структурных и аналитических группировок, если образованные группы достаточно многочисленны. Однако разнообразие реально существующих связей между признаками и явлениями и большие массивы исходных данных ставят перед исследователями задачи образования однородных совокупностей по признакам, для которых, как правило, отсутствуют какие-либо традиционные группы. Если в этом случае не руководствоваться определенными правилами, стандартами, алгоритмами, группировка одной и той же совокупности единиц может иметь очень много различных вариантов, и каждому варианту группировки будет соответствовать своя РКМ. Выводы, полученные на основе этих моделей, могут существенно отличаться, а иногда и быть прямо противоположными.

Отметим, что последовательность этапов группировки (см. гл. 6) сохраняется и в этом случае: нужно выбрать группиро-

вочные признаки, определить или наметить число групп, установить их границы и дать оценку результатам группировки.

Основой выбора группировочных признаков являющиеся изучение сущности исследуемого явления, характера изучаемой совокупности и задач, стоящих перед исследованием. Для объективного определения числа групп и их границ в последние годы разработаны различные способы и приемы, алгоритмы которых основаны на использовании статистических методов. Технический процесс расчленения совокупности на группы осуществляется, как правило, на ЭВМ. Предлагаемые алгоритмы расчленения совокупности на однородные группы имеют в своей основе различные принципы и предпосылки.

Приступая к образованию групп в исходной совокупности, исследователь не может заранее выбрать определенный способ группировки, так как ее окончательные результаты зависят как от характера изучаемой совокупности, так и от применяемого способа. Поэтому до применения того или иного способа группировки нельзя сказать, дадут ли ее результаты (образованные группы, совокупности) возможность решить задачи исследования. Решить вопрос о том, какому из способов отдать предпочтение в конкретном случае, возможно лишь сопоставив между собой результаты группировки различными способами и оценив эти результаты с точки зрения возможности решения поставленных задач.

С позиций последующего моделирования в пределах образованных совокупностей (групп) удачной можно признать такую группировку, в результате которой выполнены три основных условия: число единиц в каждой группе достаточно велико, обеспечена однородность единиц в каждой группе, а между группами имеют место существенные различия.

Проверка выполнения первого условия (достаточной наполненности групп) не требует особых расчетов. Для этого надо сопоставить число единиц в группах с числом параметров модели (см. выше).

Для оценки однородности единиц в группах используют либо дисперсию ( $\sigma^2$ ), либо — чаще — коэффициент вариации ( $v$ ), который устраняет несопоставимость признаков с различными единицами измерения и разными средними величинами. Однородность единиц подтверждается также близостью значений групповой средней к середине соответствующего интервала.

Надо отметить, что достижение абсолютной однородности единиц в группах практически неосуществимо. Речь может идти лишь о достижении такой степени однородности, при которой взаимосвязи, обнаруженные на основе обработки исходной информации, отражали бы реально существующие зависимости.

Чтобы оценить правильность группировки с точки зрения существенности различий между группами, можно использовать  $F$ -критерий Стьюдента. В зависимости от объема группы критерий рассчитывается в двух вариан-

так: 1) для многочисленных групп ( $n > 30$ ):

$$t_x = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}};$$

2) для немногочисленных групп ( $n < 30$ ):

$$t_x = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{(\sigma_1^2 n_1 + \sigma_2^2 n_2)(n_1 + n_2)}{(n_1 + n_2 - 2)n_1 n_2}}},$$

где  $n_i$ ,  $\bar{x}_i$  и  $\sigma_i^2$  — соответственно, число единиц, средняя и дисперсия в данной группе.

Сравнение вычисленных значений  $t$  с допустимыми дает основание для заключения о существенности различий между группами. При этом практически достаточно, чтобы  $t$ -расчетное было больше 3.

Наряду с указанными основными требованиями к группировке, которые должны быть выполнены во всех случаях совместного применения метода группировки и РКА, иногда имеют место требования к группировке, определяемые конкретными задачами исследования (обычно в отношении числа групп).

Рассмотрим применение метода группировки, как приема, определяющего границы применения РКА, на конкретном примере. Предположим, поставлена задача изучения влияния факторов технического прогресса на уровень производительности труда на 88 машиностроительных предприятиях одной из отраслей промышленности. Предварительная проверка показала, что по уровню производительности труда и факторам технического прогресса предприятия можно считать однородными. Однако логический анализ показал, что достоверное описание взаимосвязи факторов технического прогресса и производительности труда предполагает однородность исследуемой совокупности по масштабу производства. Известно, что крупные предприятия по сравнению с мелкими имеют большие возможности технического совершенствования производственного процесса, внедрения нового, более производительного оборудования и лучшего его использования.

В условиях научно-технического прогресса (НТП) о масштабе производства можно судить по объему производимой продукции, численности рабочих (работающих) и размеру основных производственных фондов. Возникает вопрос: можно ли судить о масштабе производства предприятий и их однородности по этому признаку, ограничившись какой-то одной из перечисленных характеристик масштаба производства?

Рассмотрим, например, показатель численности работающих. В условиях НТП оценка масштаба производства только по численности работающих не может быть точной, а группы предприятий, образованные на основе этого признака, нельзя счи-

тать однородными по масштабу производства. Ведь при этом в группу крупных предприятий с одинаковой численностью работающих попадут как действительно крупные предприятия с высоким уровнем развития техники (поточное производство, комплексная автоматизация и т. д.), так и предприятия значительно уступающие первым по масштабу производства, но имеющие большую численность работающих из-за низкого уровня используемой в производстве техники. Подобная картина качественной неоднородности предприятий может иметь место и в группе мелких предприятий.

Ели численность работающих как группировочный признак не может служить достаточным основанием для выделения качественно однородных групп по масштабу производства, то вызывает сомнение существенность этого признака. Однако нельзя признать его и несущественным, поскольку НТП имеет непосредственное отношение к затратам живого труда. «Прогресс техники в том и выражается, — писал В. И. Ленин, — что человеческий труд все более и более отстает на задний план перед трудом машин»<sup>1</sup>. Сочетание существенности и одновременно недостаточности рассматриваемого показателя как признака масштаба производства объясняется тем, что численность работающих — это характеристика лишь одной стороны процесса производства (затраты живого труда). Делать вывод о масштабе производства лишь по затратам живого труда, изолировано от других, не менее важных факторов процесса производства было бы ошибочным.

Проанализировав подобным образом другие характеристики масштаба производства, приходим к выводу, что в условиях НТП ни одна из них в отдельности не позволяет достаточно точно судить о масштабе производства. Для этого необходимо использовать комплекс всех трех признаков.

Для оценки степени однородности предприятий отрасли по масштабу производства рассчитываем коэффициенты вариации по всем трем его показателям. Допустим, что результаты расчетов таковы: объем товарной продукции —  $v = 90\%$ ; среднегодовая стоимость основных производственных фондов —  $v = 111\%$ ; среднечисловая численность работающих  $v = 120\%$ .

Это свидетельствует о неоднородности предприятий отрасли по всем показателям масштаба производства, поскольку значения коэффициентов вариации по всем показателям значительно превосходят его предельную величину (33%). В таких случаях достоверность регрессионно-корреляционной модели взаимосвязи уровня производительности труда с факторами технического прогресса, которая может быть построена для отрасли в целом, вызывает сомнение. Для адекватного описания этой взаимосвязи имеет смысл расчленить совокупность на более однородные по масштабу производства группы.

<sup>1</sup> Ленин В. И. Полн. собр. соч., т. I, с. 78.



Прежде всего необходимо предварительно наметить возможные типы предприятий, которые в последующем будут исследоваться. В данном случае имеет смысл попытка выделить три типа предприятий: мелкие, средние и крупные.

Итак, можно сформулировать следующие условия, которые должны быть выполнены в результате группировки: 1) по каждому показателю масштаба производства, взятому в качестве группировочного признака, число образованных групп должно быть равно трем; 2) в каждой отдельной группе должно быть

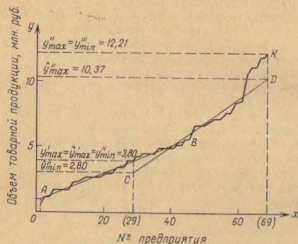


Рис. 13.5. Группировка предприятий отрасли графоаналитическим способом

достаточное число предприятий для построения надежной модели; 3) предприятия в пределах каждой группы должны быть однородны; 4) между группами должны быть существенные различия.

Представляется, что из ряда возможных способов группировки можно воспользоваться графоаналитическим<sup>1</sup>, с помощью которого границы групп определяются относительно несложно. Суть этого метода состоит в том, что первоначально вся исходная совокупность ранжируется по группировочному признаку и изображается графически в виде ломаной линии. На ось  $x$  наносятся порядковые номера единиц совокупности, а на ось  $y$  — величины группировочного признака. Определенные участки графика, не имеющие резких перепадов, визуально заменяются отрезками прямых, которые условно представляют искомые группы.

Далее визуальная группировка корректируется аналитически путем построения линейной регрессии  $y_x = a_0 + a_1x$ , где  $x$  — номер предприятия. На рис. 13.5 изображена линейная регрессия для первых 46 предприятий отрасли.

Доказано, что в случае линейной регрессии между  $\hat{y}_{\min}$ ,  $\hat{y}_{\max}$  и коэффициентом вариации  $v$  существует такое соотношение:

$$\hat{y}_{\max} = \frac{(0,574 + v) \hat{y}_{\min}}{0,574 - v} \quad (13.26)$$

Это позволяет корректировать визуальную группировку исходя из заданного значения  $v = 0,33$ . Величина расчетного минимального значения признака ( $\hat{y}_{\min}$ ) для I группы определяется при этом путем подстановки в уравнение линейной регрессии ( $y_x = a_0 + a_1x$ ) фактического минимального значения  $x$ . Так, в нашем примере  $\hat{y}_{\min} = 1,02$  млн. руб., или 1020 тыс. руб. Приняв  $v = 0,33$  по (13.26), получим для I группы:

$$\hat{y}_{\max} = \frac{(0,574 + 0,33) 1020}{0,574 - 0,33} \approx 3800 \text{ тыс. руб.}$$

Из рис. 13.5 видно, что расчетное значение  $\hat{y} = 3800$  примерно совпадает с соответствующим фактическим значением  $y$  (линия регрессии проходит здесь через соответствующую точку ломаной). Поэтому фактически верхняя граница I группы ( $y_{\max}$ ) совпадает с расчетной:  $y_{\max} = \hat{y}_{\max} = 3800$ . Таким образом, границами I группы будут  $y_{\min} = 770$  и  $y_{\max} = 3800$ . В ее состав войдут первые 29 предприятий. Коэффициент вариации для этой группы равен 0,35, т. е. незначительно отличается от заданной величины 0,33. Если бы мы ограничились визуальным определением границ I группы, то в нее следовало бы включить около 45 предприятий (см. рис. 13.5), для которых коэффициент вариации был бы значительно больше.

Для определения границ II группы находим линейную регрессию для последующих 40 предприятий. Нижняя граница II группы совпадает с верхней границей I (3800). Подставляя в уравнение регрессии для II группы значение  $x = 29$ , найдем расчетное минимальное значение для этой группы:  $\hat{y}_{\min} = 2800$ . Используя его, по (13.26), находим  $\hat{y}_{\max}$  для II группы:

$$\hat{y}_{\max} = \frac{(0,574 + 0,33) 2800}{0,574 - 0,33} \approx 10370 \text{ тыс. руб.}$$

Из рис. 13.5 видно, что это расчетное максимальное значение не совпадает с фактическим, которое равно 12210 тыс. руб.,  $y_{\max}$  можно найти графически (см. рис. 13.5).

Итак, границами II группы будут 3800 и 12210 тыс. руб. В ее состав войдут 40 предприятий отрасли, а коэффициент вариации составит 0,36. Следовательно, границами III группы

<sup>1</sup> Олесевич Л. П. Комбинаторные группировки в изучении технического прогресса. — Вестник статистики, 1975, № 7, с. 24—33.

будут 12 210 и 32 660 (максимальные  $y$  в отрасли). В нее войдут оставшиеся 19 предприятий.

Аналогично определяются границы всех групп по двум другим признакам масштаба производства (стоимости основных фондов и численности работающих). Коэффициенты вариации во всех группах по всем признакам находятся в пределах от 31 до 36 %, кроме одной группы по стоимости основных фондов ( $v=41$  %). Все группы существенно отличаются одна от другой

В нашем примере система моделей такова:  
для предприятий:

$$\begin{aligned} \text{мелких} & \rightarrow \hat{y} = \begin{cases} 1498 + 0,31x_1 + 83,8x_2 + 5,7x_3; \\ \text{средних} & 4885 + 86,5x_2 - 57,5x_3 + 105,5x_4; \\ \text{крупных} & -482,1 + 35,5x_3 + 26,1x_5 + 13,3x_6, \end{cases} \end{aligned}$$

где  $x_i$  — различные показатели технического прогресса (факторы),  $y$  — производительность труда.

Интерпретируя приведенную выше дискретно-непрерывную модель, можно отметить существенное расхождение в характере взаимосвязи факторов с результивным показателем в разрезе групп предприятий, отличающихся масштабами производства. Расхождения групповых регрессий для одного и того же результивного показателя выражаются в разном наборе аргументов, в различной степени влияния одних и тех же факторов, а в некоторых случаях в разнонаправленности их влияния. Все это еще раз подчеркивает качественное различие образованных групп объектов между собой и целесообразность исследования зависимости уровня производительности труда от факторов технического прогресса не по отрасли в целом, а по качественно однородным группам предприятий.

### 13.5. ОСНОВНЫЕ НАПРАВЛЕНИЯ ПРИМЕНЕНИЯ РЕГРЕССИОННО-КОРРЕЛЯЦИОННЫХ МОДЕЛЕЙ

Регрессионно-корреляционные модели могут быть использованы для решения различных задач: для анализа уровней социально-экономических явлений и процессов, например для анализа хозяйственной деятельности предприятий и вскрытия резервов, для прогнозирования и различных плановых расчетов.

Использование моделей позволяет значительно расширить возможности анализа, в частности анализа хозяйственной деятельности предприятий. При обычных, традиционных методах анализа и оценки деятельности предприятий показатели данного предприятия сравниваются со среднеотраслевыми (в сельском хозяйстве — со среднерайонными) показателями. Такие сравнения основаны на допущении, что все предприятия отрасли, района или той или иной группы работают примерно в одинаковых условиях, располагают более или менее одинаковыми объективными возможностями. Однако далеко не всегда дело обстоит именно так. Уровень результивного показателя на данном предприятии и средний его уровень по той или иной совокупности предприятий отличаются один от другого за счет очень многих факторов. Одни из них поддаются регулированию

Таблица 13.7. Комбинационная группировка предприятий отрасли по масштабу производства

Группы предприятий по численности работающих, чел.	Группы предприятий по стоимости основных производственных фондов, тыс. руб.	Группы предприятий по объему товарной продукции, тыс. руб.		
		770—3600	3600—12 210	12 210—32 660
234—1100	340—2 250	23 (М)	3	—
	2250—7 400	4	5	—
	7400—35 800	—	—	—
1100—3000	340—2 250	—	—	—
	2250—7 400	2	27 (С)	—
	7400—35 800	—	3	4
3000—8800	340—2 250	—	—	—
	2250—7 400	—	1	1
	7400—35 800	—	1	14 (К)
Итого		29	40	19

( $t > 3$ ). Минимальное число предприятий в группе (17) достаточно для построения надежной 3-факторной модели.

Однако для построения адекватных моделей необходима комбинационная группировка по всем трем признакам масштаба производства. Приняв границы групп, найденные рассмотренным выше способом, получим следующее распределение 88 предприятий (см. табл. 13.7).

Сочетание по всем трем признакам либо минимальных, либо средних, либо максимальных значений дает возможность выделить группу мелких ( $n=23$ ), средних ( $n=27$ ) и крупных ( $n=14$ ) предприятий (см. табл. 13.7).

Выделение этих трех групп позволяет построить не одну модель для отрасли в целом, а систему моделей, каждая из которых действительна в определенной, строго ограниченной области. Такое моделирование называется *дискретно-непрерывным*.



(управлению), целиком и полностью зависят от предприятия и отражают качество его работы, другие же лишь частично зависят от предприятия или даже почти не зависят от него (например, качество пашни в колхозе и т. п.).

Более обоснованно и правильно поэтому сравнивать фактический уровень результативного показателя (признака) на данном предприятии не со средним, а с расчетным уровнем, который вычислен по модели путем подстановки в нее значений факторов на данном предприятии. В этом случае сопоставляемые уровни ( $x_0$  и  $\hat{x}_0$ ) будут отличаться один от другого только за счет факторов, не входящих в модель. Следующим же этапом анализа будет сравнение расчетного ( $\hat{x}_0$ ) и среднего ( $\bar{x}_0$ ) уровней, различие которых обусловлено только факторами, включенными в модель. В результате отклонение фактического уровня от среднего ( $x_0 - \bar{x}_0$ ) расщепляется на две части: ( $x_0 - \hat{x}_0$ ) и ( $\hat{x}_0 - \bar{x}_0$ ).

Если в модель входят наиболее важные и существенные факторы, охватывающие основной комплекс объективных условий, от которых зависит уровень результативного показателя (признака), то остальные, не включенные в модель факторы, характеризуют главным образом индивидуальные, субъективные условия (например, организационные факторы, не поддающиеся количественной характеристике). В таких случаях общее влияние факторов, не включенных в модель, условно, с известными оговорками можно рассматривать как результат более или менее эффективного использования объективных условий и возможностей, т. е. как эффективность использования факторов, входящих в модель (оговорка заключается в том, что подразделять факторы на объективные и субъективные можно лишь с известной условностью). Поэтому расчетные уровни для отдельных предприятий (или для каких-либо их групп) выражают такие уровни результативного признака, которые были бы достигнуты при данных (фактических) значениях факторов, входящих в модель, и при средней по всей совокупности эффективности их использования. Сравнение же фактического уровня с расчетным ( $x_0 - \hat{x}_0$  либо  $\hat{x}_0 - \bar{x}_0$ ) показывает, насколько эффективность использования факторов на данном предприятии выше или ниже по сравнению со средней эффективностью их использования (если результативными признаками являются производительность труда, урожайность и т. д., то более высокая эффективность имеет место при ( $x_0 - \hat{x}_0$ ) > 0, если же результативными признаками являются себестоимость или трудоемкость единицы продукции, то при ( $x_0 - \hat{x}_0$ ) < 0).

Предприятия, у которых эффективность использования включенных в модель факторов значительно выше средней, могут быть исследованы с целью выявления передового опыта (или каких-либо благоприятных обстоятельств, способствовавших вы-

сокой эффективности), предприятия же, у которых эффективность ниже средней, необходимо исследовать, чтобы выявить причины этого и вскрыть неиспользованные резервы.

Предприятия, у которых эффективность использования включенных в модель факторов ниже средней, имеют резервы, связанные с возможным повышением уровня эффективности. Размер этих резервов зависит от того, ставится ли задача подтянуть эти предприятия до среднего по всей совокупности уровня эффективности использования факторов или же до того уровня, который уже достигнут остальными предприятиями. Резервы, связанные с достижением среднего уровня эффективности, можно определить, подставив в модель средние значения факторов на предприятиях, где эффективность ниже средней, и найдя затем отклонение среднего фактического уровня результативного признака на этих предприятиях от расчетного уровня:  $\bar{x}_0 - \hat{x}_0$ .

Определим эти резервы на основе модели, выражающей зависимость дневной выработки от уровня механизации труда и тарифного разряда бетонщиков:  $x_{012} = 0,5002 + 0,0356x_1 + 0,8351x_2$ . Рассчитанные по этой модели уровни выработки ниже фактических у 24 рабочих (№ 1, 5—9, 13 и т. д.). У этих рабочих  $x_1 = 58,4\%$ ;  $x_2 = 3$ , откуда  $x_{012} = 0,5002 + 0,0356 \cdot 58,4 + 0,8351 \times 3 = 5,08$ . Фактически же у этой группы рабочих средняя выработка составляет:  $\bar{x}_0 = 4,38$  м<sup>3</sup>. Следовательно, если бы эффективность использования факторов  $x_1$  и  $x_2$  у рабочих данной группы достигла среднего уровня, то их выработка повысилась бы на 0,7 м<sup>3</sup> (5,08—4,38), или на 16% (0,7:4,38=0,16). Общая выработка увеличилась бы за счет этого на 16,8 м<sup>3</sup> (0,7·24) и, при прочих равных условиях, составила бы 266,8 м<sup>3</sup> или 5,34 м<sup>3</sup> на одного рабочего, что на 6,8% выше, чем фактическая средняя выработка (5 м<sup>3</sup>).

Более полно можно учесть резервы, если иметь в виду подтягивание отстающих предприятий, рабочих и т. п. до того уровня эффективности использования факторов, который достигнут отстающими, передовыми в этом отношении предприятиями или рабочими. В этом случае нужно выделить группу таких единиц, у которых фактический уровень результативного признака либо вообще выше (лучше) расчетного, либо выше его не менее, чем на установленное число процентов (с учетом величины  $\sigma_p$ ). На основе первичных данных по такой группе единиц (предприятий, рабочих) нужно затем построить новую, «прогрессивную» модель. Подставив в эту новую модель средние значения факторов по остальным, т. е. отстающим и непередовым предприятиям (рабочим), можно найти новые расчетные уровни, сравнение которых с фактическими и покажет величину резервов. Следует, однако, отметить, что некоторые теоретические вопросы построения «прогрессивных» моделей изучены еще недостаточно.

В нашем примере по 26 рабочим, у которых эффективность использования факторов выше средней, новая, «прогрессивная» модель такова: пр.  $\hat{x}_{012} = 1,0094 + 0,0423 \cdot x_1 + 0,7633 \cdot x_2$  ( $R_2 = -0,824$ ). Прогрессивный расчетный уровень для остальных остающихся 24 рабочих равен:  $\hat{x}_{012} = 1,0094 + 0,0423 \cdot 58,4 + 0,7633 \times 3 = 5,77$  против 4,38 фактически. Следовательно, если бы у этих 24 рабочих эффективность использования факторов повысилась до уровня, достигнутого 26 передовыми рабочими, то выработка возросла бы на 1,39 м³ (5,77—4,38), или на 31,7 %. Подставив в новую модель средние значения факторов у всех 50 рабочих ( $\bar{x}_1 = 56$ ;  $\bar{x}_2 = 3$ ), найдем, что выработка составила бы в среднем 5,67 м³, что на 13,4 % выше фактической средней выработки (5 м³).

Анализ отклонений расчетных уровней от общего среднего уровня ( $\bar{x}_0$ ) дает возможность определить, за счет каких факторов и насколько расчетный уровень выше или ниже среднего уровня. Такой анализ может быть сделан как для отдельных единиц изучаемой совокупности, так и для тех или иных их групп.

Возьмем, например, первого по списку рабочего (см. табл. 13.1). Уровень механизации его труда ( $x_1 = 35$  %) ниже среднего ( $\bar{x}_1 = 56$  %) на 21 %. При средней эффективности использования факторов это должно было бы привести к снижению выработки по сравнению со средней на 0,748 м³ [ $(x_1 - \bar{x}_1) \cdot a_{01(2)} = (35 - 56) \times 0,0356 = -0,748$ ]. За счет более низкого, чем средний, тарифного разряда выработка должна была снизиться еще на 0,835 м³ [ $(x_2 - \bar{x}_2) \cdot a_{02(1)} = (2 - 3) \cdot 0,835$ ]. Следовательно, при фактических значениях факторов и средней эффективности их использования выработка была бы на 1,583 м³ ниже средней (5 м³), т. е. составила бы 3,417 м³. Это и есть расчетная выработка  $\hat{x}_{012}$ . Фактически же выработка первого рабочего составила только 3,0 м³, что свидетельствует о более низкой эффективности использования факторов ( $x_0 : \hat{x}_0 = 3,0 : 3,417 = 0,88$ , или 88 %).

Регрессионно-корреляционные модели широко применяются для проведения плановых расчетов, получения нормативов. Плановые нормативы в этом случае становятся более точными, объективными, чем при традиционном методе технико-экономических расчетов. Так, для предприятия одного производственного объединения были построены модели, позволяющие определить нормативное время между двумя капитальными ремонтами оборудования:

$$\hat{x}_{0123} = 44,189 + 1,861x_1^1 + 1,974x_2^1 - 17,278x_3,$$

где  $x_1$  — возраст оборудования в месяцах;  $x_2$  — количество проведенных текущих ремонтов;  $x_3$  — количество проведенных капитальных ремонтов.

Использование этих нормативов позволило оптимизировать величину ремонтного цикла, сократить расходы на капитальный ремонт оборудования.

Регрессионно-корреляционные модели находят широкое применение для прогнозирования социально-экономических явлений и процессов.

Прогнозирование представляет собой научно обоснованное выявление возможных путей и результатов предстоящего развития явлений и процессов, оценку показателей, характеризующих эти явления и процессы в будущем. Прогнозирование в условиях развитого социализма является важной формой предплановых исследований и анализа плановых показателей. Прогнозы позволяют плановым органам более глубоко и всесторонне обосновать определенный вариант плана, точнее учесть влияние имеющихся тенденций.

Для прогнозирования чаще всего используются многофакторные модели, позволяющие учесть влияние значительного числа факторов и дающие возможность получения многовариантных прогнозов. Вместе с тем в ряде случаев хорошие результаты могут быть получены и при прогнозировании по одnofакторной модели вида  $y = f(t)$ , когда в роли фактора-аргумента выступает время (подробнее об этом в гл. 14).

В качестве исходной базы для построения прогностической регрессионно-корреляционной модели может быть использована информация, характеризующая величины результативного и факторных признаков по отдельным объектам за определенный период времени (квартал, год), либо данные об изменяющихся значениях аргументов за ряд периодов. Прогнозирование осуществляется путем подстановки в построенную модель значений независимых переменных в определенном будущем; при этом предполагается, что форма взаимосвязи (значения коэффициентов регрессии) будет сохраняться неизменной на весь период прогнозирования. Таким образом, получение прогностических оценок результативного показателя предполагает прежде всего получение прогнозных значений факторов-аргументов. В случае, когда исходной базой является пространственная информация, значения независимых переменных на определенное будущее находятся, как правило, экспертным путем, на основании заключений специалистов в данной области. Когда же информация задана за ряд периодов, прогнозные значения факторов-аргументов могут быть рассмотрены на основе экстраполяции трендов (см. гл. 14).

В качестве примера прогнозирования по пространственной информации используем построенную ранее модель зависимости дневной выработки на формовке бетона у 50 рабочих от уровня механизации труда и тарифного разряда:  $\hat{x}_{012} = 0,5002 + 0,0356 \times x_1 + 0,8351 \cdot x_2$ . Предположим, что, по мнению экспертов, через 5 лет за счет улучшения организации производства и использования передового опыта средний уровень механизации повы-



сится до 65 %, а средний тарифный разряд составит 3,5. Тогда, принимая гипотезу о неизменном характере взаимосвязи, можно утверждать, что средняя дневная выработка через пять лет составит  $5,74 \text{ м}^3$  ( $0,5002 + 0,0356 \cdot 65 + 0,8351 \cdot 3,5$ ), т. е. возрастет по сравнению с нынешней ( $5 \text{ м}^3$ ) на 14,8 %.

Использование корреляционно-регрессионных моделей всегда предполагает оценку достоверности полученных результатов. Специфика этой проблемы применительно к прогнозированию очевидна: ведь объективная оценка прогноза может быть получена лишь в будущем, т. е. после его реального свершения. Но в этом случае ценность прогноза как важного элемента подготовки плана существенно снижается. Следовательно, возникает необходимость оценки достоверности прогноза уже на стадии его разработки. Эта проблема решается с помощью построения доверительных интервалов для точечных значений прогноза.

Считается, что «реальное» значение прогнозируемой величины  $\hat{x}_0$  находится в пределах  $\hat{x}_0 \pm t\hat{\sigma}_e = \hat{x}_0 \pm t \sqrt{\frac{\sum (x_0 - \hat{x}_0)^2}{n - m}}$ . Принимая  $\alpha = 0,05$ , при числе степеней свободы  $47(50 - 3)$ , определяем  $t = 2,01$ . Учитывая, что  $\sum (x_0 - \hat{x}_0)^2 = \sigma_e^2 \cdot n = 0,7038 \cdot 50 = 35,89$ , найдем  $\hat{\sigma}_e = \sqrt{35,89 : 47} = 0,8653$ . Тогда находим 95 %-ные доверительные интервалы прогнозного значения дневной выработки:  $5,74 \pm 2,01 \cdot 0,8653 = 5,74 \pm 1,74$ .

Оценка достоверности разработанных прогнозов обычно не ограничивается рассмотренной процедурой. После этого для большей точности целесообразно получить прогнозные значения с помощью другого метода и сравнить полученные оценки, которые в этом случае взаимно «проверяют» друг друга.

## ГЛАВА 14 АНАЛИЗ РЯДОВ ДИНАМИКИ

### 14.1. ДИНАМИКА СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ И ЗАДАЧИ ЕЕ СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ

Явления общественной жизни, изучаемые социально-экономической статистикой, находятся в непрерывном изменении и развитии. С течением времени — от месяца к месяцу, от года к году — изменяются численность населения и его состав, объем производимой продукции, уровень производительности труда и т. д. Поэтому одной из важнейших задач статистики является изучение изменения общественных явлений во времени — процесса их развития, их динамики. Эту задачу статистика решает путем построения и анализа рядов динамики (временных рядов).

Ряд динамики — это ряд числовых значений статистического показателя, расположенных в хронологической последовательности. Примерами динамических рядов могут служить следующие данные:

Таблица 14.1. Численность населения и число родившихся в СССР

Год	Численность населения (оценка на начало года), млн. чел.	Число родившихся, тыс.
1950	178,5	4805
1960	212,4	5341
1970	241,7 <sup>1</sup>	4226
1980	264,5	4851
1981	266,6	4961
1982	268,8	5100
1983	271,2	5456

<sup>1</sup> По переписи на 15 января.

Каждое числовое значение показателя, характеризующее величину, размер явления, называется уровнем ряда. Кроме уровня, каждый ряд динамики содержит указания о тех моментах либо периодах времени, к которым относятся уровни. В приведенном примере каждый уровень первого ряда динамики показывает, какова была численность населения СССР по состоя-

нию на 1 января соответствующего года, а каждый уровень второго ряда — сколько детей родилось в течение данного года.

При подведении итогов статистического наблюдения получают абсолютные показатели двух видов. Один из них характеризует состояние явления на определенный момент времени: наличие на этот момент каких-либо единиц совокупности или наличие того или иного объема признака. К таким показателям относятся численность населения, парк автомобилей, поголовье скота, жилищный фонд, товарные запасы и т. п. Величину таких показателей можно определить непосредственно только по состоянию на тот или иной момент времени, а потому эти показатели и соответствующие ряды динамики и называются моментными.

Другие показатели характеризуют итоги какого-либо процесса за определенный период (интервал) времени (сутки, месяц, квартал, год, пятилетку и т. п.). Такими показателями являются, например, число родившихся, количество произведенной продукции, ввод в действие жилых домов, фонд заработной платы и др. Величину этих показателей можно подсчитать только за какой-либо интервал (период) времени. Поэтому такие показатели и ряды их значений называются интервальными.

Из различного характера интервальных и моментных абсолютных показателей вытекают некоторые особенности (свойства) уровней соответствующих рядов динамики. В интервальном ряду величина уровня, представляющего собой итог какого-либо процесса за определенный интервал (период) времени, зависит от продолжительности этого периода (длины интервала). При прочих равных условиях, уровень интервального ряда тем больше, чем больше длина интервала, к которому этот уровень относится.

В моментных же рядах динамики, где тоже есть интервалы — промежутки времени между соседними в ряду датами, — величина того или иного конкретного уровня не зависит от продолжительности периода между соседними датами. Так, численность населения на начало 1980 г. не зависит, конечно, от того, приурочены ли другие уровни к началу каждого года (см. табл. 14.1) или же к его середине.

Каждый уровень интервального ряда уже представляет собой сумму уровней за более короткие промежутки времени. При этом единица совокупности, входящая в состав одного уровня, не входит в состав других уровней. Поэтому в интервальном ряду динамики уровни за примыкающие друг к другу периоды времени можно суммировать, получая итоги (уровни) за более продолжительные периоды (так, суммируя месячные уровни, получим квартальные, суммируя квартальные — получим годовые, суммируя годовые — многолетние, например, пятилетние).

Иногда путем последовательного сложения уровней интервального ряда за примыкающие друг к другу интервалы времени строится ряд нарастающих итогов, в котором каждый уро-

вень представляет собой итог не только за данный период, но и за другие периоды, начиная с определенной даты (с начала года, с начала пятилетки и т. д.). Такие нарастающие итоги нередко приводятся в отчетах предприятий. Построение ряда нарастающих итогов показано в табл. 14.2.

В моментном динамическом ряду один и те же единицы совокупности обычно входят в состав нескольких уровней. Так, большая часть жителей СССР, учтенных переписью на 15 января 1970 г., входила в состав населения в последующие годы, в том числе в состав населения на начало 1980, 1981 и 1982 гг. (см. табл. 14.1). При суммировании уровней моментного ряда обычно (в частности, в примере табл. 14.1) один единицы совокупности войдут в итог дважды, другие — большее число раз.

Поэтому суммирование уровней моментного ряда динамики само по себе не имеет смысла, так как получающиеся при этом итоги лишены самостоятельной экономической значимости.

Мы говорили выше о рядах динамики абсолютных величин, являющихся исходными, первичными. Могут быть построены также ряды динамики, уровни которых являются относительными и средними величинами. Они также могут быть либо моментными, либо интервальными:

Таблица 14.2. Жилищное строительство в СССР (млн. м<sup>2</sup> общей площади жилищ)

Год	Построено (введено в действие)	
	в данном году	с начала 11-й пятилетки
1981	106,4	106,4
1982	107,9	106,4 + 107,9 = 214,3
1983	112,4	214,3 + 112,4 = 326,7

#### Виды показателей и рядов динамики

По времени	По статистической характеристике показателя		
	Абсолютные величины	Относительные величины	Средние величины
Моментные	Поголовье коров во всех категориях хозяйств (на 1 января)	Поголовье коров в колхозах в процентах к общему поголовью (на 1 января)	Поголовье коров в среднем на один колхоз (на 1 января)
Интервальные	Производство молока во всех категориях хозяйств (по годам)	Производство молока в колхозах в процентах к общему производству (по годам)	Производство молока в среднем на один колхоз (или средний годовой удой молока от одной коровы) — по годам



В интервальных рядах динамики относительных и средних величин непосредственное суммирование уровней само по себе лишено смысла, так как относительные и средние величины являются производными и исчисляются путем деления других величин.

При построении и перед анализом ряда динамики нужно прежде всего обратить внимание на то, чтобы уровни ряда были сопоставимы между собой, так как только в этом случае динамический ряд будет правильно отражать процесс развития явления. *Сопоставимость уровней* ряда динамики — это важнейшее условие обоснованности и правильности выводов, полученных в результате анализа этого ряда. Несопоставимость уровней может возникнуть по разным причинам (см. гл. 5 и гл. 8). При построении динамического ряда вопрос о сопоставимости его уровней имеет особое значение, ибо ряд может охватывать большой период времени, в течение которого могли произойти изменения, нарушающие сопоставимость (территориальные изменения, изменения круга охвата объектов, методологии расчетов и т. д.).

При изучении динамики общественных явлений статистика решает ряд задач, чтобы дать числовую характеристику особенностей и закономерностей их развития на отдельных этапах: 1) измеряет *абсолютную и относительную скорость роста либо снижения* уровня за отдельные промежутки времени; 2) дает обобщающие характеристики уровня и скорости его изменения за тот или иной период; 3) выявляет и численно характеризует основные *тенденции развития* явлений на отдельных этапах; 4) дает сравнительную числовую характеристику развития данного явления в разных регионах или на разных этапах; 5) выявляет факторы, обуславливающие изменение изучаемого явления во времени; 6) делает прогнозы развития явления в будущем.

#### 14.2. ОСНОВНЫЕ ПОКАЗАТЕЛИ И ЭЛЕМЕНТАРНЫЕ МЕТОДЫ АНАЛИЗА ДИНАМИКИ

При изучении динамики используются различные показатели и методы анализа как элементарные, более простые, так и более сложные, требующие соответственно применения более сложных разделов математики.

*Простейшими показателями* анализа, которые используются при решении ряда задач, в первую очередь при измерении скорости изменения уровня ряда динамики, являются *абсолютный прирост, темпы роста и прироста*, а также *абсолютное значение (содержание) одного процента прироста*. Расчет этих показателей основан на сравнении между собой уровней ряда динамики. При этом уровень, с которым производится сравнение, называется *базисным*, так как он является базой сравнения. Обычно за базу сравнения принимается либо предыдущий, либо

какой-либо предшествующий уровень, например первый уровень ряда.

Если каждый уровень сравнивается с предыдущим, то полученные при этом *показатели* называются *цепными*, так как они представляют собой как бы звенья «цепи» связывающей между собой уровни ряда (рис. 14.1). Если же все уровни сравниваются с одним и тем же уровнем, выступающим как постоянная база сравнения, то полученные при этом *показатели* называются *базисными*.

Чаще построение ряда динамики начинают с того уровня, который будет использован в качестве постоянной базы сравнения. Выбор этой базы должен быть обоснован историческими социально-экономическими особенностями развития изучаемого явления. В качестве базисного целесообразно брать какой-либо характерный, типичный уровень, например конечный уровень предыдущего этапа развития (или средний его уровень, если на предыдущем этапе уровень то повышался, то понижался). В советской статистике в качестве базы сравнения часто принимаются: уровень 1913 г. — последнего «мирного» довоенного года, уровень 1922 г. — года образования СССР, 1940 г. — последнего довоенного года, уровень 1950 г. — года, когда закончилось послевоенное восстановление народного хозяйства, а также уровни за последние годы предыдущих пятилеток (1970, 1975, 1980 гг.) или средние уровни за предыдущие пятилетия.

*Абсолютный прирост* показывает, на сколько единиц увеличился (или уменьшился) уровень по сравнению с базисным, т. е. за тот или иной промежуток (период) времени. Абсолютный прирост равен разности между сравниваемыми уровнями и измеряется в тех же единицах, что и эти уровни:

$$\Delta = y_t - y_{i-t} \quad (14.1a)$$

где  $\Delta$  — абсолютный прирост за  $t$  единиц времени;  $y_t$  — сравниваемый уровень, а  $i$  — его либо хронологический, либо порядковый номер в ряду динамики;  $y_{i-t}$  — базисный уровень, а  $i-t$  — его номер;  $t$  — продолжительность периода (в тех или иных единицах времени), за который делается расчет.

Если уровень уменьшился по сравнению с базисным, то  $\Delta < 0$ ; он характеризует абсолютное уменьшение уровня.

Если за базу сравнения принимается предыдущий уровень, то цепной абсолютный прирост равен:

$$\Delta = y_t - y_{t-1} \quad (14.1b)$$

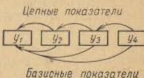


Рис. 14.1. Схема сравнения уровней при расчете цепных и базисных показателей динамики

Абсолютный прирост за единицу времени (мес., год, пятилетку) измеряет абсолютную скорость роста (или снижения) уровня.

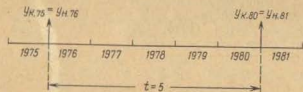
Расчет цепных и базисных абсолютных приростов показан в табл. 14.3.

Если нужно найти абсолютный прирост за какой-либо конкретный период (например, за 1976—1980 гг., т. е. за 5 лет ( $t=5$ ), то для интервального показателя в формуле (14.1а) в качестве  $y_t$  следует принять конечный уровень этого периода ( $y_{1980}$ ), а так как  $t=5$ , то  $i-t=1980-5=1975$ , т. е. в качестве базисного нужно принять  $y_{1975}$ :  $\Delta_{1976-1980} = y_{1980} - y_{1975}$ .

Если же за те же 5 лет нужно найти абсолютный прирост для моментного показателя, то в качестве  $y_t$  следует взять уровень на конец данного периода ( $y_{к. 1980}$ ), а в качестве  $y_{i-t}$  — уровень на конец предыдущего периода ( $y_{к. 1975}$ ):  $\Delta_{1976-1980} = y_{к. 1980} - y_{к. 1975}$ . Если же даны уровни на начало года, то  $\Delta_{1976-1980} = y_{н. 1981} - y_{н. 1976}$  (так как конец одного года — это начало следующего).

Во всех случаях разность между номерами уровней равна длине периода, за который делается расчет ( $t=1980-1975=1981-1976=5$ ). Это целесообразно использовать в качестве контроля правильности расчета.

Еще более наглядным и простым является контроль для моментных показателей с помощью схемы:



Годы в схеме следует подписывать под отрезками.

Возможен и другой случай: конкретный период, за который нужно сделать расчет, непосредственно не задан, а известны лишь два уровня, например  $y_{1983}$  и  $y_{1980}$ . Тогда  $i=1983$ ;  $i-t=1980$ , откуда  $t=1983-1980=3$ . Следовательно,  $\Delta_{1981-1983} = y_{1983} - y_{1980}$  (для моментных показателей уровни берутся в соответствии со схемой контроля).

Цепные и базисные абсолютные приросты связаны между собой: сумма последовательных цепных приростов равна соответствующему базисному приросту, т. е. общему приросту за весь период. Например:  $\Delta_{1981} + \Delta_{1982} + \Delta_{1983} = \Delta_{1981-1983}$ , т. е. приросту за 3 года (1981—1983 гг.).

Более полную характеристику роста можно получить только тогда, когда абсолютные величины дополняются относительными. Относительными показателями динамики являются темпы роста и темпы прироста, характеризующие *интенсивность процесса роста*.

Темп роста ( $T_p$ ) показывает, во сколько раз увеличился уровень по сравнению с базисным, а в случае уменьшения — какую часть базисного уровня составляет сравниваемый уровень:

$$T_p = \frac{y_t}{y_{i-t}}. \quad (14.2а)$$

Если за базу сравнения принимается предыдущий уровень, то цепной темп роста равен:

$$T_p = \frac{y_t}{y_{t-1}}. \quad (14.2б)$$

Как и другие относительные величины, темп роста может быть выражен не только в форме коэффициента (простого отношения уровней), но и в процентах:  $T_p (\%) = T_p \cdot 100 \%$ .

Пример расчета базисных и цепных темпов роста приведен в табл. 14.3.

Таблица 14.3. Цепные и базисные абсолютные приросты и темпы роста производства электроэнергии в СССР

Год	Производство электроэнергии, млрд. кВт.ч	Абсолютные приросты по сравнению		Темпы роста по сравнению	
		с 1980 г.	с предыдущим годом	с 1980 г. (1980—1)	с предыдущим годом
	$y_t$	$y_t - y_{80}$	$y_t - y_{t-1}$	$y_t : y_{80}$	$y_t : y_{t-1}$
1980	1294	×	...	1	...
1981	1326	32	32	1,025	1,025
1982	1367	73	41	1,056	1,031
1983	1418	124	51	1,096	1,037

Как и абсолютные приросты, темпы роста для любых рядов динамики сами по себе являются интервальными показателями, т. е. характеризуют тот или иной промежуток (интервал) времени.

Между цепными и базисными темпами роста, выраженными в форме коэффициентов, существует определенная взаимосвязь: произведение последовательных цепных темпов роста равно базисному темпу роста за весь соответствующий период.

Для большей простоты и наглядности доказательства этой взаимосвязи обратим внимание, что перемножив, например, два цепных темпа роста, мы получим:  $\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} = \frac{y_3}{y_1}$ , где  $i=3$ ,  $i-t=1$ ,  $t=2$ , т. е. число перемножаемых цепных темпов равно длине периода, за который делается расчет. Аналогичный вывод можно сделать, перемножив три и т. д. цепных темпа. Следовательно, перемножив  $t$  цепных темпов роста  $\left( \frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \dots \cdot \frac{y_{t+1}}{y_t} \right)$ ,



мы получим темп роста за  $t$  лет, который согласно принятым в формуле (14.2а) обозначениям равен  $\frac{y_t}{y_{t-1}}$ .

Взаимосвязь цепных и базисных темпов роста можно использовать для перехода от одних к другим, особенно если абсолютные уровни ряда динамики неизвестны. Если, например, производительность общественного труда в СССР в 10-й пятилетке (1976—1980 гг.) возросла на 17 %, т. е. 1,17 раза, а за 3 года 11-й пятилетки — на 9,9 %, то нетрудно определить, как она повысилась за 8 лет, т. е. в 1983 г. по сравнению с 1975 г.:

$$\frac{y_{83}}{y_{75}} = \frac{y_{80}}{y_{75}} \cdot \frac{y_{83}}{y_{80}} = 1,17 \cdot 1,099 = 1,286,$$

т. е. в 1983 г. производительность общественного труда возросла по сравнению с 1975 г. почти на 29 %.

Темп прироста характеризует относительную величину прироста, т. е. его величину по отношению к базисному уровню:

$$T_{пр} = \frac{\Delta}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1 = T_p - 1, \quad (14.3а, б)$$

где  $T_{пр}$  — темп прироста за  $t$  единиц времени, остальные обозначения прежние.

Выраженный в процентах, темп прироста показывает, на сколько процентов увеличился (или уменьшился) уровень по сравнению с базисным, принятым за 100 %:

$$T_{пр} (\%) = (T_p - 1) \cdot 100 \% = T_p (\%) - 100 \%. \quad (14.3в)$$

Следовательно, темп прироста всегда на единицу (или на 100 %) меньше соответствующего темпа роста. Так, по данным табл. 14.3 производство электроэнергии в СССР в 1983 г. по сравнению с 1980 г., т. е. за 3 года, увеличилось на 124 млрд. кВт·ч ( $\Delta$ ), в 1,096 раза ( $T_p$ ), или на 9,6 % ( $T_{пр} (\%)$ ).

При анализе темпов развития никогда не следует упускать из виду, какие абсолютные величины — уровни и абсолютные приросты — скрываются за темпами роста и прироста. Нужно, в частности, иметь в виду, что при снижении (замедлении) темпов роста и прироста абсолютный прирост может возрастать. Так, например, производство цветных телевизоров в СССР в 1978 г. по сравнению с 1974 г., т. е. за 4 года увеличилось в 3,5 раза (на 252 %), а за следующие 4 года (в 1982 г. по сравнению с 1978 г.) — только в 2,2 раза (на 118 %). Однако в абсолютном выражении прирост за 1979—1982 гг. был значительно больше, чем за 1975—1978 гг.: почти 1,7 млн. шт. против 1,0 млн. шт. Если в 1975—1978 гг. каждый из 252 % прироста в абсолютном выражении составлял 4 тыс. шт. ( $1:252=0,004$ ), то в 1979—1982 гг. каждый из 118 % прироста выражался уже величиной 14 тыс. шт. ( $1:118=0,014$ ).

Использованный здесь показатель — абсолютное значение (содержание) 1 % прироста ( $\Delta$ ) — вычисляется так:

$$\Delta = \frac{\Delta}{T_{пр} (\%)} = \Delta : \left( \frac{\Delta}{y_{t-1}} \cdot 100 \right) = \frac{y_{t-1}}{100}. \quad (14.4а, б)$$

Как уже говорилось, иногда уровни явления за одни годы непоставимы с уровнями за другие годы из-за территориальных, ведомственных и иных изменений (изменение методологии учета и исчисления показателей и т. п.). Чтобы обеспечить сопоставимость и получить пригодный для анализа временной ряд, нужно произвести прямой пересчет уровней, непоставимых с другими. Однако иногда нет необходимых для этого данных. В таких случаях можно использовать особый прием, называемый *смыканием рядов динамики*.

Пусть, например, в 1980 г. произошло изменение границ района, в связи с чем данные за 1978—1979 гг. оказались несопоставимыми с данными за 1981—1982 гг. (табл. 14.4).

Таблица 14.4. Средний удой молока от одной коровы, кг

Год	До изменения границ района (в старых границах)	После изменения границ (в новых границах)	1980 г. = 100 %		Сомкнутый ряд динамики (1980 г. = 100 %)
			В старых границах	В новых границах	
А	1	2	3	4	5
1978	2266	...	103	...	103
1979	2222	...	101	...	101
1980	2200	2100	100	100	100
1981	...	2058	...	98	98
1982	...	2205	...	105	105

Чтобы сомкнуть эти ряды и получить возможность анализа динамики среднего удоя за весь период, примем в каждом из них за базу сравнения уровень 1980 г., за который эти данные как в прежних, так и в новых границах района (см. гр. 3 и 4). Эти два ряда с одинаковой базой сравнения можно затем заменить одним сомкнутым рядом динамики (гр. 5). По данным такого сомкнутого ряда можно исчислить темпы роста по сравнению с любым годом. Можно рассчитать и абсолютные уровни за 1978—1979 гг. в новых границах. Так, в 1978 г. средний удой составлял 2163 кг ( $2100 \times 1,03$ ). Нужно, конечно, иметь при этом в виду, что результаты, полученные путем смыкания рядов динамики, содержат в себе некоторую погрешность.

Графически динамика явлений наиболее часто изображается в виде столбчатых (см. рис. 8.3) и линейных диаграмм (рис. 14.2, 14.3, 14.8, 14.9). Применяются и другие формы диаграмм — фигурные, квадратные, секторные и т. п. (см. гл. 7 и гл. 8). Аналитические графики обычно строятся в виде линейных диаграмм. Столбчатые диаграммы более целесообразны, чем

линейные, в тех случаях, когда уровни ряда взяты только за некоторые отдельные годы, а число уровней невелико (см. рис. 8.3). Порядок следования столбиков должен быть строго хронологическим.

Линейные диаграммы целесообразны, если: а) число уровней в ряду динамики велико; б) наиболее важным является изображение картины изменения в целом (его тенденции — см. ниже); в) на одном графике нужно изобразить несколько динамических рядов с целью их сравнения; г) более важным является изображение относительных изменений — темпов роста, а не самих уровней и их абсолютных приростов. Линейные графики практически более удобны для изображения динамики, так как линия лучше изображает непрерывность процесса развития.

Для построения линейной диаграммы используется прямоугольная система координат, на осях которой обычно строится числовая сетка. На оси абсцисс откладывается время, а на оси ординат строится шкала либо уровней ряда динамики, либо базисных темпов роста. Теоретически моменты времени следует

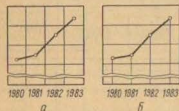


Рис. 14.2. Линейная диаграмма интервального ряда динамики

подписывать под точками шкалы времени, отмеченных черточками, а интервалы времени (месяца, годы и т. п.) — под отрезками этой шкалы. Но практически это часто не соблюдается и интервалы также подписываются под точками (рис. 14.2, б).

Линейная диаграмма — ломаная линия — наиболее соответствует характеру моментных показателей, уровни которых изображаются точками в вершинах соответствующих ординат. Интервальные же показатели теоретически правильнее изображать столбиками или ступенчатой линией (верхние границы столбиков, соединены по вертикали). Однако из соображений практического удобства динамика интервальных показателей чаще изображается в виде ломаной линии (рис. 14.3). Если при этом периоды (годы и т. п.) подписаны под отрезками шкалы, то точки, изображающие уровни, нужно наносить над серединами этих отрезков; если же периоды подписаны под точками шкалы времени, то над этими точками шкалы наносятся и точки, изоб-

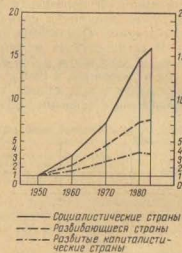


Рис. 14.3. Темпы роста промышленной продукции в социалистических и в остальных странах (1950—1)

ражающие уровни (см. рис. 14.2). При построении линейных диаграмм динамики в случаях, когда между датами или периодами (например, годами), к которым относятся уровни, промежутки времени неодинаковы, необходимо строго соблюдать масштаб при построении шкалы времени (см. рис. 14.3).

При равномерной шкале на оси ординат прямая линия изображает рост уровня ряда в арифметической прогрессии (равные абсолютные приросты за равные промежутки времени). При неравных абсолютных приростах прямая переходит в ломаную,

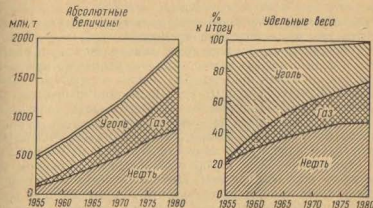


Рис. 14.4. Добыча топлива в СССР (в пересчете на условное топливо)

причем чем больше абсолютный прирост, тем больше угол наклона отрезка ломаной к оси абсцисс.

Если основной задачей линейного графика является изображение и сопоставление темпов роста (а не абсолютных приростов), то на оси ординат строится логарифмическая шкала. Тогда прямой линией будет изображаться ряд динамики с равными темпами роста за равные промежутки времени, а ломаной линией — ряд с неодинаковыми темпами роста. При этом большому темпу роста соответствует больший угол наклона отрезка ломаной к оси абсцисс.

Для изображения изменения структуры (состава) могут быть использованы наряду с секторными линейными диаграммами, называемые слоевыми. Эти диаграммы могут изображать как абсолютные величины, так и удельные веса, которые изображаются точками на ординатах соответствующих лет последовательно: одна составная часть (на рис. 14.4 — нефть) — начиная от базовой (нулевой) линии, вторая — от конца первой и т. д.

На рис. 14.4 видно, что добыча нефти (вместе с газовым конденсатом), а также газа растет и по абсолютному размеру, и по удельному весу в общей добыче, тогда как доля угля в общей добыче систематически снижается.



С течением времени изменяются не только уровни явлений, но и показатели их динамики — абсолютные приросты и темпы развития. Поэтому для обобщающей характеристики развития, для выявления и измерения типичных основных тенденций и закономерностей и решения других задач анализа используются средние показатели временного ряда — средние уровни, средние абсолютные приросты и средние темпы динамики.

К расчету средних уровней ряда динамики часто приходится прибегать уже при построении временного ряда — для обеспечения сопоставимости числителя и знаменателя при расчете средних и относительных величин. Пусть, например, нужно построить ряд динамики производства молока на душу населения в СССР. Для этого за каждый год необходимо количество произведенного в этом году молока (интервальный показатель) разделить на численность населения в этом же году (моментный показатель, величина которого непрерывно меняется на протяжении года). Ясно, что численность населения на тот или иной момент времени в общем случае несопоставима с объемом производства за весь год в целом. Для обеспечения сопоставимости нужно и численность населения как-то приурочить ко всему году, а это можно сделать, лишь рассчитав среднюю за год численность населения.

Часто приходится прибегать к средним показателям динамики и потому, что уровни многих явлений сильно колеблются от периода к периоду, например от года к году, то повышаясь, то понижаясь. Особенно это относится ко многим показателям сельского хозяйства, где год на год не приходится: благоприятные годы чередуются с менее благоприятными. Поэтому при анализе развития сельского хозяйства чаще оперируют не годовыми показателями, а более типичными и устойчивыми средне-годовыми показателями за несколько лет, например за пятилетку (см. Продовольственную программу СССР).

При исчислении средних показателей динамики необходимо иметь в виду, что к этим средним полностью относятся общие положения теории средних. Это означает прежде всего, что динамическая средняя будет типичной, если она характеризует период с однородными, более или менее стабильными условиями развития явления. Выделение таких периодов — этапов развития — в определенном отношении аналогично группировке. Если же динамическая средняя исчислена за период, в течение которого условия развития явления существенно менялись, т. е. период, охватывающий разные этапы развития явления, то такой средней нужно пользоваться с большой осторожностью, дополняя ее средними за отдельные этапы.

Средние показатели динамики должны также удовлетворять логико-математическому требованию, согласно которому при за-

мене средней величиной тех фактических величин, из которых исчислена средняя, не должна изменяться величина определяющего показателя, т. е. некоторого обобщающего показателя, связанного с осредняемым показателем.

Метод расчета среднего уровня ряда динамики зависит прежде всего от характера показателя, лежащего в основе ряда, т. е. от вида временного ряда.

Наиболее просто исчисляется *средний уровень интервального ряда абсолютных величин*. Уровни такого ряда, как уже говорилось выше, можно суммировать, получая итоговые уровни за более продолжительные периоды. Вполне логично поэтому исчислять здесь средний уровень так, чтобы при замене фактических уровней их средней величиной не изменялся итоговый уровень за весь период:

$$y_1 + y_2 + \dots + y_n = \bar{y} + \bar{y} + \dots + \bar{y} = n\bar{y}.$$

$\bar{y}$  — средняя уровней

Это приводит к простой средней арифметической:

$$\bar{y} = \frac{\sum y}{n} \quad (14.5)$$

где  $n$  — число фактических уровней за последовательные равные отрезки времени.

Так, сложив годовые уровни производства электроэнергии в СССР за 3 года (1981—1983 гг., см. табл. 14.3) и разделив сумму на 3, найдем, что в 1981—1983 гг. в среднем ежегодно производилось более 1370 млрд. кВт·ч электроэнергии.

Из примера видно, что для среднего уровня интервального ряда динамики должно быть указано: а) за какой конкретный (календарный) период исчислен средний уровень; б) в расчете на какую единицу времени он исчислен (в примере — в расчете на год, однако за те же три года можно было вычислить, например, и среднеквартальный уровень, разделив сумму не на 3 (года), а на 12 (кварталов)). Таким образом, в более общем виде динамический средний уровень интервального показателя равен:

$$\bar{y} = \frac{\text{Уровень за период}}{\text{Продолжительность периода}}.$$

Сложнее обстоит дело с исчислением *среднего уровня моментного ряда динамики абсолютных величин*. Моментный показатель может изменяться почти непрерывно. Поэтому очевидно, что чем более подробными и исчерпывающими данными о его изменении мы располагаем, тем более точно можно исчислить средний уровень. Более того, сам метод расчета зависит от того, насколько подробны имеющиеся у нас данные. Здесь возможны различные случаи.

1. Если известны полные, исчерпывающие данные об изменении моментного показателя, т. е. известно, каков был уровень в начальный момент и когда и как он изменялся, то средний уровень показателя можно исчислить совершенно точно. Выше говорилось, что сумма уровней моментного показателя сама по себе не имеет самостоятельного экономического значения. Поэтому в качестве определяющего показателя и критерия правильности расчета здесь выступает общая сумма произведений количества тех или иных единиц на продолжительность периода их наличия или функционирования ( $\sum y t$ ). Примером может служить общее количество человеко-дней, если определяется средняя численность людей; общее количество станко-дней, если определяется средний размер парка станков и т. п. Только учет тех или иных единиц (объектов) в связи со временем их наличия или функционирования дает такие показатели, которые имеют экономическую значимость и могут быть приняты в качестве определяющих.

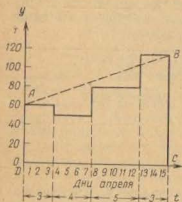


Рис. 14.5. Запас материала на складе

веко-дней ( $34 + 33 + 31 + 32 + 35$ ). Следовательно, в среднем ежедневно работали 33 человека (165 чел.-дней : 5 дней).

Аналогично можно определить, например, среднюю численность всех членов бригады, состоявших в ее списке, — среднюю списочную численность, средний размер запаса материалов на складе и т. п.

Пусть, например, в течение первых трех дней апреля запас материала составлял 60 т, 4 апреля было отпущено со склада 10 т, 8 апреля поступило 30 т, а 13 апреля — еще 35 т (больше никаких изменений запаса в первой половине апреля не было). Графически запас и его изменение можно тогда изобразить ступенчатой линией (рис. 14.5).

Найдем средний размер запаса за первую половину апреля. Так как изменения происходили относительно редко, сложение одинаковых слагаемых (уровней за отдельные дни) целесообразно заменить умножением, а расчет для наглядности можно оформить в виде таблицы (см табл. 14.5).

Отсюда средний запас составляет:  $1125 \text{ тонно-дней} : 15 \text{ дней} = 75 \text{ т}$ .

Таким образом, при наличии исчерпывающих данных об из-

менении моментного показателя его средний уровень исчисляется по формуле средней арифметической взвешенной:

$$\bar{y} = \frac{\sum y t}{\sum t}, \quad (14.6)$$

где  $y$  — уровни, сохранившиеся без изменения в течение времени  $t$ .

2. В большинстве случаев уровень моментного показателя изменяется настолько часто, что практически нецелесообразно учитывать и отражать в рядах динамики все эти изменения. Для практических потребностей оказывается достаточно учитывать уровень раз в месяц, квартал или год. В таких случаях для расчета среднего уровня мы располагаем данными о размере явления только на отдельные даты, и расчет среднего уровня может быть только приближенным, т. е. его приходится делать исходя из некоторых предположений об изменении уровня в промежутках времени между этими датами.

Таблица 14.5. Расчет количества тонно-дней запаса

Календарный период, в течение которого запас не изменялся	Продолжительность этого периода, дн. $t$	Размер запаса тонн $y$	Количество тонно-дней $y t$
1—3	3	60	180
4—7	4	50	200
8—12	5	80	400
13—15	3	115	345
Итого	15	—	1125

Допустим теперь, что нам известны только уровни на начало и конец периода, за который нужно сделать расчет. Так, если в примере с запасом материала на складе нам были бы известны только уровни на 1 апреля — 60 т и на 16 апреля — 115 т (точки А и В на рис. 14.5), то средний размер запаса за первую половину апреля можно было бы определить только приближенно, предполагая то или иное изменение уровня между этими датами. Чаще всего предполагается, что изменение происходило равномерно (в арифметической прогрессии, т. е. с постоянным абсолютным приростом; см. на рис. 14.5 прямую АВ). Тогда число тонно-дней, ранее изображавшееся площадью под ступенчатой линией, изображается площадью трапеции ABCD:  $S_{ABCD} = \frac{AD + BC}{2} \cdot CD$ . Средний запас по-прежнему получим,

разделив количество тонно-дней на длину периода в днях:

$$\bar{y} = \frac{S_{ABCD}}{CD} = \frac{AD + BC}{2} = \frac{y_{\text{нач}} + y_{\text{кон}}}{2} = \frac{60 + 115}{2} = 87,5 \text{ т}.$$

Итак, предположение о равномерном изменении уровня в период между двумя датами приводит к расчету среднего уровня как простой средней арифметической из уровней на начало ( $y_{\text{нач}}$ ) и на конец ( $y_{\text{кон}}$ ) периода:

$$\bar{y} = (y_{\text{нач}} + y_{\text{кон}}) : 2. \quad (14.7)$$



3. Допустим, далее, что, кроме уровней на начало и на конец периода, за который нужно сделать расчет, известны также уровни на некоторые промежуточные даты, промежутки времени между которыми (в общем случае) неравные.

Допустим, что численность жителей нового района города составляла (тыс. чел.): на 1.1—12, на 1.V—13, на 1.XI—15, на 1.I следующего года—18. Общая схема расчета среднего уровня в таких случаях такова. Сначала рассчитываются средние уровни за промежутки времени между двумя соседними датами. Этот расчет делается по формуле (14.7). Затем из полученных таким путем промежуточных средних уровней ( $\bar{y}_i$ ) вычисляется средний уровень за весь период по формуле средней арифметической взвешенной, где весами являются величины промежутков времени между датами:

$$\bar{y} = \frac{\sum \bar{y}_i t_i}{\sum t_i}, \quad (14.8)$$

где  $\bar{y}_i$ —средние уровни за промежутки времени между соседними датами;  $t_i$ —величины этих промежутков.

В нашем примере промежуточные средние уровни равны (тыс. чел.):  $y_{1-IV} = (12+13):2=12,5$ ;  $y_{V-X} = (13+15):2=14$ ;  $y_{XI-XII} = (15+18):2=16,5$ . Промежутки времени между соседними датами равны 4, 6 и 2 месяцам. Средний уровень за год в целом составляет:

$$\bar{y} = \frac{12,5 \cdot 4 + 14 \cdot 6 + 16,5 \cdot 2}{4 + 6 + 2} = 13,9 \text{ тыс. чел.}$$

4. Допустим, теперь, что промежутки времени между соседними датами равны друг другу, т. е. когда мы имеем дело с равными (или примерно равными) интервалами между датами (например, когда известны уровни на начало каждого месяца или квартала, года). Средний уровень ряда можно вычислить и по предыдущей формуле (14.8), однако более целесообразно преобразовать ее:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i} = \frac{1/2(y_1 + y_2)t_1 + 1/2(y_2 + y_3)t_2 + \dots + 1/2(y_{n-1} + y_n)t_{n-1}}{t_1 + t_2 + \dots + t_{n-1}}$$

(здесь  $n$ —число уровней (дат), а число промежутков между датами всегда на 1 меньше).

Учтя, что  $t_1=t_2=\dots=t$ , вынесем этот множитель в числитель за скобки, а знаменатель представим как  $t(n-1)$ . Тогда дробь сократим на  $t$  и получим:

$$\bar{y} = \frac{1/2 y_1 + y_2 + \dots + y_{n-1} + 1/2 y_n}{n-1}, \quad (14.9)$$

где  $y_1$ —уровень на начало периода, за который делается расчет;  $y_n$ —уровень на конец этого периода;  $n$ —число уровней на равностоящие даты, а  $(n-1)$ —число промежутков между датами, численно равное длине периода, выраженной в тех или

иных единицах времени, что полезно иметь в виду для контроля.

Формулу (14.9) часто называют *средней хронологической*. Пусть нужно найти среднюю стоимость основных фондов за года за I квартал на начало месяца по следующим данным, тыс. руб.:

	январь	февраль	март	апрель
Стоимость основных фондов	2 500	2 560	2 599	2 650

Считая, что интервалы между датами приближенно равны, применим формулу (14.9):

$$\bar{y} = \frac{0,5 \cdot 2\,500 + 2\,560 + 2\,599 + 0,5 \cdot 2\,650}{4-1} = 2\,578 \text{ тыс. руб.}$$

В качестве  $y_n$  мы использовали уровень на конец I квартала (на начало II квартала), а  $n-1=3$  измеряет длину квартала в месяцах.

Таблица 14.6. Производство молока и поголовье коров в колхозе

Месяц	Поголовье коров на начало месяца, % годов	Производство молока, т		% выполнения плана	Средний месячный удой молока от одной коровы, кг
		по плану	фактически		
Апрель	494	80,6	83,8	104,0	167,9
Май	504	96,7	96,7	100,0	186,0
Июнь	536	109,0	104,6	96,0	191,9
Итого		286,3	285,1	99,6	182,3

\* Поголовье на начало июля — 554 головы.

При прочих равных условиях (14.7), (14.8) и (14.9) дают тем более точные результаты, чем меньше колеблется уровень и чем короче промежутки времени между датами.

До сих пор речь шла о среднем уровне рядов динамики абсолютных величин. Для рядов динамики средних и относительных величин средний уровень нужно исчислять исходя из содержания и смысла этих средних и относительных показателей (см. гл. 8 и гл. 9). Обозначив относительный (или средний) показатель, являющийся уровнем ряда динамики, через  $k=a:b$ , где  $a$  и  $b$ —абсолютные величины, порядок расчета динамического среднего уровня можно выразить формулой

$$\bar{k} = \bar{a} : \bar{b}. \quad (14.10)$$

Следовательно, сначала производится динамическое осреднение абсолютных величин, а затем путем деления вычисляется средний уровень ряда относительного (или среднего) показателя. Такая последовательность вычисления обеспечивает получение взвешенной средней, тогда как прямое осреднение  $k$  даю бы простую среднюю.

Допустим, по следующим данным нужно найти средний процент выполнения плана производства молока и среднемесечный удой молока от одной коровы за II квартал (см. табл. 14.6).

Среднемесечное производство молока равно (формула 14.5): по плану — 95,4 т, фактически — 95,0 т. Отсюда средний процент выполнения плана (совпадающий с процентом выполнения плана за II квартал в целом) равен:  $(95 : 95,4) \cdot 100 = 99,6 \%$ .

Среднее поголовье найдем по формуле (14.9):  $y = (0,5 \times 494 + 504 + 536 + 0,5 \cdot 554) : 3 = 521$  голова.

Тогда средний месячный удой молока от одной коровы составляет (14.10):  $\bar{k} = 95\,000 : 521 = 182,3$  кг (невзвешенная средняя равна 181,9 кг).

Средний абсолютный прирост показывает, на сколько единиц увеличивался или уменьшался уровень по сравнению с предыдущим в среднем за единицу времени (в среднем ежемесячно, ежегодно и т. д.).

Средний абсолютный прирост характеризует среднюю абсолютную скорость роста (или снижения) уровня и всегда является интервальным показателем. Он вычисляется путем деления общего прироста за весь период на длину этого периода в тех или иных единицах времени:

$$\bar{\Delta} = \frac{\text{Прирост за весь период}}{\text{Длина периода}} = \frac{y_t - y_{t-t}}{t} = \frac{\Sigma \Delta}{t}, \quad (14.11a, б)$$

где  $\Delta$  — цепные абсолютные приросты за последовательные промежутки времени;  $t$  — длина периода (или число цепных приростов); остальные обозначения — прежние.

Особенности исчисления  $\bar{\Delta}$  зависят от вида ряда — моментного или интервального (см. пояснения к формуле 14.1а).

Определим, на сколько киловатт-часов увеличивалось в среднем ежегодно производство электроэнергии в СССР в 1981—1983 гг., т. е. найдем среднегодовой абсолютный прирост за эти три года ( $t=3$ ). При этом  $y_t = y_{83}$ ;  $y_{t-t} = y_{80}$ , отсюда (см. табл. 14.3):  $\Delta_{81-83} = (y_{83} - y_{80}) : 3 = (1\,418 - 1\,294) : 3 = 41,3$  млрд. кВт·ч (либо  $\Delta_{81-83} = (\Delta_{81} + \Delta_{82} + \Delta_{83}) : 3 = (32 + 41 + 51) : 3 = 41,3$ ).

Для контроля полезно иметь в виду, что разность между номерами уровней в числителе формулы (14.1а) равна длине периода, за который производится расчет:  $t = 83 - 80 = 3$ . В этот период не входит базисный год (1980 г.).

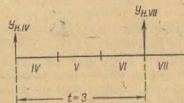
Из примера видно, что, говоря о среднем абсолютном приросте, нужно указывать: а) за какой календарный период ис-

числен (или нужно исчислить) средний прирост; б) в расчете на какую единицу времени он исчислен (в примере — за 1981—1983 гг. в расчете на год).

Для моментных показателей в качестве  $y_t$  в (14.1а) принимается уровень на конец периода, за который делается расчет. Так, если по данным табл. 14.6 нужно найти среднемесечный абсолютный прирост поголовья коров за II квартал, т. е. за 3 месяца (апрель—июнь), то  $t=3$ ;  $y_t = y_{\text{июнь}} = y_{\text{июль}} = 554$ ;  $y_{t-t} = y_{\text{апрель}} = 494$  (контроль:  $t=7-4=3$ ). Тогда  $\Delta_{t-t} = (y_{\text{июль}} - y_{\text{апрель}}) : 3 = (554 - 494) : 3 = 20$  голов, т. е. поголовье коров во II квартале в среднем ежемесячно увеличивалось на 20 голов.

Для моментных показателей контроль облегчается при использовании схемы, аналогичной схеме на 2.70.

В нашем примере:



Для моментных показателей (рядов) длина периода ( $t$ ) в (14.11) может быть выражена в любых единицах времени. При изменении единицы времени будет соответственно изменяться смысл среднего прироста (среднемесечный, среднегодовой и т. п.).

При замене цепных абсолютных приростов их средней величиной, вычисленной по (14.1а, б), не меняется прирост за весь период, выступающий здесь как определяющий показатель. При данном базисном уровне ( $y_{t-t}$ ) в результате такой замены не изменятся и конечный уровень ( $y_t$ ), однако остальные (промежуточные) уровни, рассчитанные на основе  $y_{t-t}$  и  $\bar{\Delta}$ , могут отличаться от фактических. Поэтому и сумма расчетных уровней обычно не совпадает с суммой фактических уровней ряда динамики.

Расчетные уровни ( $y_t'$ ) могут быть вычислены так:  $y_t' = y_{t-t} + \bar{\Delta}t$ , где  $t=1$ ,  $t$  — промежуток времени между базисным годом (месяцем) и годом (месяцем), для которого рассчитывается уровень. При каждом отдельном расчете  $t$  — разность между номером расчетного года и номером базисного года.

В связи с этим в тех случаях, когда сумма уровней ряда динамики имеет особое значение (например, когда в плане задан не конечный уровень планируемого периода (пятилетки и т. п.), а сумма уровней за весь этот период или среднегодовой уровень), среднегодовой абсолютный прирост целесообразно рассчитывать не по формуле (14.1а, б), а иначе — так,



чтобы при замене цепных приростов их средней величиной не изменялась сумма уровней за весь период, для которого делается расчет. Иными словами, в качестве определяющего показателя при расчете среднего прироста целесообразно принять не прирост за весь период, а сумму уровней. Это приводит к такой формуле:

$$\bar{\Delta}_x = \frac{2 \left( \sum_{i=1}^t y_i - y_0 t \right)}{t(t+1)}, \quad (14.11в)$$

где  $y_0$  — базисный уровень;  $t$  — длина периода, за который делается расчет;  $i=1, t$ .

Применение этой формулы покажем на таком примере. Государственные закупки сахарной свеклы в СССР составляли (млн. т.): в 1970 г. — 71,4, в 1971—1980 гг. — 724,1, в том числе в 1980 г. — 65,2. Если использовать обычную формулу (14.11а), то  $\bar{\Delta}_{71-80} = (y_{80} - y_{70}) : 10 = (65,2 - 71,4) : 10 = -0,62$  млн. т. т. е. закупки в 1971—1980 гг. в среднем ежегодно уменьшались на 0,62 млн. т. Однако сумма расчетных уровней при этом равна 679,9 вместо 724,1 фактически. Такие результаты объясняются тем, что конечный 1980 г. был неурожайным, а потому нехарактерным.

Используем (14.11в):

$$\bar{\Delta}_x = \frac{2 \cdot (724,1 - 71,4 \cdot 10)}{10 \cdot 11} = 0,18,$$

т. е. в среднем ежегодно закупки увеличивались почти на 0,2 млн. т.

В качестве основы и критерия правильности исчисления среднего темпа роста (как и среднего абсолютного прироста) также могут быть приняты различные определяющие показатели. Обычно в качестве определяющего показателя применяется произведение цепных темпов роста, равное темпу роста за весь рассматриваемый период.

Выше (параграф 14.2) было показано, что, перемножив  $t$  цепных темпов роста, мы получим темп роста за весь период  $t$  лет:

$$\frac{y_0}{y_1} \cdot \frac{y_1}{y_2} \cdot \dots \cdot \frac{y_{t-1}}{y_t} \cdot \frac{y_t}{y_{t+1}} = \frac{y_{t+1}}{y_1} = \frac{y_t}{y_{t-1}}$$

(если обозначить  $t+1=i$ , откуда  $1=i-t$ ). Следовательно,

$$T_1 \cdot T_2 \cdot \dots \cdot T_t = \frac{y_i}{y_{i-t}}. \quad (14.12)$$

Поставим задачу найти такой средний темп роста ( $\bar{T}_p$ ), чтобы при замене им фактических цепных темпов в формуле

(14.12) остался без изменения темп роста за весь период ( $y_i : y_{i-t}$ ). Следовательно, должно соблюдаться равенство:

$$\underbrace{\bar{T}_p \cdot \bar{T}_p \cdot \dots \cdot \bar{T}_p}_{\text{произведение } t \text{ средних темпов роста}} = \frac{y_i}{y_{i-t}},$$

из которого следует:

$$\bar{T}_p^t = \frac{y_i}{y_{i-t}}.$$

Отсюда

$$\bar{T}_p = \sqrt[t]{\frac{y_i}{y_{i-t}}}; \quad \bar{T}_p = \sqrt[t]{T_1 \cdot T_2 \cdot \dots \cdot T_t}, \quad (14.13а, б)$$

где  $T_1, T_2, \dots, T_t$  — цепные темпы роста.

Формула (14.13б) носит название *простой средней геометрической*, (14.13а) — средняя геометрическая в неясном виде.

Особенности исчисления  $\bar{T}_p$  зависят от вида ряда (см. пояснения к формуле 14.11).

*Средний темп роста, выраженный в форме коэффициента, показывает, во сколько раз увеличивался уровень по сравнению с предыдущим в среднем за единицу времени* (в среднем ежегодно, ежемесячно и т. п.).

Для средних темпов роста и прироста сохраняет силу та же взаимосвязь, которая имеет место между обычными темпами роста и прироста:

$$\bar{T}_{np} = \bar{T}_p - 1 \quad \text{или} \quad \bar{T}_{np}(\%) = \bar{T}_p(\%) - 100 \%, \quad (14.14а, б)$$

*Средний темп прироста (или снижения), выраженный в процентах, показывает, на сколько процентов увеличивался (или уменьшался) уровень по сравнению с предыдущим в среднем за единицу времени* (в среднем ежегодно, ежемесячно и т. п.). Средний темп прироста характеризует среднюю интенсивность роста, т. е. среднюю относительную скорость изменения уровня.

Из двух видов формулы среднего темпа роста чаще используется (14.13а), так как она не требует вычисления всех цепных темпов роста. По формуле (14.13б) расчет целесообразно производить лишь в тех случаях, когда не известны ни уровни ряда динамики, ни темп роста за весь период, а известны только цепные темпы роста (или прироста). Например, прирост производительности общественного труда в СССР составлял: в 8-й пятилетке (1966—1970 гг.) — 39 %, в 9-й пятилетке — 25 %, в 10-й пятилетке — 17 %. Измеряя время ( $t$ ) в 5-летиях ( $t=3$ ), по формуле (14.13б) получим:  $\bar{T}_p = \sqrt[3]{139 \cdot 125 \cdot 117}$ ;  $\lg \bar{T}_p = (\lg 139 + \lg 125 + \lg 117) : 3 = 2,1027$ , откуда  $\bar{T}_p = 126,7 \%$ . Следо-

вательно, производительность общественного труда в 1966—1980 гг. повышалась в среднем за пятилетку на 26,7 %.

Формула (14.13а) позволяет исчислять средний темп роста, если известны: а) базисный ( $y_{t-1}$ ) и конечный ( $y_t$ ) уровни; б) темп роста (или прироста) за весь период ( $y_t : y_{t-1}$ ). Например, в СССР потребление мяса и мясопродуктов на душу населения в 1980 г. составляло 58 кг, а в 1990 г. в соответствии с Продовольственной программой СССР должно составить 70 кг. Найдем среднегодовой темп роста, используя формулу (14.13а):

$t = 1990 - 1980 = 10$ ;  $\bar{T}_p = \sqrt[10]{\frac{y_{90}}{y_{80}}} = \sqrt[10]{\frac{70}{58}}$ . Далее расчет производится либо с помощью логарифмов, либо по специальным таблицам<sup>1</sup>, что более удобно. В нашем случае  $\bar{T}_p = 1,019$ , или 101,9 %;  $\bar{T}_{пр} = 1,9$  %, т. е. потребление мяса и мясопродуктов на душу населения в 1981—1990 гг. должно увеличиваться в среднем ежегодно на 1,9 %.

Другой пример. Производство предметов потребления (группа «Б») в СССР в 1983 г. по сравнению с 1975 г. возросло на 35 %. Здесь  $t = 1983 - 1975 = 8$  и  $\bar{T}_p = \sqrt[8]{\frac{y_{83}}{y_{75}}} = \sqrt[8]{1,35}$ , откуда  $\bar{T}_p = 1,038$  или 103,8 %.

Если уровень ряда динамики снижается, то темп роста за весь период и средний темп роста будут меньше 1 (или 100 %), а средний темп прироста — отрицательной величиной. Так, прямые затраты труда на производство 1 ц зерна (без кукурузы) в колхозах снизились с 1,8 чел.-ч в 1971—1975 гг. до 1,4 чел.-ч в 1976—1980 гг. Темп снижения за 5 лет (1980—1975 = 5; 1976—1971 = 5) равен:  $1,4 : 1,8 = 0,7778$  (снижение на 22,22 %), а среднегодовой темп снижения составляет:  $\bar{T}_{сн} = \sqrt[5]{0,7778} = -1 = 0,951 - 1 = -0,049$ , или -4,9 %.

При замене ценных темпов роста их средней величиной, вычисленной по формуле (14.13), не изменяется темп роста за весь период, а при данном базисном уровне при такой замене не изменяется конечный уровень. Однако промежуточные уровни и сумма всех уровней могут при этом отличаться от фактических.

Так, производство телевизоров в СССР в 1981—1982 гг. в среднем ежегодно увеличивалось в 1,0529 раза, или на 5,29 % ( $\bar{T}_p = \sqrt[2]{8345 : 7926} = 1,0529$ ). Но если бы каждый год производство увеличивалось на 5,29 %, то в 1981 г. было бы произведено 7926 тыс. шт. (7528 · 1,0529) и в 1982 г. — 8346 тыс. шт. (7926 · 1,0529). Расчетный уровень 1981 г. ( $y_{81} = 7926$ ) на 264 тыс. шт. меньше фактического (8190—7926=264). Соответственно и сумма расчетных уровней за 1981—1982 гг. меньше фактической.

Поэтому если сумма уровней ряда по каким-либо причинам (см. пояснение к формуле 14.11а) имеет особое значение, средний темп роста целесообразно рассчитывать так, чтобы при замене ценных темпов их средней

величиной не изменялась сумма уровней за весь период. Это приводит к уравнению

$$y_0 (\bar{T} + \bar{T}^2 + \bar{T}^3 + \dots + \bar{T}^t) = \sum_{i=1}^t y_i,$$

где  $y_0$  — базисный уровень;  $t$  — длина периода, за который производится расчет.

В общем виде это уравнение не решается, но есть специальные таблицы<sup>1</sup>, позволяющие найти  $\bar{T}$ , зная  $y_0$  и  $\sum y_i$ .

В нашем примере получим:  $y_{80}(\bar{T} + \bar{T}^2) = y_{81} + y_{82}$  или  $7528 \cdot (\bar{T} + \bar{T}^2) = 8190 + 8345$ , откуда  $\bar{T}^2 + \bar{T} - 2,19647 = 0$  и  $\bar{T}_p = 1,0641$ .

На разных этапах развития рост или снижение уровня могут происходить с разной скоростью — быстрее или медленнее. Для числовой характеристики ускорения или замедления роста (снижения) уровня можно сопоставлять одноименные показатели динамики за разные этапы развития явления. Если эти этапы имеют одинаковую продолжительность, можно сопоставлять показатели за весь период. Однако часто длина сравниваемых периодов неодинакова, и тогда следует сопоставлять средние показатели динамики. Наиболее выразительные, легко воспринимаемые и поддающиеся истолкованию результаты дает сопоставление средних абсолютных приростов, а также средних темпов прироста. Однако с позиций изучения взаимосвязей целесообразно сопоставление темпов роста (в том числе средних).

Средние абсолютные приросты за отдельные периоды можно сравнивать один с другим путем вычитания ( $\delta = \Delta_2 - \Delta_1$ ), а если они имеют одинаковые знаки, то и путем деления ( $k_\Delta = \Delta_2 : \Delta_1$  либо  $k_\Delta = \Delta_1 : \Delta_2$ ).

Если сравниваемые приросты положительные ( $\Delta > 0$ ), то  $\delta$  характеризует *увеличение* или *уменьшение средней абсолютной скорости*, т. е. *абсолютное ускорение (замедление) прироста*. Если же сопоставляемые приросты отрицательные ( $\Delta < 0$ ), то  $\delta$  характеризует *абсолютное ускорение (замедление) снижения* уровня. Так, по данным табл. 14.7 видно, что обеспеченность населения часами и магнитофонами в 1976—1982 гг. по сравнению с 1971—1975 гг. в абсолютном выражении росла быстрее, т. е. ускоренно: ежегодный прирост обеспеченности часами увеличился на 8,5 шт., а магнитофонами — на 1,3 шт. Обеспеченность же телевизорами росла медленнее, чем в 1971—1975 гг., — ежегодный прирост (средняя абсолютная скорость) уменьшился на 6 шт.

При сравнении средних приростов путем деления  $k_\Delta$  характеризует *относительное ускорение* или *замедление средней скорости* роста. Так, средняя скорость роста обеспеченности часами в 1976—1982 гг. по сравнению с 1971—1975 гг. увеличилась более чем в 1,3 раза, а магнитофонами — почти в 1,3 раза.

<sup>1</sup> См.: Айрапетов А. М. Таблицы исчисления среднегодовых темпов роста, прироста и снижения. — М.: Статистика, 1979.

<sup>1</sup> См., например: Казинцев Л. С. Темпы роста и абсолютные приросты. — М.: Статистика, 1975.



Что касается обеспеченности телевизорами, то ее средняя скорость роста уменьшилась в 1,7 раза (для наглядности и удобства истолкования здесь больший (по модулю) прирост разделен на меньший).

Если уровень ряда динамики уменьшается и средние абсолютные приросты будут меньше нуля (табл. 14.8), то также

Таблица 14.7. Обеспеченность населения СССР предметами длительного пользования (на конец года; в расчете на 1000 человек населения), шт.

	Обеспеченность			$\bar{\Delta}$		$\frac{\Delta}{\bar{\Delta}} = \frac{\Delta_{79-82}}{\Delta_{71-75}}$	$k_{\Delta} = \frac{\bar{\Delta}_{79-82}}{\bar{\Delta}_{71-75}}$
	1970	1975	1982	1971—1975	1976—1982		
Часы	1193	1319	1555	25,2	33,7	8,5	1,34
Магнитофоны	21	46	90	5,0	6,3	1,3	1,26
Телевизоры	143	215	274	14,4	8,4	-6,0	1,71*

\* 14,4 : 8,4 = 1,71.

могут быть найдены  $\delta$  и  $k_{\Delta}$ . Так, потребление картофеля в 1971—1982 гг. снижалось более быстро, чем в 1961—1970 гг. (на 0,4 кг, или в 1,3 раза), а снижение потребления хлебных продуктов в 1971—1982 гг. замедлилось (на 0,5, или в 1,5 раза (табл. 14.8)).

Таблица 14.8. Потребление некоторых продуктов питания в СССР (на душу населения в год)

	1960	1970	1982	$\bar{\Delta}$		$\delta$	$k_{\Delta}$
				1961—1970	1971—1982		
Яйца, шт.	118	159	249	4,1	7,5	3,4	1,83
Картофель, кг	143	130	110	-1,3	-1,7	-0,4	1,31
Хлебные продукты*, кг	164	149	137	-1,5	-1,0	0,5	1,50

\* Хлеб в пересчете на муку, мука, крупа, бобовые, макаронные изделия.

Средние темпы прироста (или снижения) целесообразно сравнивать только путем деления большего (по модулю) темпа на меньший, причем оба сравниваемых темпа должны характеризовать одинаковые по направлению процессы, т. е. иметь одинаковые знаки. Получаемые при этом показатели являются

коэффициентами ускорения (или замедления) средних темпов прироста:

$$\bar{k}_{\text{уск} \bar{T}_{\text{пр}}} = \bar{T}_2 : \bar{T}_1 \text{ (если } |\bar{T}_2| > |\bar{T}_1|); \quad (14.15a)$$

$$k_{\text{зам} \bar{T}_{\text{пр}}} = \bar{T}_1 : \bar{T}_2 \text{ (если } |\bar{T}_1| > |\bar{T}_2|). \quad (14.15b)$$

Так, по данным табл. 14.7 среднегодовые темпы прироста обеспеченности населения составляют (в процентах):

	1971—1975	1976—1982
Часы	2,0	2,4
Магнитофоны	17,0	10,1
Телевизоры	8,5	3,5

Следовательно, обеспеченность населения часами в 1976—1982 гг. росла ускоренными по сравнению с 1971—1975 гг. темпами (коэффициент ускорения равен 1,2 ( $2,4 : 2$ ), а обеспеченность магнитофонами и телевизорами — замедленными (магнитофоны — 1,7, телевизоры — 2,4).

Коэффициенты ускорения (замедления) можно исчислить и для случаев снижения уровня. Так, потребление картофеля (см. табл. 14.8) в 1971—1982 гг. снижалось по сравнению с 1961—1970 гг. ускоренно ( $k_{\text{уск} \bar{T}_{\text{пр}}} = \bar{T}_{\text{пр} (71-82)} : \bar{T}_{\text{пр} (61-70)} = (-1,4) : (-0,9) = 1,6$ ), а потребление хлебных продуктов, наоборот, снижалось замедленно ( $k_{\text{зам} \bar{T}_{\text{пр}}} = \bar{T}_{\text{пр} (61-70)} : \bar{T}_{\text{пр} (71-82)} = (-1,0) : (-0,7) = 1,4$ ).

Как и другие показатели динамики, рассмотренные здесь показатели ускорения и замедления находят применение в практике экономической работы на любом ее уровне.

#### 14.4. ВЫЯВЛЕНИЕ И ХАРАКТЕРИСТИКА ОСНОВНОЙ ТЕНДЕНЦИИ РАЗВИТИЯ

Одна из важнейших задач анализа динамики — выявление и количественная характеристика основной тенденции развития явления. Для этого необходимо выделить такие периоды развития, которые достаточно однородны в отношении взаимосвязи данного явления с другими и в отношении условий его развития. Будем называть такие периоды — *этапами развития явления*.

Выделение этапов развития — это задача, находящаяся на стыке науки, изучающей данное явление (экономики, социологии и т. п.), и статистики. Решение этой задачи осуществляется не только и даже не столько с помощью статистических методов (хотя и они могут принести здесь определенную пользу), сколько на базе содержательного анализа сущности, природы явления и общих законов его развития.

Для каждого этапа развития нужно далее выявить и численно охарактеризовать основную тенденцию изменения уровня явления. Под тенденцией понимается общее направление к росту, снижению или стабилизации уровня явления с течением времени. Если уровень непрерывно растет или непрерывно снижается, то тенденция к росту или снижению является явной и отчетливой: она легко обнаруживается визуально по графику временного ряда. Нужно однако иметь в виду, что и рост, и снижение уровня могут происходить по-разному: либо равномерно, либо ускоренно, либо замедленно. Под *равномерным ростом* (или снижением) здесь понимается рост (снижение)

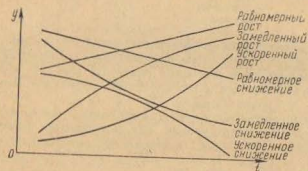


Рис. 14.6. Графическое изображение основной тенденции развития (тренда)

с постоянной абсолютной скоростью, когда цепные абсолютные приросты ( $\Delta_i$ ) одинаковы. При *ускоренном росте* или *снижении* цепные приросты систематически увеличиваются по абсолютной величине, а при *замедленном росте* или *снижении* — уменьшаются (то же по модулю). Однако практически уровни ряда динамики очень редко растут (или снижаются) строго равномерно. Нечасто встречается и систематическое — без единого отклонения — увеличение или снижение цепных приростов. Именно наличие таких отклонений от строгой закономерности и вынуждает нас говорить не просто о росте или снижении уровня, а о его тенденции к росту или снижению.

Такие отклонения могут объясняться тем, что с течением времени может несколько изменяться либо комплекс основных причин и факторов, от которых зависит уровень явления, либо сила их действия, либо внешние условия, в которых происходит развитие явления. Могут также изменяться направление и сила влияния второстепенных, в том числе чисто случайных обстоятельств и факторов. Поэтому при анализе динамики и идет речь не просто о тенденции развития, а об основной тенденции, достаточно стабильной (устойчивой) на протяжении данного этапа развития.

Основной тенденцией (трендом) называется достаточно плавное и устойчивое изменение уровня явления во времени, более или менее свободное от случайных колебаний. Основную тенденцию можно представить либо аналитически — в виде уравнения (модели) тренда, либо графически (рис. 14.6).

Для количественной характеристики явной и отчетливо выраженной тенденции можно использовать средний абсолютный прирост и средний темп прироста, исчисляемые по формулам (14.11а, б; 14.13а, б), если только ценные показатели динамики не имеют от года к году очень резких колебаний. Абсолютный прирост и темпы роста и прироста за весь период (формулы 14.1а; 14.2а; 14.3) можно в этих случаях применять для характеристики общих результатов действия основной тенденции.

Однако, если тенденция недостаточно ясна и отчетлива, так как ценные показатели динамики от года к году резко и сильно колеблются, расчет показателей динамики по указанным выше (обычным) формулам может дать недостаточно точную характеристику основной тенденции и общих результатов ее действия. К еще большим искажениям и ошибкам при характеристике основной тенденции может привести использование обычных формул в тех случаях, когда уровень ряда динамики то повышается, то понижается. Во всех таких случаях при использовании обычных формул всегда есть риск принять за сравнимый ( $y_i$ ) или за базисный ( $y_{i-1}$ ) уровни такие, которые не являются достаточно характерными и типичными. Это хорошо видно на рис. 14.7, где показаны среднегодовые уровни за две пятилетки и уровни базисного (1970 г.) и конечного (1980 г.) года.

Если за базу принять уровень 1970 г., то и прирост за 10 лет, и среднегодовой прирост при использовании обычных формул будут отрицательными (см. рис. 14.7), тогда как из сопоставления средних уровней видна тенденция к росту.

Безусловно, что среднегодовые уровни, при расчете которых в значительной мере устраняются случайные колебания, более характерны и типичны для характеристики основной тенденции, чем годовые уровни, находящиеся под воздействием кратковременных и преходящих обстоятельств и факторов. Поэтому в тех случаях, когда уровень ряда динамики то повышается, то понижается или же изменяется в одном направлении, но очень неравномерно, целесообразно сравнивать и использовать при расчете показателей динамики не годовые, а более

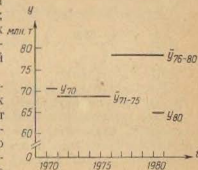


Рис. 14.7. Государственные закупки сахарной свеклы в СССР, млн. т

типичные среднегодовые уровни, изменение которых лучше отражает основную тенденцию развития.

Формулы показателей динамики в этих случаях таковы:

$$\Delta \bar{y} = \bar{y}_{\text{за } t \text{ сравн. лет}} - \bar{y}_{\text{за } t \text{ базисных лет}}; \quad (14.16)$$

$$T_p(\bar{y}) = \frac{\bar{y}_{\text{за } t \text{ сравн. лет}}}{\bar{y}_{\text{за } t \text{ базисных лет}}}; \quad (14.17)$$

$$\bar{\Delta y} = \frac{\Delta \bar{y}}{t}; \quad \bar{T}_p(\bar{y}) = \sqrt[t]{T_p(\bar{y})}. \quad (14.18a, б)$$

Так, в нашем примере (см. рис. 14.7) получим:

$$\bar{\Delta y} = \frac{76,9 - 67,9}{5} = 1,8 \text{ млн. т}; \quad \bar{T}_p(\bar{y}) = \sqrt[5]{\frac{76,9}{67,9}} = 1,025,$$

откуда  $T_{np}(\bar{y}) = 2,5\%$ .

Следовательно, среднегодовые закупки свеклы увеличивались в среднем ежегодно на 1,8 млн. т, или на 2,5%.

Отметим, что показатели роста производства продукции сельского хозяйства, роста производительности труда в сельском хозяйстве и др. рассчитываются именно путем сопоставления среднегодовых уровней.

При расчетах по формуле (14.18) происходит усреднение  $\Delta$  и  $T_p$  скользящих средних уровней, о которых будет сказано ниже. Если по какому-либо причинам необходимо использовать в расчетах средние уровни за неравные по продолжительности периоды ( $t_1 \neq t_2$ ), то длина периода ( $t$ ) в (14.18) принимается равной простой средней арифметической из этих величин:  $\bar{t} = (t_1 + t_2) : 2$ .

Возможен и иной более углубленный подход к характеристике основной тенденции, при котором используются специальные способы ее выявления. Эти способы применяются прежде всего в тех случаях, когда тенденция развития из-за колебаний уровня оказывается как бы затухающей и недостаточно отчетливой, а также тогда, когда необходимо точное ее измерение. При этом выявление и характеристика основной тенденции могут представлять собой не только самостоятельную задачу, но преследовать и иные цели: выделить в динамическом ряду регулярные, систематические колебания, накладывающиеся на тренд (например, сезонные колебания), построить прогнозную модель и т. п. В статистике используются различные приемы и способы выявления и характеристики основной тенденции — и элементарные, и более сложные.

Рассмотрим способ *укрупнения интервалов* и их характеристики средними уровнями. Этот способ заключается в переходе от интервалов менее продолжительных к более продолжительным, например от суток — к неделям или декадам, от декад — к месяцам, от месяцев — к кварталам или годам, от годовых интервалов — к многолетним. Если уровни ряда динамики колеблются с более или менее определенной периодично-

стью (волнообразно), то укрупненный интервал целесообразно взять равным периоду колебаний (длине «волны» цикла). Если же такая периодичность отсутствует, то укрупнение производится постепенно от малых интервалов ко все более крупным, пока общее направление тренда не станет достаточно отчетливым.

Пусть, например, имеются следующие данные:

Таблица 14.9. Производство зерна в СССР

Год	Валовой сбор, млн. т	Год	Валовой сбор, млн. т	Год	Валовой сбор, млн. т
1966	171	1971	181	1976	224
1967	148	1972	168	1977	196
1968	170	1973	223	1978	237
1969	162	1974	196	1979	179
1970	187	1975	140	1980	189
Итого в среднем за год	838 167,6	Итого в среднем за год	908 181,6	Итого в среднем за год	1025 205,0

Как видно из табл. 14.9 итоги по пятилетиям уже обнаруживают отчетливую тенденцию к росту. Но добавление еще двух-трех лет могло бы привести к несопоставимости с пятилетними уровнями. Поэтому для получения более или менее сравнимых показателей целесообразно перейти от пятилетних уровней к средним уровням (см. последнюю строку табл. 14.9).

Если ряд динамики является моментным, а также в тех случаях, когда уровень ряда является относительной или средней величиной, суммирование уровней не имеет смысла, и также следует охарактеризовать укрупненные периоды средними уровнями.

При укрупнении интервалов число членов динамического ряда сильно сокращается, в результате чего движение уровня внутри укрупненного интервала выпадает из поля зрения. В связи с этим для выявления основной тенденции и более детальной ее характеристики используется сглаживание ряда с помощью скользящей средней.

*Сглаживание* ряда динамики с помощью *скользящей средней* заключается в том, что вычисляется средний уровень из определенного числа первых по порядку уровней ряда, затем — средний уровень из такого же числа уровней, начиная со второго, далее начиная с третьего и т. д. Таким образом, при расчетах среднего уровня как бы «скользят» по временному ряду от его начала к концу, каждый раз отбрасывая один уровень в начале и добавляя один следующий. Отсюда название — *скользящая средняя*.



Каждое звено скользящей средней — это средний уровень за соответствующий период. При графическом изображении и при некоторых расчетах каждое звено принято условно относить к центральному интервалу того периода, за который сделан расчет (для моментного ряда — к центральной дате). Если, например, интервальный ряд динамики за 1966—1980 гг. сглаживается с помощью 3-летней скользящей средней, то первое звено (за 1966—1968 гг.) следует отнести к 1967 г. и т. д. Если каж-



Рис. 14.8. Валовой сбор зерна в СССР

тер: цепные абсолютные приросты звеньев центрированной скользящей средней ( $\Delta = \bar{y}_t - \bar{y}_{t-1}$ ) сначала возрастают, а затем убывают.

Вопрос о том, за какой период следует вычислять звенья скользящей средней, решается в зависимости от конкретных особенностей динамики. Как и при укрупнении интервалов, если в колебаниях уровня есть определенная периодичность, то период сглаживания целесообразно принять равным периоду колебаний или кратной ему величине. Так, при наличии квартальных уровней, испытывающих ежегодно сезонные спады и повышения, целесообразно применять четырех- или восьмиквартальную среднюю и т. п. Если же колебания уровней являются беспорядочными, то целесообразно постепенно укрупнять интервал сглаживания, пока не выявится отчетливая картина тренда.

Уменьшение числа звеньев скользящей средней по сравнению с числом исходных уровней ряда несколько сужает, конечно, возможности изучения характера выявленной тенденции в начале и в конце этапа развития. Тем не менее скользящая средняя обладает достаточной гибкостью, позволяющей все же уловить особенности изменения тенденции. Однако скользящая средняя не дает аналитического выражения тренда.

Таблица 14.10. Сглаживание ряда динамики с помощью 8-летней скользящей средней и ее центрирование

Восьмилетия	Валовой сбор зерна, млн. т		Центрированная 8-летняя скользящая средняя $\bar{y}_t$	Центральный год двух 8-летних	$\Delta = \bar{y}_t - \bar{y}_{t-1}$
	за 8 лет	в среднем за год			
1966—1973	1410	176,3			
1967—1974	1435	179,4	177,8*	1970	...
1968—1975	1427	178,4	178,9	1971	1,1
1969—1976	1481	185,1	181,8	1972	2,9
1970—1977	1515	189,4	187,3	1973	5,5
1971—1978	1565	195,6	192,5	1974	5,2
1972—1979	1563	195,4	195,5	1975	3,0
1973—1980	1584	198,0	196,7	1976	1,2

\*  $(176,3 + 179,4) : 2$  и т. д.

Аналитическое выравнивание ряда динамики позволяет получить аналитическую модель тренда. Оно производится следующим образом.

1. На основе содержательного анализа выделяется этап развития явления (см. с. 289) и устанавливается характер динамики на этом этапе.

2. Исходя из предположения о той или иной закономерности роста и из характера динамики выбирается форма аналитического выражения тренда, вид аппроксимирующей функции, которой графически соответствует определенная линия — прямая, парабола, показательная кривая и т. п. Эта линия (функция) выражает предполагаемую закономерность плавного изменения уровня во времени, т. е. основную тенденцию. При этом каждый уровень ряда динамики условно рассматривается как сумма двух составляющих (компонент):  $y_t = f(t) + e_t$ . Одна из них ( $\bar{y}_t = f(t)$ ), выражающая тренд, характеризует влияние постоянно действующих, основных факторов и называется систематической

регулярной компонентой. Другая составляющая ( $\varepsilon_t$ ) отражает влияние случайных факторов и обстоятельств и называется *случайной компонентой*. Эту компоненту называют также *остаточной* (или просто *остатком*), так как она равна отклонению фактического уровня от тренда. Таким образом, допускается (условно предполагается), что основная тенденция (тренд) формируется под влиянием постоянно действующих главных факторов, а второстепенные, случайные факторы вызывают отклонение уровня от тренда.

Выбор формы тренда (вида кривой  $\hat{y}_t = f(t)$ ) практически редко можно сделать на основе одного только содержательного анализа. Обычно на 1-м этапе выбора отбирают функции, пригодные с позиций содержательного анализа, а на 2-м этапе вид функции конкретизируется с помощью иных подходов и приемов, имеющих эмпирический характер.

Наиболее простой эмпирический прием — визуальный: выбор формы тренда на основе графического изображения ряда — ломаной линии. В случае очень сильных и резких колебаний уровня целесообразно использовать график скользящей средней. Нередко, однако, ни график уровней, ни график скользящей средней не могут дать ответ об оптимальной форме тренда. В таких случаях целесообразен анализ цепных абсолютных приростов и темпов прироста (включая их сглаживание с помощью скользящей средней).

Если *цепные абсолютные приросты* относительно стабильны, не имеют отчетливой тенденции к росту или снижению, т. е. если уровень явления изменяется с достаточно постоянной абсолютной скоростью ( $\Delta_t \approx \text{const}$ ), то в качестве *формы тренда* нужно принять *прямую линию (линейную функцию)*:  $\hat{y}_t = a_0 + a_1 t$ . Если же относительно стабильными являются *цепные темпы прироста*, т. е. если уровень явления растет с более или менее постоянной относительной скоростью ( $T_t \approx \text{const}$ ), то в качестве формы тренда следует принять *показательную кривую* ( $\hat{y}_t = a_0 a_1^t$ ).

В тех же случаях, когда *цепные абсолютные приросты* более или менее равномерно увеличиваются (или уменьшаются), т. е. если уровень ряда динамики изменяется с равномерно возрастающей (или убывающей) абсолютной скоростью, в качестве формы тренда (аппроксимирующей функции) можно принять *параболу второй степени* ( $\hat{y}_t = a_0 + a_1 t + a_2 t^2$ ).

Однако и этот второй подход к выбору формы тренда, т. е. анализ цепных показателей динамики, не всегда приводит к достаточно уверенному выбору конкретной формы тренда (особенно в случае, если у цепных абсолютных приростов есть слабо выраженная тенденция к увеличению или снижению). Поэтому в тех случаях, когда требуется особо точное измерение тенденции (например, если предполагается использование модели тренда для прогнозирования), при выборе вида кривой воз-

можно также использование специальных критериев математической статистики, о которых будет сказано ниже.

3. После выбора вида кривой вычисляются ее параметры. Расчет параметров обычно производится *методом наименьших квадратов*. Это означает, что ставится и решается задача: из множества кривых данного вида найти ту, которая обращает в минимум сумму квадратов отклонений фактических уровней динамического ряда от соответствующих им во времени выравненных (расчетных) уровней, т. е. уровней, лежащих на искомой кривой:

$$\sum (y_t - \hat{y}_t)^2 = \min, \quad (14.19)$$

где  $y_t$  — фактические,  $\hat{y}_t$  — выравненные (расчетные) уровни.

Параметры кривой ( $a_i$ ), удовлетворяющей этому условию, могут быть найдены путем решения системы нормальных уравнений (см. гл. 13).

4. На основе найденного уравнения кривой (модели тренда) вычисляются выравненные уровни. Таким образом, технически выравнивание ряда динамики заключается в замене фактических уровней такими плавными изменяющимися уровнями ( $\hat{y}_t$ ), которые в среднем менее всего<sup>1</sup> отклонялись бы от фактических и имели бы определенное аналитическое выражение, соответствующее общему направлению и характеру тренда.

Рассмотрим технику *выравнивания* ряда динамики по *прямой*:  $\hat{y}_t = a_0 + a_1 t$ . Параметры  $a_0$  и  $a_1$  искомой прямой, удовлетворяющие методу наименьших квадратов, находятся путем решения такой системы нормальных уравнений:

$$\left. \begin{aligned} a_0 n + a_1 \sum t &= \sum y_t \\ a_0 \sum t + a_1 \sum t^2 &= \sum y_t t \end{aligned} \right\}$$

где  $t$  — время (порядковый номер интервала или момента времени).

Расчет параметров значительно упрощается, если за начало отсчета ( $t=0$ ) времени принять центральный интервал (или момент). При нечетном числе уровней значения  $t$  устанавливаются так, как это сделано в тр. 2 табл. 14.11. Если же количество уровней четное, значения  $t$  будут такими (это равнозначно измерению времени не в годах, а в полугодиях)

1975	1976	1977	1978	1979	1980	1981	1982
-7	-5	-3	-1	1	3	5	7

В обоих случаях  $\sum t = 0$ , так что система уравнений принимает вид  $a_0 n = \sum y_t$  и  $a_1 \sum t^2 = \sum y_t t$ , откуда

$$a_0 = \sum y_t / n; \quad a_1 = \sum y_t t / \sum t^2. \quad (14.20a, б)$$

<sup>1</sup> В смысле наименьшей средней квадратической величины отклонений.

Пронлюстрируем выравнивание ряда динамики по прямой на следующем примере (табл. 14.11). В порядке первого приближения здесь можно принять линейную форму тренда (см. рис. 14.8).

Подставляя в формулу (14.20) итоговые суммы из табл. 14.11, получим:

$$a_0 = 2771 : 15 = 184,73 \approx 184,7; \quad a_1 = 840 : 280 = 3,0.$$

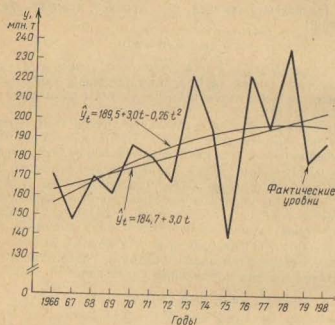


Рис. 14.9. Валовой сбор зерна в СССР

Отсюда уравнение искомой прямой таково:

$$\hat{y}_t = 184,7 + 3t.$$

Подставляя в это уравнение значения  $t$  из табл. 14.11, найдем выравненные (расчетные) уровни (см. гр. 5, табл. 14.11). Графическое изображение найденного тренда показано на рис. 14.9.

Аналитическое выравнивание ряда динамики не только делает более четким направление основной тенденции, но одновременно дает также числовую ее характеристику. В частности, при выравнивании по прямой параметр  $a_1$  — это абсолютный прирост выравненного уровня за единицу времени  $t$ , или средний абсолютный прирост с учетом тенденции к равномерному росту (росту в арифметической прогрессии). Так, в нашем при-

мере  $a_1 = 3$  означает, что выравненный валовой сбор ежегодно увеличивался на 3 млн. т или с учетом тенденции к равномерному росту сбор зерна в среднем ежегодно возрастал на эту величину.

Таблица 14.11. Расчетная таблица при выравнивании по прямой и по параболе ряда динамики производства зерна в СССР

Год	Валовой сбор зерна, млн. т $y$	Обозначение времени $t$	$at$	$t^2$	Выравненные (расчетные) уровни $\hat{y}_t = a_0 + a_1 t$	$t^4$	$at^3$
A	1	2	3	4	5	6	7
1966	171	-7	-1197	49	163,7	2401	8 379
1967	148	-6	-888	36	166,7	1296	5 328
1968	170	-5	-850	25	169,7	625	4 250
1969	162	-4	-648	16	172,7	256	2 592
1970	187	-3	-561	9	175,7	81	1 683
1971	181	-2	-362	4	178,7	16	724
1972	168	-1	-168	1	181,7	1	168
1973	223	0	0	0	184,7	0	0
1974	196	1	196	1	187,7	1	196
1975	140	2	280	4	190,7	16	560
1976	224	3	672	9	193,7	81	2 016
1977	196	4	784	16	196,7	256	3 136
1978	237	5	1185	25	199,7	625	5 925
1979	179	6	1074	36	202,7	1296	6 444
1980	189	7	1323	49	205,7	2401	9 261
Итого	2771	0	840	280	2770,5	9352	50 662

При четном числе уровней ежегодный прирост будет равен  $2a_1$  ( $a_1$  тогда характеризует прирост за половину исходного интервала).

При выравнивании ряда динамики по показательной кривой  $\hat{y}_t = a_0 a_1^t$  параметр  $a_1$  характеризует темп роста выравненного уровня за единицу времени  $t$ , или средний темп роста с учетом тенденции к росту с постоянной относительной скоростью. Пример выравнивания по показательной кривой приведен в 14.7.

Показательная кривая путем логарифмирования приводится к линейному виду относительно  $\lg y_t$ ,  $\lg a_0$  и  $\lg a_1$ , а именно:  $\lg y_t = \lg a_0 + t \lg a_1$ .

Поэтому техника выравнивания по показательной кривой аналогична технике выравнивания по прямой с той лишь разницей, что после расчета  $\lg a_0$ ,  $\lg a_1$  и  $\lg y_t$  параметры  $a_0$ ,  $a_1$  и выравненные уровни определяются путем потенцирования.



При выравнивании ряда динамики по параболе второй степени ( $y_t = a_0 + a_1 t + a_2 t^2$ ) ее параметры (при  $\sum t = 0$ ) рассчитываются так:

$$\left. \begin{aligned} a_1 &= \frac{\sum y t}{\sum t^2}; & a_0 n + a_2 \sum t^2 &= \sum y; \\ a_0 \sum t^2 + a_2 \sum t^4 &= \sum y t^2. \end{aligned} \right\} \quad (14.21a, б)$$

Абсолютная скорость роста выравненного уровня в этом случае с течением времени изменяется (она равна  $\hat{y}'_t = a_1 + 2a_2 t$  — первой производной  $\hat{y}_t$  по  $t$ ). При нечетном числе уровней за единицу времени  $t$  скорость увеличивается (или уменьшается) на  $2a_2$ . Таким образом, в этом случае  $2a_2$  — это ускорение (или замедление), т. е. прирост скорости. При четном числе уровней и при  $\sum t = 0$  ускорение (замедление) равно  $8a_2$ .

Если производство зерна в СССР выравнивать по параболе второй степени, то  $a_1 = 840 : 280 = 3$ , а система уравнений (14.21б) такова:

$$\left. \begin{aligned} 15a_0 + 280a_2 &= 2771; \\ 280a_0 + 9352a_2 &= 50662. \end{aligned} \right\}$$

Это дает:  $a_0 = 189,54 \approx 189,5$ ;  $a_2 = -0,258 \approx -0,26$ . Следовательно, уравнение параболы таково:  $\hat{y}_t = 189,5 + 3,0t - 0,26t^2$ . Подставляя в него значения  $t$ , найдем выравненные уровни. Так, для 1966 г. ( $t = -7$ ) получим:  $\hat{y}_{66} = 189,5 + 3 \cdot (-7) - 0,26 \cdot 49 = 155,9$  и т. д. (см. табл. 14.12 и рис. 14.9).

Средняя скорость изменения уровня (как и для прямой) здесь 3 млн. т. а ежегодное ее уменьшение, т. е. замедление, равно 0,52 млн. т. Следовательно, найденная параболa характеризует замедленный рост.

Отклонения фактических уровней ряда динамики от выравненных (расчетных) используются для характеристики колеблемости фактических уровней около тренда. Абсолютным показателем этой колеблемости является среднее квадратическое отклонение:

$$\sigma_{\text{ост}} = \sqrt{\frac{\sum (y - \hat{y}_t)^2}{n}}. \quad (14.22)$$

Относительной мерой колеблемости является модифицированный коэффициент вариации:

$$v = \frac{\sigma_v}{\bar{y}} \cdot 100\%. \quad (14.23)$$

В нашем примере колеблемость сбора зерна около линейного тренда составляет 23,2 млн. т. или 12,6 %, а около параболического тренда — 22,8 млн. т. или 12,4 % (расчет сумм квадратов остаточных отклонений показан в табл. 14.12).

Суммы квадратов остаточных отклонений можно также использовать для выбора формы тренда с помощью статистических критериев. В качестве

грубого критерия иногда применяют суммы квадратов отклонений, скорректированные на числа степеней свободы:  $\hat{S}_e^2 = \sum (y - \hat{y}_t)^2 : (n - m)$ , где  $n$  — число уровней,  $m$  — число параметров тренда. Выбирается тот тренд, для которого  $\hat{S}_e^2$  меньше. Однако более надежные результаты дает использование  $F$ -критерия Фишера, основанного на процедуре проверки гипотез. Проверяется нулевая гипотеза о том, что уменьшение  $\hat{S}_e^2$  за счет включения в уравнение тренда члена  $a_2 t^2$  статистически незначимо. Нами разработана основная на критерии  $F$  таблица, в которой приведены критические значения

Таблица 14.12. Расчет сумм квадратов остаточных отклонений

Год	y	Линейный тренд		Параболический тренд		
		$y - \hat{y}_t$	$(y - \hat{y}_t)^2$	$\hat{y}_t$	$y - \hat{y}_t$	$(y - \hat{y}_t)^2$
1966	171	7,3	53,29	155,9	15,1	228,01
1967	148	-18,7	349,69	162,3	-14,3	204,49
1968	170	0,3	0,09	168,1	1,9	3,61
1969	162	-10,7	114,49	173,4	-11,4	129,96
1970	187	11,3	127,69	178,3	8,7	75,69
1971	181	2,3	5,29	182,5	-1,5	2,25
1972	168	-13,7	187,69	186,3	-18,3	334,89
1973	223	38,3	1466,89	189,5	33,5	1122,25
1974	196	8,3	68,89	192,3	3,7	13,69
1975	140	-50,7	2570,49	194,5	-54,5	2970,25
1976	224	30,3	918,09	196,2	27,8	772,84
1977	196	-0,7	0,49	197,4	-1,4	1,96
1978	237	37,3	1391,29	198,1	38,9	1513,21
1979	179	-23,7	561,69	198,3	-19,3	372,49
1980	189	-16,7	277,89	197,9	-8,9	79,21
Итого	2771	0,5	8094,95	2770,0	0	7824,80

величины  $k = S_e^2$  (лин.) :  $S_e^2$  (пар.), где  $S_e^2$  (лин.) и  $S_e^2$  (пар.) — суммы квадратов остаточных отклонений соответственно для линейного тренда и параболы второй степени. Если эмпирическое значение  $k$  больше табличного, то нулевая гипотеза отклоняется при данном уровне значимости  $\alpha$ , т. е. параболический тренд предпочтительнее, чем линейный. Если же фактическое значение  $k$  меньше табличного, то нулевая гипотеза не отклоняется, т. е. следует предпочесть линейный тренд.

Критические значения  $k = S_e^2$  (лин.) :  $S_e^2$  (пар., 2ст.).

n	5	6	7	8	9	10	11	12	13	14	15	16
$\alpha = 5\%$	10,26	4,38	2,93	2,32	2,00	1,80	1,67	1,57	1,50	1,44	1,40	1,36
$\alpha = 2\%$	25,26	7,87	4,51	3,26	2,65	2,28	2,05	1,88	1,76	1,67	1,60	1,54

В нашем примере  $k_0 = 8094,95 : 7824,8 = 1,035$ , что меньше табличного  $k$  (для  $n = 15$ ) и при  $\alpha = 2\%$  (1,60), при  $\alpha = 5\%$  (1,40). Поэтому можно считать, что линейный тренд предпочтительнее параболы второй степени (криволинейность не улучшает форму тренда).

Отметим, что если расчеты делаются на ЭВМ, то парабола второй степени предпочтительнее, чем линейный тренд, если статистически значим параметр  $a_2$  (по  $t$ -критерию Стьюдента).

Некоторые общественные явления и процессы подвержены *сезонным колебаниям*, т. е. их уровень из года в год в определенные месяцы повышается, а в другие — понижается. Эти внутригодовые колебания, имеющие более или менее регулярный характер, и называются сезонными. Особенно наглядно сезонные колебания проявляются в сельском хозяйстве, но они имеют место и в других отраслях.

Сезонные подъемы и спады часто наносят ущерб народному хозяйству, вызывая неравномерное в течение года использование трудовых ресурсов и производственных мощностей, что нередко ведет к снижению производительности труда, повышению себестоимости продукции и т. п. Поэтому обычно стремятся ликвидировать либо уменьшить сезонные колебания.

Отсюда вытекает необходимость изучения сезонности и измерения сезонных колебаний. Такое измерение позволяет судить об эффективности мер по ликвидации или уменьшению сезонности. Вместе с тем измерение сезонных колебаний необходимо также для того, чтобы при внутригодовом планировании предвидеть и учесть такие сезонные колебания, которые пока еще не удается устранить. Механическая разбивка годового плана по кварталам и месяцам могла бы привести к серьезным неувязкам.

Существуют различные способы выявления и измерения сезонных колебаний. Их мерой обычно является *индекс сезонности* — процентное отношение фактического уровня за тот или иной месяц (квартал) к выравненному уровню за тот же период:

$$I_{\text{сез}} = (y_{\text{факт}} : y_{\text{выравн}}) \cdot 100 \% \quad (14.24)$$

Поскольку в каждом году сезонные колебания могут быть смешаны с чисто случайными, индексы сезонности исчисляются обычно за несколько лет и находят их среднюю величину за одноименные месяцы ( $I_{\text{сез}} = \sum I_{\text{сез}} : n$ , где  $n$  — число лет).

В зависимости от отсутствия или наличия отчетливой тенденции к росту (снижению) уровня от года к году выравненный уровень по формуле (14.24) вычисляется по-разному. Если тенденции нет (или она слабо выражена), в качестве выравненного уровня можно принять среднюю арифметическую из месячных уровней за все годы. В этом случае в числителе индекса также можно брать средние уровни за одноименные месяцы. Это сразу дает средние индексы сезонности. Если же уровень явления из года в год существенно повышается (снижается), следует предварительно выделить тренд — либо путем сглаживания с помощью 12-месячной (4-квартальной) скользящей средней, либо путем аналитического выравнивания. Выравненные уровни (т. е. либо звенья скользящей средней, либо  $\hat{y}_t$ ) и

используются тогда при расчете индексов сезонности по формуле (14.24). Используются и другие методы выравнивания уровней, в частности по ряду Фурье.

Измерим сезонные колебания числа заключенных браков в СССР в 1979—1981 гг. (табл. 14.13).

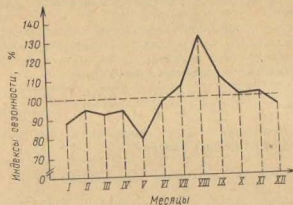


Рис. 14.10. Сезонные колебания числа браков в СССР (средние индексы сезонности за 1979—1981 гг.; %)

Выравнивание месячных данных по прямой ( $\sum t = 0$ ) показывает тенденцию к некоторому росту:  $\hat{y}_t = 33,17 + 0,0884t$ . Расчетные уровни и индексы сезонности по годам показаны в табл.

Таблица 14.13. Расчет индексов сезонности числа заключенных браков в СССР

Месяц	Число браков, тыс.			$\hat{y}_t = 233,17 + 0,0884t$			$I_{\text{сез}}, \%$			$I_{\text{сез}}, \%$
	1979	1980	1981	1979	1980	1981	1979	1980	1981	
I	199	190	223	230,1	232,2	234,3	86,5	81,8	95,1	88
II	224	208	229	230,2	232,4	234,5	97,2	89,5	97,6	95
III	221	201	219	230,4	232,5	234,7	95,8	86,4	93,2	92
IV	211	250	196	230,6	232,7	234,8	91,4	107,3	83,5	94
V	180	184	188	230,8	232,9	235,0	80,0	79,0	80,0	80
VI	248	215	232	231,0	233,1	235,2	107,3	92,1	98,6	99
VII	259	243	243	231,1	233,3	235,4	112,0	104,1	103,1	106
VIII	300	320	316	231,3	233,4	235,6	129,6	137,0	134,1	111
IX	283	246	253	231,5	233,6	235,7	122,2	105,2	107,2	102
X	236	226	251	231,7	233,8	235,9	101,8	96,6	106,3	103
XI	259	241	222	231,8	234,0	236,1	111,6	102,9	93,9	103
XII	261	201	216	232,0	234,1	236,3	112,4	85,9	91,3	96
Итого	2881	2725	2788	2772,5	2798,0	2823,5	—	—	—	1200

14.13. В последней графе этой таблицы даны средние индексы за соответствующие месяцы, округленные до целых процентов.

Размер и характер сезонных колебаний особенно наглядны при графическом изображении (рис. 14.10). На графике хорошо видна сезонная «волна» с пиком в августе, когда число браков на 34 % выше среднего, и спадом в мае (на 20 % ниже среднего уровня).

В качестве обобщающих показателей, позволяющих судить о том, уменьшились или усилились сезонные колебания, могут служить либо среднее линейное, либо среднее квадратическое отклонение индексов сезонности от 100 %:

$$\sigma = \sqrt{\frac{\sum (I_{\text{сез}} - 100)^2}{12}} \quad (14.25)$$

Эти же показатели можно использовать для сравнения сезонных колебаний разных явлений.

#### 14.5. АНАЛИЗ НЕСКОЛЬКИХ РЯДОВ ДИНАМИКИ

Сравнительный анализ нескольких рядов динамики занимает важное место в изучении развития социально-экономических явлений. При этом могут ставиться задачи сопоставления: а) динамики одного и того же явления в разных регионах; б) динамики явления в целом и его составных частей; в) динамики различных явлений, характеризующих одну сторону или область производства, обслуживания и т. п. (например, производство различных товаров культурно-бытового назначения); г) динамики различных явлений, между которыми существует либо функциональная, либо корреляционная связь.

При сравнительном анализе одноименных явлений (пункты «а» и «б») можно сопоставлять не только направление и характер основных тенденций и относительные показатели динамики (темпы развития), но также и абсолютные величины — уровни, абсолютные приросты, абсолютные ускорения.

Сравним, например, производство искусственных и синтетических волокон в СССР (см. табл. 14.14).

В 1960 г. синтетические волокна и нити составляли только 7 % общего производства. Их было произведено в 13 раз меньше, чем искусственных. За 10 лет (1961—1970) производство искусственных волокон возросло на 260 тыс. т (в среднем ежегодно — на 26 тыс. т), тогда как производство синтетических волокон — только на 152 тыс. т (в среднем ежегодно — на 15,2 тыс. т). Однако по темпам роста производство синтетических волокон опережало производство искусственных: если производство искусственных увеличилось в 2,3 раза (в среднем ежегодно — на 8,8 %), то производство синтетических волокон возросло в 11 раз (в среднем ежегодно — на 27,3 %). В результате в 1970 г. производство синтетических волокон составляло уже около 27 % общего производства (вместо 7 % в 1960 г.).

За последующие 13 лет (1971—1983 гг.) производство искусственных волокон увеличилось на 189 тыс. т, или на 41,4 % (в среднем ежегодно — на 2,9 %), а производство синтетических — на 541 тыс. т, или в 4,2 раза (в среднем ежегодно — на 11,8 %). Таким образом, производство синтетических волокон опережало производство искусственных уже не только по темпам роста, но и по абсолютному приросту. Прирост производства синтетических волокон составил почти  $\frac{3}{4}$  общего прироста производства химических волокон (541 из 730 тыс. т).

Таблица 14.14. Производство химических волокон и нитей в СССР

Год	Производство всего, тыс. т	в том числе		Среднегодовой абсолютный прирост, тыс. т		Среднегодовой темп прироста, %	
		искусственные	синтетические	искусственные	синтетические	искусственные	синтетические
1960	211	196	15				
1970	623	456	167	26,0	15,2	8,8	27,3
1983	1353	645	708	14,5	41,6	2,9	11,8

а само их производство — более половины общего производства. В 1971—1983 гг. по сравнению с 1961—1970 гг. среднегодовой абсолютный прирост производства синтетических волокон увеличился более чем в 2,7 раза, тогда как по искусственным волокнам он снизился почти вдвое.

Для одноименных явлений при характеристике процесса опережения можно сравнивать любые показатели динамики, исчисленные за один и тот же период. Абсолютные приросты (за весь период или среднегодовые) можно сравнивать либо путем вычитания, либо путем деления (в последнем случае сравниваемые приросты должны иметь одинаковые знаки). Коэффициент, показывающий, во сколько раз абсолютный прирост одного явления больше, чем прирост другого явления, может быть назван коэффициентом опережения по абсолютному приросту, или коэффициентом абсолютного опережения:

$$K_{\text{асс, опер}} = \Delta_A : \Delta_B, \quad (\Delta_A > \Delta_B), \quad (14.26)$$

где А и Б — сравниваемые части целого или сравниваемые регионы.

Аналогично можно сравнивать и средние абсолютные приросты. Так, в нашем примере в 1961—1970 гг. (см. табл. 14.14) по средней абсолютной скорости производство искусственных волокон опережало производство синтетических волокон в 1,7 раза (26:15,2). В последующие же 13 лет (1971—1983 гг.) по средней абсолютной скорости производство синтетических



волокон опережало производство искусственных волокон почти в 3 раза (41,6 : 14,5 = 2,9).

Темпы развития целесообразно сравнивать только путем деления большего из них на меньший. Коэффициент, показывающий, во сколько раз один темп больше другого, называется *коэффициентом опережения по темпам* (роста или прироста), или *коэффициентом относительного опережения*:

$$K_{\text{отн. опер}} = T_A : T_B, \quad (T_A > T_B). \quad (14.27)$$

При этом сравниваемые темпы должны характеризовать одинаковую по направлению тенденцию. Аналогично можно сравнивать среднегодовые темпы роста и прироста. Так, в нашем примере производство синтетических волокон опережало производство искусственных волокон по средней относительной скорости в 1961—

Таблица 14.15. Темпы роста производительности труда в промышленности СССР и США (в расчете на одного работающего)

Год	$T_p$ 1960 г. = 100 %		$\bar{T}_{пр}$ , %	
	СССР	США	СССР	США
1960	100	100		
1970	166	137	5,2	3,2
1982	272	169	4,2	1,7

1970 гг. более чем в 3 раза (27,3 : 8,8 = 3,1), а в 1971—1983 гг.— в 4 раза (11,8 : 2,9 = 4,1).

Если анализируется динамика различных явлений, то можно сопоставлять только направление и характер основных тенденций и относительные показатели динамики (темпы развития и коэффициенты ускорения или замедления).

Ряды динамики, уровни которых непосредственно не сопоставимы, обычно

приводят к *одному основанию*, т. е. выражаются в виде *базисных* темпов роста. На основе этих темпов роста можно вычислить и сопоставить не только коэффициенты ускорения темпов, но и коэффициенты ускорения абсолютных приростов. Можно для таких рядов вычислить и коэффициенты относительного опережения.

Рассмотрим, например, следующие данные (см. табл. 14.15).

Абсолютные уровни производительности труда здесь неизвестны. Однако можно найти абсолютные приросты, измеряя их в так называемых *пунктах роста*, т. е. в процентах базисного уровня. Так, в СССР абсолютный прирост за 1961—1970 гг. составил 66 пунктов (166—100), или 6,6 пункта ежегодно. В 1971—1982 гг. абсолютный прирост составил уже 106 пунктов (272—166), или 8,8 пункта в среднем ежегодно. Следовательно, среднегодовой абсолютный прирост ускорился более чем в 1,3 раза (8,8 : 6,6 = 1,33), тогда как в США этот показатель снизился с 3,7 пункта до 2,7 пункта. В 1961—1970 гг.

СССР опережал США по среднегодовой относительной скорости в 1,62 раза (5,2 : 3,2), а в 1971—1982 гг.— в 2,47 раза.

При сравнительном анализе динамики часто строят ряды динамики производных показателей — относительных и средних величин. Так, сопоставляя динамику производства зерна и посевных площадей зерновых культур, можно получить динамический ряд средней урожайности зерновых и т. п. Это позволяет измерять роль отдельных факторов в изменении зависящего от них показателя (см. гл. 15).

#### 14.6. РЕГРЕССИЯ И КОРРЕЛЯЦИЯ В РЯДАХ ДИНАМИКИ

Если содержательный социально-экономический анализ показывает, что между уровнями двух или нескольких рядов динамики может существовать корреляционная зависимость, то возникает задача исследования этой зависимости, ее количественной характеристики с помощью показателей регрессии и корреляции. Полученные результаты могут быть использованы как для анализа влияния факторов на результирующий показатель, так и для прогнозирования. Однако применение методов анализа регрессии и корреляций к рядам динамики связано с определенными трудностями и имеет поэтому ряд особенностей.

Первая трудность возникает в связи с тем, что в социально-экономических рядах динамики обычно имеет место та или иная тенденция (тренд), обусловленная действием постоянных факторов. Это приводит к тому, что последующие уровни ряда зависят от предыдущих, т. е. в ряду динамики имеется *автокорреляция и авторегрессия*. Так, например, размер поголовья скота в данном году зависит, естественно, от его размера в предшествующие годы; аналогичная зависимость имеется и между объемом производства (или потребления) определенного продукта и т. п.

Автокорреляция уровней ряда динамики приводит к нарушению предпосылок и требований, которым должна удовлетворять информация (исходные данные) при регрессионно-корреляционном анализе и использовании в нем метода наименьших квадратов. Здесь имеется в виду прежде всего нарушение *независимости наблюдений друг от друга* (см. гл. 13). В результате наличия автокорреляции и авторегрессии в каждом изучаемом ряду или хотя бы одним из них при коррелировании уровней этих рядов показатели корреляции и регрессии могут быть искажены: они будут отражать не только влияние изменения признака-фактора (т. е. колебаний уровня соответствующего ряда динамики около тренда), но и влияние автокорреляции и авторегрессии в обоих коррелируемых рядах. При этом может оказаться «выявленной» зависимости между явлениями, которой в действительности нет, или, наоборот, пре-

уменьшена или преувеличена зависимость, которая объективно существует.

Чтобы избежать подобных искажений и измерить влияние колебаний признака-фактора ( $x$ ) на колебания результативного признака ( $y$ ) в более или менее чистом виде, нужно проверить наличие автокорреляции в каждом временном ряду и — если она будет обнаружена — каким-то путем устранить или ослабить ее.

Для проверки наличия или отсутствия автокорреляции уровней ряда динамики обычно вычисляется коэффициент автокорреляции 1-го порядка, измеряющий тесноту связи между соседними уровнями ряда (зависимости данного уровня от одного предыдущего). Этот коэффициент вычисляется по обычным формулам (13.14), в которых в качестве одного признака, принимается уровень данного периода (например, года), а в качестве другого — предыдущий уровень того же ряда динамики. Иными словами, вычисляется коэффициент корреляции между рядом  $y_1, y_2, \dots, y_n$  и рядом  $y_2, y_3, \dots, y_{n+1}$ , сдвинутым по отношению к первому ряду на один период (год):

$$r_{y_t y_{t+1}} = \frac{(n-1) \sum y_t y_{t+1} - \sum y_t \cdot \sum y_{t+1}}{\sqrt{[(n-1) \sum y_t^2 - (\sum y_t)^2][(n-1) \sum y_{t+1}^2 - (\sum y_{t+1})^2]}}, \quad (14.28)$$

где  $n$  — число уровней в первоначальном ряду.

Если второй ряд динамики сдвинуть по отношению к первому на  $t$  периодов (лет), то при вычислении коэффициента корреляции между такими рядами получим коэффициент автокорреляции порядка  $t$ .

Найдем, например, коэффициент автокорреляции 1-го порядка для ряда государственных закупок подсолнечника (табл. 14.16); суммы для расчета взяты из строки «Итого», т. е. без данных за 1983 г.

$$r_{y_t y_{t+1}} = \frac{14 \cdot 6028 - 280 \cdot 293}{\sqrt{(14 \cdot 5814 - 280^2)(14 \cdot 6321 - 293^2)}} = 0,836.$$

Аналогичный расчет для ряда производства подсолнечника дает  $r_{x_t x_{t+1}} = 0,615$ . Таким образом, в обоих динамических рядах есть довольно высокая положительная автокорреляция. Поэтому показатели корреляции и регрессии, вычисленные непосредственно для уровней этих рядов ( $y_x = -23,3 + 1,65x$ ;  $r_{yx} = 0,973$ ), дают неправильное представление о зависимости ежегодных колебаний объема закупок от колебаний объема производства подсолнечника. Эти показатели искажены автокорреляцией, наличием тенденции к росту в обоих рядах. Ее влияние нужно элиминировать или хотя бы ослабить.

Есть различные методы исключения тенденции в рядах динамики. Первый способ заключается в том, что исследуется зависимость не между уровнями ряда, а между их разностями. Если в динамическом ряду есть тенденция к равномерному росту (снижению) уровней, т. е. по прямой, с постоянной аб-

солютной скоростью, то вместо уровней для построения уравнения регрессии и расчета коэффициента корреляции используются ценные абсолютные приросты (первые разности); если же в ряду есть тенденция к росту (снижению) уровня с равно-

Таблица 14.16. Производство и государственные закупки подсолнечника в регионе, тыс. т

Год	Производство $x_t$	Закупки $y_t$	$y_t + 1$	$y_t y_{t+1}$	$y_t^2$	$y_{t+1}^2$	$\sum x_t^2$	$\sum y_t^2$	$\sum x_t'$	$\sum y_t'$	$\sum x_t'^2$	$\sum y_t'^2$
1969	22	13	16	208	169	256	484	286	—7	—154	—91	—96
1970	24	16	14	224	256	196	576	384	—6	—144	—36	—144
1971	23	14	19	266	196	361	529	322	—5	—115	—70	—76
1972	26	19	17	323	361	289	676	494	—4	—104	—76	—104
1973	24	17	19	323	289	361	576	408	—3	—72	—51	—51
1974	26	19	20	380	361	400	676	494	—2	—52	—38	—38
1975	25	20	22	440	400	484	625	500	—1	—25	—25	—20
1976	28	22	20	440	484	400	784	616	0	0	0	0
1977	26	20	22	440	400	484	676	520	1	26	20	20
1978	28	22	22	484	484	484	784	616	2	56	44	44
1979	27	22	25	550	484	625	729	594	3	81	66	66
1980	30	25	24	600	625	576	900	750	4	120	100	100
1981	28	24	27	648	576	729	784	672	5	140	120	120
1982	30	27	26	702	729	676	900	810	6	180	162	162
Итого	367	280	293	6028	5814	6321	—	—	—	—	—	—
1983	29	26	—	—	676	—	841	754	7	203	182	—
Всего (1969—1983 гг.)	396	306	—	—	6490	—	10 540	8220	0	140	252	—

мерно возрастающей или убывающей абсолютной скоростью, т. е. по параболе второй степени, то следует использовать приросты цепных приростов, т. е. вторые разности.

В нашем примере в обоих рядах имеется четкая тенденция к росту по прямой (рис. 14.11), поэтому следует исследовать зависимость между цепными абсолютными приростами. Расчеты (табл. 14.17) дают такие результаты:

$$\begin{aligned} 14a_0 + 7a_1 &= 13; & \text{откуда } a_0 &= 0,5824; & a_1 &= 0,6923; \\ 7a_0 + 59a_1 &= 62; & \Delta_y &= 0,58 + 0,69\Delta x; \end{aligned}$$

$$r_{\Delta y \Delta x} = \frac{14 \cdot 62 - 7 \cdot 13}{\sqrt{(14 \cdot 79 - 13^2)(14 \cdot 59 - 7^2)}} = 0,911.$$

Таким образом, и коэффициент корреляции, и особенно коэффициент регрессии ( $a_1=0,69$ ), оказались меньше, чем при коррелировании уровней. С увеличением абсолютного прироста производства ( $\Delta_x$ ) на 1 тыс. т, абсолютный прирост закупок ( $\Delta_y$ ) увеличивается в среднем на 0,7 тыс. т (а не на 1,65, как

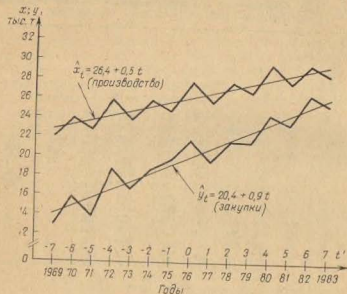


Рис. 14.11. Производство и государственные закупки подсолнечника в регионе

было при коррелировании уровней — см. выше). 83 % общей дисперсии приростов закупок объясняется вариацией приростов объема производства ( $0,911^2=0,83$ ).

Отметим также, что в ряду  $\Delta_x$  коэффициент автокорреляции 1-го порядка равен  $(-0,12)$  (вместо 0,62 для уровней), а в ряду  $\Delta_y$  соответственно 0,05, т. е. переход от уровней к цепным приростам свел автокорреляцию почти на нет.

Второй способ элиминирования тенденции — переход от коррелирования уровней к коррелированию отклонений от трендов — остатков ( $y_t - \hat{y}_t = e_t$ ). Предполагается, что в рядах отклонений от трендов ( $e_t$ ) автокорреляции нет или она значительно уменьшена и не существенна (незначима). Однако это обстоятельство требует проверки.

Дело в том, что причинами наличия автокорреляции в остатках (отклонения фактических уровней от рассчитанных по той или иной модели) могут быть: а) не включение в модель существенного фактора, оказывающего систематическое влияние на зависимый (результативный) признак; б) неправильный выбор формы связи (для рядов динамики — формы тренда); в) некоторые другие причины.

Таблица 14.17. К расчету  $r_{\Delta y \Delta x}$  и критерия  $d$  для ряда  $x$

Год	$x_t - \hat{x}_t = e_t$	$y_t - \hat{y}_t = e_t$	$\Delta y_t$	$\Delta x_t$	$\Delta y_t^2$	$\Delta x_t^2$	Для ряда $x$				
							$\hat{x}_t$	$\hat{y}_t$	$e_t$	$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$
							$\hat{x}_t$	$\hat{y}_t$	$e_t$	$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$
1969	...	...	...	...	...	...	22,9	0,9	1	—	0,81
1970	2	3	6	4	9	16	23,4	0,6	1,5	2,25	0,36
1971	-1	-2	2	1	4	1	23,9	-0,9	-1,5	2,25	0,81
1972	-3	-5	15	9	25	81	24,4	1,6	2,5	6,25	2,56
1973	-2	-2	4	4	4	16	24,9	-0,9	-2,5	6,25	0,81
1974	-2	-2	4	4	4	16	25,4	0,6	1,5	2,25	0,36
1975	-1	1	1	1	1	1	25,9	-0,9	-1,5	2,25	0,81
1976	3	2	6	9	4	81	26,4	1,6	2,5	6,25	2,56
1977	-2	-2	4	4	4	16	26,9	-0,9	-2,5	6,25	0,81
1978	2	2	4	4	4	16	27,4	0,6	1,5	2,25	0,36
1979	-1	0	0	1	0	1	27,9	-0,9	-1,5	2,25	0,81
1980	3	3	9	9	9	81	28,4	1,6	2,5	6,25	2,56
1981	-2	-1	2	4	1	16	28,9	-0,9	-2,5	6,25	0,81
1982	2	3	6	4	9	16	29,4	0,6	1,5	2,25	0,36
1983	-1	-1	1	1	1	1	29,9	-0,9	-1,5	2,25	0,81
Итого	7	13	62	59	79	396,0	0	—	55,50	15,60	

Наличие автокорреляции остатков приводит к тому, что оценки показателей регрессии и корреляции не обладают оптимальными статистическими свойствами, в связи с чем затрудняется проверка их значимости.

Проверка наличия и существенности автокорреляции остатков производится обычно с помощью критерия Дарбина — Уотсона:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}, \quad (14.29)$$

где  $e_t = y_t - \hat{y}_t$  — остатки (отклонения от тренда или от уровней, рассчитанных по какой-либо модели).

Вычисленное (эмпирическое) значение  $d$  сопоставляется с табличными значениями при принятом уровне значимости  $\alpha$ , данном  $n$  (числе уровней ряда динамики или числе наблюдений) и  $m'$  (числе независимых переменных (факторов) в модели (уравнении тренда или регрессии)). Для линейного тренда  $m'=1$ , а для параболы второй степени  $m'=2$ .

Приведем некоторые табличные значения критерия  $d$  при  $\alpha=5\%$  (односторонняя проверка). Особенностью критерия  $d$  является указание двух критических значений: нижнего ( $d_n$ ) и верхнего ( $d_u$ ), между которыми находится область неопределенности, где критерий  $d$  не работает, т. е. с его



помощью нельзя прийти ни к какому решению о значимости (существенности) автокорреляции.

Следующая схема поясняет применение критерия Дарбина — Уотсона (числа в скобках — применительно к нашему примеру:  $n=15$ ;  $m'=1$ ;  $\alpha=5\%$ ).

0	$d_n$ (1,08)	$d_{\beta}$ (1,36)	2	$4-d_{\beta}$ (2,64)	$4-d_n$ (2,92)	4
Есть значимая положительная автокорреляция	?	Автокорреляция не существенна (при принятом $\alpha$ )	?	Автокорреляция не существенна (при принятом $\alpha$ )	?	Есть значимая отрицательная автокорреляция

Найдем уравнение тренда для каждого ряда в нашем примере. Для производства подсолнечника ( $x$ ; см. табл. 14.16) получим:  $a_0 = \Sigma x : n = 396 : 15 = 26,4$ ;  $a_1 = \Sigma x t' : \Sigma (t')^2 = 140 : 280 = 0,5$ . Отсюда  $x_t = 26,4 + 0,5t$ . Аналогичные расчеты для объема

Таблица 14.18. Критерий Дарбина — Уотсона

Число наблюдений (уровней ряда динамики)	$m' = 1$		$m' = 2$	
	$d_n$		$d_{\beta}$	
	$d_n$	$d_{\beta}$	$d_n$	$d_{\beta}$
15	1,08	1,36	0,95	1,54
20	1,20	1,41	1,10	1,54

закупок дают:  $\hat{y}_t = 20,4 + 0,9t$ . Найдем далее  $\hat{x}_t$  (см. табл. 14.17),  $\hat{y}_t$  (табл. 14.19), отклонения от них фактических уровней и критерий  $d$  для каждого ряда. При расчете  $d$  величины  $(e_t - e_{t-1})$  находим так (см. стр. 2 табл. 14.19):  $e_t - e_{t-1} = 1,0 - (-1,1) = 2,1$  и т. д.

Для остатков  $y_t - \hat{y}_t = e_t$  критерий  $d$  составляет:  $d = 66,94 : 20,8 = 3,22$ . Это значение  $d$  больше, чем  $(4 - d_n) = 2,92$ , что свидетельствует о значимой (при  $\alpha = 5\%$ ) отрицательной автокорреляции (см. схему).

Действительно, если рассчитать коэффициент автокорреляции 1-го порядка не для уровней, а для остатков  $e_t = y_t - \hat{y}_t$ , то он оказывается равным  $(-0,677)$ . Следовательно, после исключения из ряда  $y$  линейного тренда (т. е. при переходе от  $y_t$  к  $y_t - \hat{y}_t = e_t$ ) автокорреляция сохранилась, изменив знак (для ряда  $y$  коэффициент автокорреляции равнялся 0,836).

Для остатков  $x_t - \hat{x}_t = e_t$  получим:  $d = 55,5 : 15,6 = 3,56$ , т. е. и здесь есть значимая отрицательная автокорреляция.

Таким образом, мы столкнулись здесь со случаем, когда исключение тренда не привело к существенному ослаблению автокорреляции остатков. Это объясняется, очевидно, тем, что на колебания объема закупок  $(e_t = y_t - \hat{y}_t)$  существенное систе-

Таблица 14.19. К расчету критерия  $d$  для ряда  $y$ ,  $e_{y,t}$  и  $e_{x,t}^2$

Год	Для ряда $y$					Для ряда $x$ *				
	$y_t$	$e_t = y_t - \hat{y}_t$	$e_t^2$	$e_t - e_{t-1}$	$(e_t - e_{t-1})^2$	$e_t$	$e_{y,t}$	$\hat{y}_t$	$y - \hat{y}$	$(y - \hat{y})^2$
1969	14,1	-1,1	1,21	—	—	-0,9	0,99	13,2	-0,2	0,04
1970	15,0	1,0	1,00	2,1	4,41	0,6	0,60	15,6	0,4	0,16
1971	15,9	-1,9	3,61	-2,9	8,41	-0,9	1,71	15,0	-1,0	1,00
1972	16,8	2,2	4,84	4,1	16,81	1,6	3,52	18,4	0,6	0,36
1973	17,7	-0,7	0,49	-2,9	8,41	-0,9	0,63	16,8	0,2	0,04
1974	18,6	0,4	0,16	1,1	1,21	0,6	0,24	19,2	-0,2	0,04
1975	19,5	0,5	0,25	0,1	0,01	-0,9	-0,45	18,6	1,4	1,96
1976	20,4	1,6	2,56	1,1	1,21	1,6	2,56	22,0	0	0
1977	21,3	-1,3	1,69	-2,9	8,41	-0,9	1,17	20,4	-0,4	0,16
1978	22,2	-0,2	0,04	1,1	1,21	0,6	-0,12	22,8	-0,8	0,64
1979	23,1	-1,1	1,21	-0,9	0,81	-0,9	0,99	22,2	-0,2	0,04
1980	24,0	1,0	1,00	2,1	4,41	1,6	1,60	25,6	-0,6	0,36
1981	24,9	-0,9	0,81	-1,9	3,61	-0,9	0,81	24,0	0	0
1982	25,8	1,2	1,44	2,1	4,41	0,6	0,72	26,4	0,6	0,36
1983	26,7	-0,7	0,49	-1,9	3,61	-0,9	0,63	25,8	0,2	0,04
Итого	306,0	0	20,80	—	66,94	0	15,60	306,0	0	5,20

\* См. табл. 14.17.

матическое влияние оказывают колебания объема производства ( $e_t = x_t - \hat{x}_t$ ), что проявляется в почти полной их синхронности (см. рис. 14.11). Коэффициент корреляции между отклонениями от трендов (остатками  $e_y = y_t - \hat{y}_t$  и  $e_x = x_t - \hat{x}_t$ ) вычисляется так<sup>1</sup>:

$$r_{e_y e_x} = \frac{\Sigma e_y e_x}{\sqrt{\Sigma e_y^2 \cdot \Sigma e_x^2}}. \quad (14.30)$$

В нашем примере получим  $(\Sigma e_y e_x$  и  $\Sigma e_y^2$  — из табл. 14.19;  $\Sigma e_x^2$  — из табл. 14.17):

$$r_{e_y e_x} = \frac{15,6}{\sqrt{20,8 \cdot 15,6}} = \sqrt{0,75} = 0,866.$$

Это значительно меньше, чем  $r_{yx} = 0,973$  и  $r_{\Delta y \Delta x} = 0,911$ .

В уравнении линейной регрессии  $e_y = a_0 + a_1 e_x$  параметр  $a_0$  всегда равен нулю, так как  $\Sigma e_y = \Sigma e_x = 0$ , а параметр  $a_1 = \frac{\Sigma e_y e_x}{\Sigma e_x^2}$ .

<sup>1</sup> Формула (14.30) проще, чем обычная формула коэффициента корреляции и формула  $r_{\Delta y \Delta x}$ , так как  $\Sigma e_x = \Sigma e_y = 0$ , что не имеет места в других случаях.

В нашем примере  $a_1 = 15,6; 15,6 = 1$ , т. е. отклонение от тренда объема закупок ( $e_y$ ) увеличивается в среднем на столько же, на сколько возрастает отклонение объема производства ( $e_x$ ):  $e_y = e_x$ . Различие значений параметров  $a_1$ , а также коэффициентов корреляции при коррелировании цепных приростов ( $\Delta$ ) и коррелировании отклонений от тренда ( $\epsilon$ ) объясняется тем, что эти показатели имеют различный статистический смысл: в одном случае имеется в виду отклонение от предыдущего уровня ( $\Delta = y_t - y_{t-1}$ ), а в другом — отклонение от тренда ( $e_y = y_t - \hat{y}_t$ ). Большой интерес представляют, очевидно, колебания уровней около тренда, их отклонения от основной тенденции, поэтому предпочтительнее коррелирование отклонений (остатков).

Точно такие же результаты, как и коррелирование отклонений от тренда, дает включение фактора времени ( $t$ ) в уравнение регрессии<sup>1</sup>:

$$\hat{y}_{xt} = a_0 + a_1 x + a_2 t. \quad (14.31)$$

Параметры этой регрессии вычисляются путем решения системы уравнений:

$$\left. \begin{aligned} a_0 n + a_1 \sum x + a_2 \sum t &= \sum y; \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum tx &= \sum yx; \\ a_0 \sum t + a_1 \sum xt + a_2 \sum t^2 &= \sum yt. \end{aligned} \right\}$$

Если начало отсчета времени перенести в центр ряда динамики, то  $\sum t = 0$  и система упрощается:

$$\left. \begin{aligned} a_0 n + a_1 \sum x &= \sum y; \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum t'x &= \sum yx; \\ a_1 \sum t' + a_2 \sum (t')^2 &= \sum y t'. \end{aligned} \right\}$$

В нашем примере в табл. 14.16 и 14.19 уже есть все необходимые суммы (кроме  $\sum (t')^2$ , которая равна 280):

$$\left. \begin{aligned} 15a_0 + 396a_1 &= 306; \\ 396a_0 + 10540a_1 + 140a_2 &= 8220; \\ 140a_1 + 280a_2 &= 252. \end{aligned} \right\}$$

Решение дает:  $a_0 = -6$ ;  $a_1 = 1$ ;  $a_2 = 0,4$ , откуда  $\hat{y}_{xt} = -6 + x + 0,4t$ .

Как и в  $\hat{e}_y = e_x$ , параметр  $a_1 = 1$ . Параметр  $a_2 = 0,4$  показывает, что если исключить влияние объема производства ( $x$ ) на

<sup>1</sup> Если хотя бы в одном из рядов тренд выражается параболой второй степени, в уравнение регрессии нужно включить также  $+a_3 t^2$ .

объем заготовок ( $y$ ), то объем заготовок имеет тенденцию к ежегодному росту на 0,4 тыс. т (за счет уменьшения потерь и потребления в хозяйствах производителей).

Тесноту связи следует здесь измерять с помощью *частного коэффициента корреляции между  $y$  и  $x$  при устранении влияния  $t$*  (см. также формулу 13.24):

$$r_{yx(t)}^2 = \frac{\sigma_{yx}^2 - \sigma_{yt}^2}{\sigma_y^2 - \sigma_{yt}^2}, \quad (14.32)$$

где  $\sigma_y^2$  — общая дисперсия  $y$ ;  $\sigma_{yt}^2$  — факторная дисперсия для модели тренда  $\hat{y}_t = a_0 + a_1 t$ ;  $\sigma_{yx}^2$  — факторная дисперсия для регрессии  $\hat{y}_{xt} = a_0 + a_1 x + a_2 t$ .

Расчеты дисперсий произведем по модифицированным рабочим формулам, необходимые для которых суммы есть в табл. 14.6—14.19:

$$\begin{aligned} \sigma_y^2 &= \bar{y}^2 - \bar{y}^2 = (6490 : 15) - (20,4)^2 = 16,5067; \\ \sigma_{yt}^2 &= \sigma_y^2 - \frac{\sum (y - \hat{y}_t)^2}{n} = 16,5067 - \frac{20,8}{15} = 15,12; \\ \sigma_{yx}^2 &= \sigma_y^2 - \frac{\sum (y - \hat{y}_{xt})^2}{n} = 16,5067 - \frac{5,2}{15} = 16,16. \end{aligned}$$

Тогда  $r_{yx(t)}^2 = \frac{16,6 - 15,12}{16,5067 - 15,12} = 0,750$ , откуда  $r_{yx(t)} = \sqrt{0,75} = 0,866$ , что совпадает с  $r_{e_y e_x}$ .

Зависимость колебаний уровней одного ряда динамики от колеблемости уровней другого имеет еще одну существенную особенность. Она заключается в том, что изменения одного явления нередко отстают по времени от изменений другого, запаздывают по сравнению с ними на тот или иной промежуток времени, который может измеряться месяцами и годами. Так, колебания в числе родившихся, особенно первенцев, отстают от колебаний в числе браков; прирост продукции может отставать от прироста основных производственных фондов, так как требуется определенное время для освоения нового оборудования и доведения новых сооружений до полной проектной мощности; колебания размера поголовья скота запаздывают по сравнению с колебаниями урожайности и объема заготовок кормов и т. п.

Отставание (запаздывание) колебаний одного ряда динамики от колебаний другого называется *лагом*. Если такое отставание имеет место, то, естественно, его нужно учесть при построении уравнений регрессии и расчета коэффициентов корреляции, сдвигая один ряд по отношению к другому на величину лага, например  $\hat{y}_{xt} = a_0 + a_1 x_{t-1} + a_2 t$ ;  $r_{y_{t-1} x_t}(n)$ . Положение осложняется при этом тем, что точная величина лага обычно

известна, и, кроме того, она может не совпадать точности интервалами, к которым относятся уровни ряда динамики. Величину лага пужно определять на основе содержательного анализа, однако нередко ее определяют эмпирически, вычисляя коэффициенты корреляции без лага и с различными по величине лагами. Максимальное значение коэффициента корреляции указывает ориентировочную приближенную величину лага. Если в регрессию включено несколько факторов, то они могут иметь различный лаг, например  $y_{t+1} = a_0 + a_1 x_{1t} + a_2 x_{2t-1} + \dots + a_n x_{nt} + a_d$ .

Зависимость колебаний уровней одного временного ряда от колебаний другого может изменяться с течением времени. Поэтому для более глубокого ее изучения может быть построен ряд уравнений регрессии и коэффициентов корреляции за различные периоды. При этом показатели регрессии и корреляции могут быть получены как за примыкающие один к другому периоды, так и за перекрывающиеся периоды — так, как это делается при расчетах скользящей средней (например, 1963—1967 гг., 1964—1968 гг. и т. д.). Выбор интервала сглаживания должен производиться с учетом условий развития явлений.

#### 14.7. ЭКСТРАПОЛЯЦИЯ И ИНТЕРПОЛЯЦИЯ РЯДОВ ДИНАМИКИ

Изучение закономерностей развития социально-экономических явлений, выявление и характеристика трендов и моделей взаимосвязи создают базу для *прогнозирования*, т. е. для *определения ориентировочных размеров явлений в будущем*. В связи с развитием и совершенствованием планирования и управления вопросы прогнозирования становятся особенно актуальными. На июньском (1983 г.) Пленуме ЦК КПСС подчеркивалась необходимость ориентации на будущее, с учетом назревающих тенденций, что позволяет надежно прогнозировать, принимать обоснованные решения<sup>1</sup>.

Важное место в системе методов прогнозирования занимают статистические методы. Любой метод прогнозирования предполагает, что та или иная закономерность развития, действовавшая в прошлом, сохранится и в прогнозируемом будущем, т. е. прогноз основан на *экстраполяции (распространении)* этой закономерности на будущее. Поэтому надежность и точность прогноза зависят от того, насколько близкими к действительности окажутся эти предположения, а также от того, насколько точно удалось охарактеризовать выявленную в прошлом закономерность.

Особое значение при прогнозировании методом экстраполяции имеют вопросы о ее базе и сроках, т. е. о длине базис-

ного периода, характерная для которого закономерность будет распространена на будущее, и о длине будущего периода, на который эту закономерность можно распространить. В качестве базы для экстраполяции нельзя, очевидно, брать ни очень короткий, ни слишком длинный период. Короткий период может оказаться недостаточно типичным из-за действия временных, преходящих условий и факторов, а излишне длинный период нецелесообразен потому, что на его протяжении могли значительно измениться условия развития явления, и то, что было характерно для динамики 15—20 лет назад, может стать совершенно нехарактерным в будущем. Оптимальная по длине база экстраполяции должна быть выбрана на основе содержательного анализа с учетом конкретных исторических условий развития. Что же касается *срока экстраполяции (периода предупреждения прогноза)*, то чем он короче, тем более надежные и точные результаты (при прочих равных условиях) дает прогноз, так как за короткий срок не успевают сильно измениться условия развития явления и характер его динамики.

Наиболее простым методом статистического прогнозирования является экстраполяция на будущее средней абсолютной или относительной скорости изменения уровня ( $\bar{\Delta}$  или  $\bar{T}_{\text{пр}}$ ), т. е. метод, основанный на предположении о равномерном изменении уровня (под равномерностью здесь понимается стабильность либо  $\bar{\Delta}$ , либо  $\bar{T}$ ). Если в базисном периоде ценные показатели динамики не имели резких колебаний и  $\bar{\Delta}$  и  $\bar{T}_p$  рассчитывались по формулам (14.11) и (14.13), экстраполяцию можно делать с помощью таких формул:

$$y_{i+t} = y_i + \bar{\Delta}t; \quad y_{i+t} = y_i (\bar{T}_p)^t, \quad (14.33a, б)$$

где  $y_{i+t}$  — экстраполируемый уровень ( $i+t$ ) — номер этого уровня (года);

$i$  — номер конечного уровня (года) базисного периода, за который рассчитаны  $\bar{\Delta}$  или  $\bar{T}_p$ ;

$t$  — срок прогноза (период предупреждения).

Так, если за 1971—1983 гг. (базисный период;  $i=83$ ) был вычислен среднегодовой прирост протяженности автомобильных дорог с твердым покрытием в СССР ( $\Delta_{71-83}=21,8$  тыс. км), то при сохранении этой скорости протяженность дорог в 1985 г. должна составлять:  $y'_{85} = y_{83} + \bar{\Delta}t = 773 + 21,8 \cdot 2 = 816,6$  тыс. км.

Если ценные показатели динамики в базисном периоде сильно колебались, и ежегодный прирост (или темпы роста и прироста) вычислялись на базе средних уровней (рассчитанных за равные периоды), то для экстраполяции можно использовать аналогичные формулы:

$$\bar{y}_{i+t} = \bar{y}_i + \bar{\Delta}_p t; \quad \bar{y}_{i+t} = \bar{y}_i (\bar{T}_p(\bar{y}))^t, \quad (14.34)$$

где  $\bar{y}_{i+t}$  — экстраполируемый средний уровень;

<sup>1</sup> См.: Материалы Пленума ЦК КПСС, 14—15 июня 1983 г.—М.: Политиздат, 1983, с. 34.



$i$  — номер последнего года, учтенного при расчете  $\bar{\Delta}_y$  или  $T_y$ ;

$t$  — длина периода упреждения.

Так, среднегодовые уровни производства мяса в СССР составляли (млн. т):  $y_{61-70} = 10,45$ ;  $y_{71-80} = 14,4$ . Следовательно,  $\bar{\Delta}_y = (14,4 - 10,45) : 10 = 0,395 \approx 0,4$  млн. т. Предполагая, что в 11-й и 12-й пятилетках сохранится тенденция к ежегодному увеличению среднего уровня на 0,4 млн. т, найдем:  $y_{81-90} = y_{71-80} + \Delta y = 14,4 + 0,4 \cdot 10 = 18,4$  млн. т.

В Продовольственной программе СССР на период до 1990 г. предусмотрено довести среднегодовое производство мяса в 11-й пятилетке до 17—17,5 млн. т и в 12-й пятилетке до 20—20,5 млн. т. Таким образом, в соответствии с Продовольственной программой СССР среднегодовой уровень за 1981—1990 гг. должен составить 18,5—19 млн. т [(17+20) : 2 = 18,5; (17,5+20,5) : 2 = 19].

Следовательно, если в 1981—1990 гг. сохранится тенденция к ежегодному увеличению среднего уровня на 0,4 млн. т, то это почти обеспечит достижение нижней его границы, предусмотренной программой. Реализация же мер, предусмотренных в Продовольственной программе СССР, позволит обеспечить более быстрый рост производства мяса.

С помощью приведенных формул можно рассчитать и среднегодовые уровни отдельно за 11-ю и 12-ю пятилетки:  $y_{81-90} = (y_{81-85} + y_{86-90}) : 2$ ;  $y_{81-85} = y_{81-85} + 5\bar{\Delta}_y$ . Следовательно,  $y_{81-90} = (y_{81-85} + y_{81-85} + 5\bar{\Delta}_y) : 2 = y_{81-85} + 2,5\bar{\Delta}_y$ . Отсюда  $y_{81-85} = y_{81-90} - 2,5\bar{\Delta}_y = 18,4 - 2,5 \cdot 0,4 = 17,4$  млн. т. Таким образом, в 11-й пятилетке даже простое сохранение выявленной тенденции позволит довести среднегодовое производство мяса до уровня, предусмотренного майским (1982 г.) Пленумом ЦК КПСС.

Если для измерения основной тенденции производилось аналитическое выравнивание ряда динамики, то для экстраполяции используется уравнение тренда, в которое подставляются значения  $t$  в будущие годы. Так, для среднегодового объема государственных закупок овощей в СССР за 4 пятилетия был получен параболический тренд:  $y_t = 11,16 + 1,76t + 0,0875t^2$  (табл. 14.20):

$$a_1 = \sum y_t : \sum t^2 = 35,2 : 20 = 1,76;$$

$$4a_0 + 20a_2 = 46,4;$$

$$20a_0 + 164a_2 = 237,6;$$

$$a_0 = 11,1625;$$

$$a_2 = 0,0875.$$

Значение  $t$  для следующего пятилетия (1981—1985 гг.) равно 5. Отсюда  $y_{81-85} = 11,16 + 1,76 \cdot 5 + 0,0875 \cdot 25 = 22,15$  млн. т.

Таблица 14.20. К расчету уравнения тренда среднегодового объема государственных закупок овощей в СССР

Годы	Закупки (в среднем за год), млн. т $y$	$t$	$yt$	$yt^2$	$t^2$	$t^4$
1961—1965	6,7	-3	-20,1	60,3	9	81
1966—1970	9,4	-1	-9,4	9,4	1	1
1971—1975	13,1	1	13,1	13,1	1	1
1976—1980	17,2	3	51,6	154,8	9	81
Итого	46,4	0	35,2	237,6	20	164

Поскольку в базисном периоде уровень колеблется около тренда, очевидно, колебания весьма вероятны и в будущем. Поэтому трудно ожидать, что фактический уровень в будущем точно совпадает с экстраполированным по тренду. Более вероятно, что будущий уровень окажется в определенном интервале, центром которого является значение экстраполированного уровня. Границы этого интервала приближенно равны  $y' \pm t_{\alpha} \hat{\sigma}_e$ , где, как и при расчете ошибки выборки,  $t_{\alpha}$  — коэффициент доверия по распределению Стьюдента, зависящий от

доверительной вероятности и от  $n-m$ ;  $\hat{\sigma}_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n-m}}$  — скорректированное (на число степеней свободы) остаточное среднее квадратическое отклонение ( $m$  — число параметров тренда).

В нашем примере  $n-m = 4-3 = 1$ ;  $\sigma_e = 0,134$ ;  $t_{\alpha} = t_{0,95} = 6,314$ , откуда  $t_{\alpha} \hat{\sigma}_e = 0,134 \cdot 6,314 = 0,85$  млн. т. Следовательно, с вероятностью 0,9 фактический средний уровень будет находиться в пределах от 21,3 до 23,0 млн. т. Нужно, однако, отметить, что такие расчеты носят не только приближенный, но и в известной мере условный характер, связанный с распространением на временные ряды положений и принципов регрессионно-корреляционного анализа выборочной совокупности.

Для прогнозирования используются также различные регрессионно-корреляционные модели (уравнения регрессии), в том числе авторегрессионные с включением фактора времени (вида  $y_t = a_0 + a_1 y_{t-1} + \dots + a_k t$ ), и авторегрессионно-факторные (смешанные):  $y_t = a_0 + a_1 y_{t-1} + \dots + a_k x_{1t} + a_{k+1} x_{2t} + \dots + a_n x_{nt}$  и т. д.

Если экстраполяция заключается в определении уровней ряда динамики, лежащих за его пределами (обычно в будущем), то интерполяция — это приближенный расчет уровней, лежащих внутри ряда динамики, но почему-либо неизвестных. Как и экстраполяция, интерполяция тоже основана на том или ином предположении о тенденции изменения уровней, но характер этого предположения несколько иной: здесь уже не прихо-

дится предполагать, что тенденция, характерная для прошлого, сохранится и в будущем; при интерполяции предполагается, что ни выявленная тенденция, ни ее характер не претерпели существенных изменений в том промежутке времени, уровень (или уровни) которого нам не известен. Такое предположение обычно является более обоснованным, чем предположение о будущей тенденции. Точность результатов интерполяции зависит также и от степени колеблемости уровней ряда около предполагаемого тренда.

Как и экстраполяция, интерполяция может производиться на основе среднего абсолютного прироста, среднего темпа ро-

Таблица 14.21. К расчету показательной кривой для численности специалистов

Год	Численность специалистов, млн. чел. у	$t$	$\lg y$	$t \lg y$	$t^2$
1975	22,8	1	1,3579	1,3579	1
1976	...	—	—	—	—
1977	25,2	3	1,4014	4,2042	9
1978	...	—	—	—	—
1979	...	—	—	—	—
1980	28,6	6	1,4564	8,7384	36
Итого	76,6	10	4,2157	14,3005	46

ста, а также с помощью аналитического выравнивания. В первых двух случаях формулы могут быть выведены из (14.11) и (14.13):

$$y'_i = y_{i-t} + \Delta \bar{t}; \quad y'_i = y_{i-t} (\bar{T}_p)^t,$$

где  $y'_i$  — интерполируемый уровень ( $i$  — номер этого уровня (года);  $y_{i-t}$  — базисный уровень, использованный при расчете  $\Delta$  или  $\bar{T}_p$ ;  $t$  — длина промежутка времени между годом, для которого делается интерполяция, и базисным годом.

Так, численность специалистов с высшим и средним специальным образованием, занятых в народном хозяйстве СССР, составляла на 15.XI (млн. чел.): в 1975 г. — 22,8, в 1980 г. — 28,6. Среднегодовой темп роста равен 104,65 %. Предполагая, что ценные темпы роста мало отличаются один от другого, найдем уровень 1977 г.:  $y'_{77} = 22,8 \cdot (1,0465)^2 = 25,0$  млн. чел. Это лишь на 0,2 млн. чел. меньше действительной численности (табл. 14.21).

Если, кроме конечного и базисного уровней, известны также некоторые промежуточные уровни, интерполяцию можно осуществлять на основе уравнения тренда. Добавив к уровням

1975 и 1980 г. фактический уровень 1977 г., выравнивая этот ряд динамики по показательной кривой (экспоненте):  $y_i = a_0 a_1^i$ . В отличие от выравнивания полного ряда (где известны уровни за все годы) здесь отсчет времени ( $t$ ) нужно вести по порядку, начиная с 1-го года. Однако строки, где уровни отсутствуют, не заполняются. Система уравнений здесь такова:

$$\left. \begin{aligned} n \lg a_0 + \lg a_1 \sum t &= \sum \lg y; & 3 \lg a_0 + 10 \lg a_1 &= 4,2157; \\ \lg a_0 \sum t + \lg a_1 \sum t^2 &= \sum t \lg y; & 10 \lg a_0 + 46 \lg a_1 &= 14,3005. \end{aligned} \right\}$$

Отсюда  $\lg a_0 = 1,3399$ ;  $a_0 = 21,87$ ;  $\lg a_1 = 0,0196$ ;  $a_1 = 1,046$ . Уравнение тренда:  $y_t = 21,87 \cdot (1,046)^t$ . Подставив  $t = 2$ , найдем  $y'_{76} = 21,87 \times (1,046)^2 = 23,93$  млн. чел. и т. д.

## ГЛАВА 15 ИНДЕКСНЫЙ АНАЛИЗ

### 15.1. ОБЩЕЕ ПОНЯТИЕ ОБ ИНДЕКСАХ И ИНДЕКСНОМ МЕТОДЕ

Под **индексом** (лат. *index*) в широком смысле понимается относительный показатель, который характеризует соотношение уровней социально-экономического явления во времени, по сравнению с планом или в пространстве. Во всех этих случаях сопоставляются между собой размеры одноименных показателей, имеющих одинаковое экономическое содержание. Следовательно, индексами являются любые *относительные величины* трех видов: *динамики* (включая плановые темпы роста), *степени выполнения плана и сравнения* (в пространстве). Так например, если мы говорим, что производство электроэнергии в СССР в 1983 г. увеличилось по сравнению с 1940 г. более чем в 29 раз, то эту относительную величину в широком смысле можно назвать индексом динамики производства электроэнергии. Если план добычи газа в СССР в 1983 г. выполнен на 101,3 %, то эта относительная величина является индексом выполнения плана. Многие задания пятилетних планов и Продовольственной программы СССР также выражаются в виде индексов.

По охвату элементов совокупности (ее объектов, единиц и их признаков) различают индексы индивидуальные (элементарные) и сводные (сложные), которые в свою очередь делятся на общие и групповые. *Индивидуальный индекс* характеризует соотношение уровней только одного элемента совокупности (например, изменение объема производства одного определенного вида продукции или себестоимости единицы продукции данного ее вида). Так, приведенный выше индекс динамики производства электроэнергии в СССР является индиви-

дуальным; индивидуальным является и индекс выполнения плана добычи газа.

*Сводный индекс* характеризует соотношение уровней нескольких элементов совокупности (например, изменение объема выпуска нескольких видов продукции, имеющих различную натурально-вещественную форму, или изменение уровня производительности труда при производстве нескольких видов продукции). Так, сводными являются приведенные выше индексы планового задания по росту промышленной продукции и производительности труда в 1981—1985 гг. Если изучаемая совокупность состоит из нескольких групп, то сводные индексы, каждый из которых характеризует изменение уровней отдельной группы единиц, являются *групповыми (субиндексами)*, а сводный индекс, охватывающий всю совокупность единиц,— *общим (тотальным) индексом*.

Показатель, изменение которого характеризует индекс, называется *индексируемым*. Индексируемые показатели могут быть двоякого рода. Одни из них измеряют общий, суммарный размер (объем) того или иного явления и условно называются *объемными, экстенсивными* (количество, т. е. физический объем продукции данного вида, численность рабочих, общие затраты труда (рабочего времени) на производство продукции, общая себестоимость продукции, общий расход материала данного вида и т. п.). Эти показатели получаются как итог непосредственного подсчета или суммирования и являются исходными, первичными.

Другие показатели измеряют уровень явления или признака в расчете на ту или иную единицу совокупности и условно называются *качественными, интенсивными*: выработка продукции в единицу времени (или на одного работника), затраты рабочего времени на единицу продукции, себестоимость единицы продукции и т. д. Эти показатели получаются путем деления объемных показателей, т. е. не носят расчетный, вторичный характер. Они измеряют интенсивность, эффективность явления или процесса и, как правило, являются либо средними, либо относительными величинами.

При использовании *индексного метода* применяется определенная символика, т. е. система условных обозначений. Каждый индексируемый показатель обозначается определенной буквой (обычно латинской). Ниже будут использоваться такие условные обозначения:

$Q$  — количество (объем) произведенной продукции (или количество проданного товара) данного вида в натуральном выражении;

$T$  — общие затраты рабочего времени (труда) на производство продукции данного вида, измеряемые в человеко-часах или человеко-днях; в некоторых случаях этой же буквой будет обозначаться среднее списочное число рабочих (или всех работников);

$z$  — себестоимость единицы продукции;

$p$  — цена единицы продукции или товара;

$M$  — общий расход сырья, материала или топлива на производство продукции данного вида и объема.

На основе этих показателей могут быть получены другие:

$\frac{Q}{T} = q$  — производство продукции данного вида в единицу времени (либо выработка продукции на одного списочного рабочего или работника), т. е. уровень производительности труда;

$\frac{T}{Q} = t$  — затраты рабочего времени (труда) на производство единицы продукции данного вида, т. е. трудоемкость единицы продукции;

$zQ$  — общая себестоимость продукции данного вида, т. е. денежные затраты на ее производство;

$pQ$  — общая стоимость произведенной продукции данного вида<sup>1</sup> (либо общая стоимость проданного товара данного вида, т. е. товарооборот);

$\frac{M}{Q} = m$  — расход сырья, материала или топлива данного вида на производство единицы продукции, т. е. удельный расход сырья, материала или топлива.

Показатели за базисный период имеют в формулах подстрочный знак «0», а за сравниваемый (текущий, отчетный) период — знак «1»; если сравниваются данные за несколько периодов, то указываются либо их порядковые номера, либо конкретные хронологические периоды (годы, кварталы и т. п.). Плановые показатели снабжаются знаком «пл».

Индивидуальные индексы обозначаются буквой  $i$  и тоже снабжаются подстрочным знаком — обозначением индексируемого показателя. Так,  $i_z$  означает индивидуальный индекс количества (физического объема) произведенной продукции (или проданного товара) данного вида;  $i_t$  — индивидуальный индекс себестоимости единицы продукции данного вида;  $i_q$  — индивидуальный индекс общей себестоимости продукции отдельного вида и т. п.

Сводные индексы обозначаются буквой  $I$  и также сопровождаются подстрочными знаками показателей, изменение которых они характеризуют. Например,  $I_t$  — сводный индекс трудоемкости единицы продукции,  $I_T$  — сводный индекс общих затрат времени (или численности работников) и т. д.

Индивидуальные индексы являются обычными относительными величинами, т. е. могут быть названы индексами только в широком понимании этого термина.

<sup>1</sup> Речь идет о денежном (ценовном) выражении стоимости продукции.



Формулы индивидуального индекса объема продукции в натуральном выражении, т. е. индексы физического ее объема, можно записать так:

$$i_Q = Q_1 : Q_0 \text{—индекс динамики (темп роста);} \quad (15.1a)$$

$$i_Q = Q_{пл} : Q_0 \text{ (или } Q_{1пл} : Q_{0пл}) \text{—индекс планового задания (плановый темп роста);} \quad (15.1б)$$

$$i_Q = Q_1 : Q_{пл} \text{—индекс (степень) выполнения плана;} \quad (15.1в)$$

$$i_Q = Q_A : Q_B \text{—территориальный индекс,} \quad (15.1г)$$

где А и В—разные территории или объекты.

Аналогично строятся индивидуальные индексы других показателей—себестоимости единицы продукции, производительности труда и т. п.

Поскольку индивидуальные индексы—обычные относительные величины, для них сохраняют силу соответствующие взаимосвязи:

а) индекс планового задания, умноженный на индекс выполнения плана, равен индексу динамики (темпу роста):

$$\frac{z_{пл}}{z_0} \cdot \frac{z_1}{z_{пл}} = \frac{z_1}{z_0};$$

б) произведение цепных индексов динамики равно соответствующему базисному индексу:

$$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} = \frac{q_3}{q_0}.$$

Для индивидуальных индексов всех четырех типов (динамики, планового задания и т. д.) справедлива также следующая взаимосвязь: если произведение двух или нескольких показателей представляет собой новый показатель, имеющий реальный экономический смысл, то произведение индексов этих показателей-составителей равно индексу их произведения, т. е. индексу нового показателя.

Так, если

$$qT = Q, \text{ то } \frac{q_1}{q_0} \cdot \frac{T_1}{T_0} = \frac{q_1 T_1}{q_0 T_0} = \frac{Q_1}{Q_0},$$

или

$$i_q \cdot i_T = i_Q, \quad (15.2)$$

т. е. индекс производительности труда, умноженный на индекс

общих затрат времени, равен индексу физического объема продукции.

Аналогично

$$i_q \cdot i_Q = i_{tQ}; \quad (15.3)$$

$$i_m \cdot i_Q = i_{mQ} = i_M. \quad (15.4)$$

Такого рода взаимосвязи позволяют по двум или нескольким известным индексам находить связанный с ними индекс, величина которого неизвестна, т. е. производить косвенное индексирование. Если, например, известно, что объем произведенной продукции возрос на 11,0 %, а общие затраты времени на ее производство увеличились на 2,2 %, то  $i_Q = 1,11$ ,  $i_T = 1,022$  и согласно (15.2)  $i_q \cdot 1,022 = 1,11$ , откуда  $i_q = 1,11 : 1,022 = 1,086$ , т. е. производительность труда повысилась на 8,6 %. В то же время, так как  $t = T : Q$ , то  $i_t = i_T : i_Q = 1,022 : 1,11 = 0,921$ , т. е. трудоемкость единицы продукции снизилась на 7,9 %. Отметим также, что уровень производительности труда ( $q = Q : T$ ) и трудоемкость единицы продукции ( $t = T : Q$ )—показатели обратные, а потому и их индивидуальные индексы—обратные величины:

$$i_q = 1 : i_t = i_Q : i_T. \quad (15.5)$$

Индексы в узком смысле, или собственно индексы,—это тоже показатели относительные, но особого рода. Они имеют более сложную методику построения и расчета, а специфические приемы их построения и составляют суть индексного метода.

Социально-экономические явления и характеризующие их показатели могут быть соизмеримыми, т. е. иметь общую меру, и несоизмеримыми. Так, объем продукции или товара одного и того же вида и сорта, произведенные на разных предприятиях или проданные в разных магазинах, соизмеримы и могут суммироваться, а объемы разных видов продукции или товаров несоизмеримы и непосредственно суммироваться не могут. Нельзя, например, складывать килограммы хлеба с литрами молока, метрами ткани и парами обуви. Несоизмеримость и невозможность непосредственного суммирования при построении и расчете сводного индекса объясняется здесь не столько различием натуральных единиц измерения, сколько различием потребительных свойств, неодинаковой натурально-вещественной формой этих продуктов или товаров. Количество проданного мяса измеряется в тех же единицах веса (массы), что и количество рыбы, но непосредственное суммирование килограммов мяса и рыбы все же экономически бессмысленно, так как полученная величина представляла бы собой в прямом смысле «ни рыбу ни мясо».

Итак, причиной несонизмеримости является здесь неоднородность, различие натуральной формы и свойств.

В связи с этим для разнородных продуктов или товаров сводный индекс физического объема (количества) нельзя построить и вычислить как отношение простых сумм:  $I_q = \frac{\sum Q_1}{\sum Q_0}$ . Здесь требуется использование специальных приемов, составляющих специфику *собственно индексов и индексного метода*.

«Различные вещи становятся количественно сравнимыми лишь после того, как они сведены к одному и тому же единству. Только как выражения одного и того же единства они являются одноименными, а следовательно, соизмеримыми величинами», — говорит К. Маркс<sup>2</sup>.

Единство различных видов продукции или разных товаров состоит в том, что они являются продуктами общественного воспроизводства, продуктами труда, имеют определенную стоимость и ее денежное выражение — цену ( $p$ ). Каждый продукт имеет также ту или иную себестоимость ( $z$ ) и трудоемкость ( $t$ ). Эти качественные показатели и могут быть использованы в качестве *общей меры — коэффициентов соизмерения* разнородных продуктов. Умножая объем продукции каждого вида ( $Q$ ) на соответствующую цену, себестоимость или трудоемкость единицы продукции, мы сведем различные продукты к одному и тому же единству и получим сравнимые показатели, которые можно суммировать ( $Qp, Qz, Qt = T$ )<sup>3</sup>.

Аналогично обстоит дело и при построении сводных индексов качественных показателей. Пусть, например, нас интересует изменение общего уровня цен на различные проданные товары. Хотя формально цены разных товаров соизмеримы, однако непосредственное их суммирование, без учета количества проданного товара каждого вида, дает величину, лишенную самостоятельного практического значения. Так, например, сложив цены 1 кг хлеба, 1 м ткани, 1 пары обуви и т. д., мы получим общую стоимость совершенно произвольного набора товаров. Однако нас интересует изменение общего уровня цен не того или иного произвольного набора товаров, а уровня цен фактически проданного их количества, образующего реально

существующую совокупность. Поэтому сводный индекс цен нельзя построить как отношение простых сумм:  $I_p = \frac{\sum p_1}{\sum p_0}$ . Цены отдельных товаров не учитывают конкретного количества проданных товаров и их статистического веса и роли в процессе товарооборота<sup>4</sup>. Простые суммы цен отдельных товаров непригодны для построения сводного индекса также и потому, что цены зависят от единицы измерения товаров, изменение которых даст иные суммы и иную величину индекса.

Все это относится и к другим качественным показателям — себестоимости и трудоемкости единицы продукции ( $z, t$ ), удельному расходу сырья или материала ( $m$ ) и т. д. Следовательно, при построении сводных индексов качественных показателей их также нельзя рассматривать в отрыве от связанных с ними объемных показателей, в расчете на единицу которых вычислены эти качественные показатели. Только умножив тот или иной качественный показатель ( $p, z, t, m$ ) на непосредственно связанный с ним объемный показатель ( $Q$ ), мы сможем учесть роль и статистический вес каждого вида продукции (или товара) в том или ином экономическом процессе — процессе образования общей стоимости ( $pQ$ ), общей себестоимости ( $zQ$ ), общих затрат рабочего времени ( $tQ = T$ ) и т. п. Вместе с тем мы получим показатели, суммирование которых имеет практическую значимость.

Таким образом, *первая особенность индексного метода* и собственно индексов состоит в том, что индексируемый показатель рассматривается не изолированно, а *во взаимосвязи* с другими показателями.

Умножая индексируемый показатель на другой, связанный с ним, мы сводим различные явления к их единству, обеспечиваем их количественную сравнимость и учитываем их вес в реальном экономическом процессе. Поэтому показатели-сумможители, связанные с индексируемыми показателями, принято называть *весами индексов*, а умножение на них — *взвешиванием*.

Однако умножение значений индексируемого показателя на связанные с ним значения другого показателя (веса) еще не решает проблему собственно индекса. Умножив, например, цены на соответствующие им количества товаров, мы найдем стоимость этих товаров в каждом периоде и тем самым решим проблему соизмерения и взвешивания. Однако сопоставление полученных сумм произведений ( $\sum p_1 Q_1$  и  $\sum p_0 Q_0$ ) дает показатель, который характеризует изменение стоимости товаров, т. е. товарооборота, зависящего от двух факторов — цен и количества (объемов) товаров:

$$I_{pq} = \frac{\sum p_1 Q_1}{\sum p_0 Q_0} \quad (15.6)$$

<sup>4</sup> По тем же причинам не решает вопроса и простое осреднение индивидуальных индексов:  $I_p = \frac{\sum I_p}{n}$ .

<sup>1</sup> Нельзя построить сводный индекс и путем простого осреднения индивидуальных индексов:  $I_p = \frac{\sum I_p}{n}$ , так как эта средняя является неавансированной и не учитывает удельные веса отдельных продуктов. Как будет показано ниже, идея осреднения индивидуальных индексов может быть успешно реализована именно путем их соответствующего взвешивания.

<sup>2</sup> Маркс К., Энгельс Ф. Соч., т. 23, с. 58—59.

<sup>3</sup> Если продукты имеют одинаковое потребительное назначение (например, различные виды топлива или кормов), то коэффициентами соизмерения могут выступать и другие качественные показатели (теплотворная способность, кормовые единицы). Это используется при измерении продукции в условных натуральных единицах (см. гл. 8).

Для того чтобы индекс охарактеризовал изменение только одного фактора, нужно устранить в формуле (15.6) изменение другого фактора, зафиксировав его как в числителе, так и в знаменателе на уровне одного и того же периода. Так, если товары, проданные в обоих периодах, оценить по одним и тем же, например базисным ценам ( $p_0$ ), то индекс отразит изменение только одного фактора — индексируемого показателя  $Q$ :

$$I_Q = \frac{\sum Q_1 p_0}{\sum Q_0 p_0}. \quad (15.7)$$

Если же в формуле индекса предусмотреть лишь изменение цен ( $p$ ), а количество товаров ( $Q$ ) зафиксировать в числителе и знаменателе индекса на одном и том же уровне (либо на отчетном, либо на базисном), то индекс будет характеризовать только изменение цен, т. е. индексируемого показателя, так как изменение весов ( $Q$ ) будет устранено (элиминировано) благодаря их фиксированию:

$$a) I_p = \frac{\sum p_1 Q_1}{\sum p_0 Q_1}; \quad б) I_p = \frac{\sum p_1 Q_0}{\sum p_0 Q_0}. \quad (15.8a, б)$$

В обоих случаях ( $I_Q$  и  $I_p$ ) индекс отразил изменение только одного фактора — индексируемого показателя — благодаря фиксированию другого (весов) на одном и том же уровне. Элиминирование влияния изменения весов путем их фиксирования в числителе и знаменателе индекса на одном и том же уровне — вторая особенность индексов и индексного метода.

Рассматривая проблемы, возникающие при построении собственно индексов, мы ставили задачу дать сравнительную характеристику уровней сложного явления, состоящего из различных элементов (разные виды продукции и т. п.). Так,  $I_p$  должен показать, как изменился в целом уровень цен, т. е. измерить динамику цен различных товаров в виде одного обобщающего показателя. Исторически собственно индексы появились как результат решения именно этой экономической задачи — задачи обобщения, синтеза динамики отдельных элементов сложного явления в одном обобщающем показателе — сводном индексе.

Однако собственно индексы используются для решения и другой задачи — анализа влияния изменения отдельных показателей-факторов на изменение показателя, представляющего функцию этих факторов-аргументов. Так, в нашем примере общая стоимость проданных товаров (товарооборот —  $\sum pQ$ ) есть функция их цен ( $p$ ) и количеств (объемов —  $Q$ ). Поэтому можно поставить задачу измерить влияние каждого из этих факторов на изменение товарооборота: определить, как он изменился отдельно за счет изменения каждого фактора. Индексы, применяемые для решения подобных аналитических задач, также строятся с использованием специфических особен-

ностей индексного метода — взвешивания и элиминирования изменения весов.

Таким образом, *собственно индекс* — это относительный показатель особого рода, в котором уровни социально-экономического явления рассматриваются в связи с другим (или другими) явлением, изменение которого при этом элиминируется. Показатели, связанные с индексируемым показателем, используются в качестве весов индекса, а взвешивание и элиминирование изменения весов (фиксирование в числителе и знаменателе индекса на одном и том же уровне) составляют специфику собственно индексов и индексного метода.

## 15.2. АГРЕГАТНЫЕ ИНДЕКСЫ КАЧЕСТВЕННЫХ ПОКАЗАТЕЛЕЙ

Каждый качественный показатель связан с тем или иным объемным показателем, в расчете на единицу измерения которого он исчисляется (или к единице измерения которого относится). Так, цена единицы товара связана с его количеством ( $Q$ ); с объемом произведенной продукции связаны такие качественные показатели, как цена ( $p$ ), себестоимость ( $z$ ) и трудоемкость ( $t = T:Q$ ) единицы продукции, а также удельный расход сырья, материала ( $m = M:Q$ ).

Сводные индексы качественных показателей должны характеризовать не их изменение вообще, применительно к какому-либо произвольному набору товаров или продукции, а изменение цен, себестоимости, трудоемкости или удельных расходов влode определенного количества произведенной продукции или проданных товаров. Это и достигается путем взвешивания — умножения уровней индексируемого качественного показателя на значения связанного с ним объемного показателя (веса) — и фиксирования весов в числителе и знаменателе индекса на одном и том же уровне. Сопоставление сумм таких произведений дает *агрегатный индекс* (см., например, выше формулы 15.8a, б). Аналогично могут быть построены агрегатные индексы динамики себестоимости и трудоемкости единицы продукции, а также индекс удельного расхода сырья или материала.

Основной проблемой при построении этих сводных индексов является экономически обоснованный выбор уровня, на котором нужно зафиксировать веса индекса, т. е. в данном случае объем продукции (или товаров) —  $Q$ .

Обычно перед сводным индексом динамики качественного показателя ставится задача измерить не только относительное изменение уровня, но и абсолютную величину того *экономического эффекта*, который получен в текущем периоде в результате этого изменения: сумму экономии покупателей за счет снижения цен (или сумму их *дополнительных расходов*, если цены повсились), сумму экономии (или *дополнительных затрат*) за счет изменения себестоимости и т. п.



Такая постановка задачи приводит к индексам динамики качественных показателей с весами текущего периода: во-первых, нас интересует изменение себестоимости или трудоемкости той продукции, которая выпущена в настоящее время, а не в прошлом; во-вторых, экономический эффект должен быть увязан с фактическими результатами текущего, отчетного, а не предыдущего (базисного) периода.

Пусть, например, имеются следующие данные по двум видам продукции.

Таблица 15.1. Объем производства и себестоимость продукции

Вид продукции	Исходные данные				Расчет		
	Базисный период		Отчетный период		Общая себестоимость продукции, руб.		
	произведено	себестоимость единицы, руб.	произведено	себестоимость единицы, руб.	базисный период	отчетный период	в отчетном периоде при базисной себестоимости единицы продукции
	$Q_0$	$z_0$	$Q_1$	$z_1$	$z_0 Q_0$	$z_1 Q_1$	$z_0 Q_1$
А, шт.	800	92	1000	90	73 600	90 000	92 000
Б, т	600	66	500	67	39 600	33 500	33 000
Итого	—	—	—	—	113 200	123 500	125 000

В нашем примере себестоимость 1 шт. продукции А снизилась на 2 руб. ( $z_1 - z_0 = 90 - 92 = -2$ ). Следовательно, при производстве каждой единицы этой продукции завод сэкономил в отчетном периоде 2 руб., а выпустив 1000 шт., сэкономил 2000 руб. [ $(z_1 - z_0) Q_1 = -2 \cdot 1000 = -2000$ ] (здесь и далее знак минус означает уменьшение затрат, т. е. экономию, а знак плюс — увеличение затрат). Соответственно дополнительные затраты на каждую тонну продукции Б составляли 1 руб., а на весь выпуск было дополнительно израсходовано 500 руб. [ $(z_1 - z_0) Q_1 = (67 - 66) \cdot 500 = 500$ ]. В целом по обоим видам продукции экономический эффект составил:  $\sum (z_1 - z_0) Q_1 = \sum z_1 Q_1 - \sum z_0 Q_1 = -2000 + 500 = -1500$  руб. (экономия). Таким образом, реальный, фактически достигнутый в отчетном периоде эффект за счет изменения себестоимости единицы продукции выражается как разность показателей  $(\sum z_1 Q_1 - \sum z_0 Q_1)$ , путем деления которых может быть вычислен агрегатный индекс себестоимости, в котором веса ( $Q$ ) зафиксированы на уровне отчетного периода:

$$I_z = \frac{\sum z_1 Q_1}{\sum z_0 Q_1} \quad (15.9)$$

В нашем примере

$$I_z = \frac{90 \cdot 1000 + 67 \cdot 500}{92 \cdot 1000 + 66 \cdot 500} = \frac{123\,500}{125\,000} = 0,988, \text{ или } 98,8 \%,$$

т. е. в целом по обоим видам продукции себестоимость единицы продукции снизилась на 1,2 %, за счет чего и было сэкономлено 1500 руб. ( $123\,500 - 125\,000$ )<sup>1</sup>.

В этом индексе числитель  $(\sum z_1 Q_1)$  — сумма фактических затрат на продукцию в отчетном периоде, а знаменатель  $(\sum z_0 Q_1)$  — условная сумма затрат на ту же продукцию отчетного периода при базисных уровнях себестоимости единицы продукции каждого вида. Иными словами, знаменатель — условная величина, которая показывает, сколько средств было бы затрачено на продукцию отчетного периода, если бы себестоимость единицы каждого вида продукции сохранялась на базисном уровне.

Как уже было показано, разность между суммами в числителе и знаменателе индекса  $(\sum z_1 Q_1 - \sum z_0 Q_1)$  или  $\sum (z_1 - z_0) Q_1$  характеризует реальный экономический эффект, полученный за счет изменения себестоимости единицы продукции. Таким образом, взвешивание по весам отчетного (текущего) периода увязывает индекс качественного показателя с показателем экономического эффекта, который получен за счет изменения индексируемого показателя. Поэтому агрегатные индексы динамики качественных показателей строятся и вычисляются обычно с весами отчетного периода:

$$I_l = \frac{\sum t_1 Q_1}{\sum t_0 Q_1}; \quad I_m = \frac{\sum m_1 Q_1}{\sum m_0 Q_1} \quad (15.10; 15.11)$$

Соответственно в индексах планового задания веса фиксируются на уровне плана:

$$I_z = \frac{\sum z_{пл} Q_{пл}}{\sum z_0 Q_{пл}}; \quad I_l = \frac{\sum t_{пл} Q_{пл}}{\sum t_0 Q_{пл}}; \quad I_m = \frac{\sum m_{пл} Q_{пл}}{\sum m_0 Q_{пл}} \quad (15.12; 15.13; 15.14)$$

В этих индексах разность между числителем и знаменателем характеризует размер плановой экономии (или плановых дополнительных затрат) за счет изменения соответствующего качественного показателя: в  $I_z$  — плановой экономии денежных средств, в  $I_l$  — плановой экономии рабочего времени, в  $I_m$  — плановой экономии сырья или материала определенного вида (в натуральном выражении).

<sup>1</sup> Отметим, что фиксирование весов устраняет лишь влияние их изменения, но не освобождает индекс от влияния весов вообще. Так, в нашем

примере  $I_z = \frac{\sum z_1 Q_0}{\sum z_0 Q_0} = \frac{112\,200}{92 \cdot 800 + 66 \cdot 600} = 0,991$ , или 99,1 %, что отличается от результата, полученного выше по (15.9). Изменно это обстоятельство и требует экономического обоснования выбора весов индекса и периода, на уровне которых они фиксируются.

Индексы динамики и планового задания для качественных показателей ( $z, t, m$ ) вычисляются применительно к сравнимой продукции, т. е. охватывают те ее виды, которые произвелись как в базисном, так и в отчетном периоде (либо их производство предусматривалось планом).

В агрегатных индексах выполнения плана (по снижению  $z, t, m$ ) веса фиксируются на уровне отчетного периода:

$$I_z = \frac{\sum z_{пл} Q_1}{\sum z_{отч} Q_1}; \quad I_t = \frac{\sum t_1 Q_1}{\sum t_{отч} Q_1} \quad \text{и т. п.} \quad (15.15; 15.16)$$

Разность между числителем и знаменателем в этих индексах характеризует размер сверхплановой экономии (или сверхплановых затрат) за счет изменения качественного показателя.

Казалось бы, что для сводных индексов планового задания, выполнения плана и динамики можно ожидать такой же взаимосвязи, как и для индивидуальных индексов (см. формулы 15.2—15.4). Однако в том случае, если в двух из этих индексов веса зафиксированы на уровне отчетного периода, а в одном (индекс планового задания) — на уровне плана, взаимосвязь сводных индексов усложняется. Действительно, при указанном фиксировании весов:

$$\frac{\sum z_{пл} Q_{пл}}{\sum z_0 Q_{пл}} \cdot \frac{\sum z_1 Q_1}{\sum z_{пл} Q_1} \neq \frac{\sum z_1 Q_1}{\sum z_0 Q_1}. \quad (A)$$

По этой же причине сумма плановой и сверхплановой экономии не равна экономии в отчетном периоде по сравнению с базисным:

$$(\sum z_{пл} Q_{пл} - \sum z_0 Q_{пл}) + (\sum z_1 Q_1 - \sum z_{пл} Q_1) \neq \sum z_1 Q_1 - \sum z_0 Q_1. \quad (B)$$

Для того чтобы обеспечить указанную взаимосвязь ( $I_{пл. задания} \times I_{выполн. плана} = I_{динамики}$ ) и баланс экономии, нужно во всех трех индексах зафиксировать веса на одном и том же уровне:

$$\frac{\sum z_{пл} Q_1}{\sum z_0 Q_1} \cdot \frac{\sum z_1 Q_1}{\sum z_{пл} Q_1} = \frac{\sum z_1 Q_1}{\sum z_0 Q_1}. \quad (15.17a)$$

Здесь, в отличие от изложенного выше, индекс планового задания взвешен по отчетным (а не по плановым) весам<sup>1</sup>. Это равносильно умножению левой части неравенства (A) на индекс влияния ассортиментных сдвигов<sup>2</sup>:

$$\frac{\sum z_{пл} Q_{пл}}{\sum z_0 Q_{пл}} \cdot \left( \frac{\sum z_{пл} Q_1}{\sum z_1 Q_1} \cdot \frac{\sum z_{пл} Q_{пл}}{\sum z_0 Q_{пл}} \right) \cdot \frac{\sum z_1 Q_1}{\sum z_{пл} Q_1} = \frac{\sum z_1 Q_1}{\sum z_0 Q_1} \quad (15.17b)$$

$$\text{или } I_{пл. задания} \cdot I_{асс. сдвигов} \cdot I_{выполн. плана} = I_{динамики}. \quad (15.17b)$$

<sup>1</sup> При составлении плана, когда  $Q_1$  еще неизвестны, в индексе планового задания веса фиксируются на уровне плана (см. формулы 15.12—15.14), но по окончании отчетного периода этот индекс можно перевзвесить.

<sup>2</sup> Под ассортиментными сдвигами здесь понимается изменение соотношения между выпуском отдельных видов продукции в отчетном периоде по сравнению с этим соотношением по плану. Индекс влияния ассортиментных сдвигов (в формуле 15.17b он заключен в скобки) характеризует степень влияния этих сдвигов на динамику себестоимости продукции.

Таким образом, общая экономия (или дополнительные затраты) в отчетном периоде по сравнению с базисным складываются так:

$$\begin{aligned} \text{Плановая экономия} &= \sum z_{пл} Q_{пл} - \sum z_0 Q_{пл} \\ \text{Экономия за счет} \\ \text{ассортиментных сдвигов} &= (\sum z_{пл} Q_1 - \sum z_0 Q_1) - \\ &\quad - (\sum z_{пл} Q_{пл} - \sum z_0 Q_{пл}) \\ \text{Сверхплановая экономия} &= \sum z_1 Q_1 - \sum z_{пл} Q_1 \\ \text{Общая экономия} &= \sum z_1 Q_1 - \sum z_0 Q_1 \end{aligned}$$

Сложив плановую экономию и экономию за счет ассортиментных сдвигов, получим плановую экономию в расчете на отчетный ассортимент ( $\sum z_{пл} Q_1 - \sum z_0 Q_1$ ).

Рассмотрим агрегатный индекс производительности труда. В некоторых случаях умножение индексируемого качественного показателя на непосредственно связанный с ним объемный показатель, т. е. взвешивание, дает такой объемный показатель, элементы которого являются несоизмеримыми и не могут суммироваться. Это усложняет построение сводного индекса в агрегатной форме.

Уровень производительности труда ( $q$ ) измеряется выработкой продукции за единицу времени (или выработкой продукции в расчете на одного работника):  $q = Q : T$ . Поэтому в качестве весов агрегатного индекса производительности труда нужно принять общие затраты времени (или число работников) —  $T$ . Зафиксировав веса на уровне отчетного периода, получим:

$$I_q = \frac{\sum q_1 T_1}{\sum q_0 T_1}. \quad (15.18a)$$

Эта формула применима только для одноименной продукции. Поскольку  $qT = Q$ , суммирование в формуле (15.18a) возможно только в том случае, если оно производится по одному и тому же виду продукции (или по одноименным видам продукции), производимой на отдельных участках производства (например, добыча угля в шахтах и в разрезах (открытый способ добычи), производство молока на отдельных фермах, колхозах и т. п.).

В случае разноименной продукции, отдельные виды которой в натуральном выражении суммироваться не могут, агрегатный индекс производительности труда приходится строить и исчислять иначе — как обратную величину сводного агрегатного индекса трудоемкости единицы продукции. При этом исходят из того, что если уровень производительности труда ( $q = Q : T$ ) и трудоемкость единицы продукции ( $t = T : Q$ ) являются обратными показателями, то и их сводные индексы (как

и индивидуальные — см. формулу (15.5)) тоже являются показателями обратными:

$$I_q = \frac{1}{I_t} = \frac{\sum t_0 Q_1}{\sum t_1 Q_1}. \quad (15.186)$$

Пусть, например, имеются следующие данные (табл. 15.2). Для расчета сводного индекса производительности труда найдем сначала затраты времени на 1 ц продукции каждого вида в 1980 г., т. е. базисную трудоемкость (см. табл. 15.2).

Таблица 15.2. Производство продукции в затраты рабочего времени в колхозе

Вид продукции	Исходные данные				Расчет	
	1980		1983		Затрачено чел.-ч на 1 ц в 1980 г.	Было бы затрачено чел.-ч на 1 ц в 1983 г. при трудоемкости 1980 г., тыс.
	Произведено, тыс. ц	Затрачено, чел.-ч	Произведено, тыс. ц	Затрачено, чел.-ч		
	$Q_0$	$T_0$	$Q_1$	$T_1$	$I_t = T_0 : Q_1$	$t_0 Q_1$
Зерно	40	56,0	46	57,5	1,40	64,40
Сахарная свекла*	30	32,4	33	33,0	1,08	35,64
Молоко	20	189,2	22	198,0	9,46	208,12
Итого	—	277,6	—	288,5	—	308,16

\* Фабричная.

Умножив ее на объем продукции в 1983 г., получим величину условных затрат времени (последняя графа табл. 15.2). Учитывая, что  $t_1 Q_1 = T_1$ , вычислим сводный индекс:

$$I_q = \frac{\sum t_0 Q_1}{\sum t_1 Q_1} = \frac{308,16}{288,5} = 1,068, \text{ или } 106,8 \, \%$$

Следовательно, в целом по всем трем видам продукции производительность труда повысилась на 6,8%. В результате этого вместо более 308 тыс. чел.-ч, которые потребовались бы для производства продукции в 1983 г., если бы трудоемкость 1 ц осталась на уровне 1980 г., фактически в 1983 г. было затрачено лишь 288,5 тыс. чел.-ч. Таким образом, за счет снижения

трудоемкости единицы продукции, т. е. за счет роста производительности труда, было сэкономлено 19,66 тыс. чел.-ч ( $\sum t_1 Q_1 - \sum t_0 Q_1 = 288,5 - 308,16 = -19,66$ ), в том числе при производстве зерна — 6,9 тыс. ч ( $T_1 - t_0 Q_1 = 57,5 - 64,4$ ), сахарной свеклы — 2,64 тыс., молока — 10,12 тыс.

### 15.3. АГРЕГАТНЫЕ ИНДЕКСЫ ОБЪЕМНЫХ ПОКАЗАТЕЛЕЙ

Объемные показатели могут быть соизмеримыми ( $T$ ;  $pQ$ ;  $zQ$ ) и несоизмеримыми (объем продукции или товаров разного вида —  $Q$ ). Соизмеримые объемные показатели могут непосредственно суммироваться, и построение агрегатных индексов не вызывает трудностей.

Как уже говорилось выше (см. 15.1), для получения общего итога и построения агрегатного индекса несоизмеримого объемного показателя нужно предварительно соизмерить отдельные значения этого показателя. Исходя из экономической сущности явления нужно найти *общую меру* и использовать ее в качестве *коэффициента соизмерения*. Такой общей мерой для объемных показателей являются связанные с ними качественные показатели. Так, объемы различных видов продукции могут быть соизмерены с помощью цены ( $p$ ), себестоимости ( $z$ ) и трудоемкости ( $t$ ) единицы этих продуктов. Умножая индексируемый объемный показатель на тот или иной качественный показатель-соизмеритель, мы не только обеспечиваем возможность суммирования, но одновременно учитываем также роль каждого элемента, например продукта, в реальном экономическом процессе, т. е. его статистический вес в этом процессе.

Поскольку в индексе объемного показателя в качестве весов могут выступать различные качественные показатели, возникает вопрос о том, какой же именно из них следует использовать. Этот вопрос в каждом конкретном случае должен решаться в соответствии с той познавательной экономической задачей, которая ставится перед индексом, т. е. выбор тех или иных весов-соизмерителей должен быть обоснован экономически.

В практике экономической и статистической работы в качестве весов агрегатного индекса объема продукции обычно используются цены. Так строятся индексы объема промышленной и сельскохозяйственной продукции, а также индексы физического объема товарооборота.

В ряде случаев изменение объема продукции интересует нас не само по себе, а с точки зрения его влияния на изменение показателя более сложного порядка — общей стоимости продукции, общей ее себестоимости, общих затрат рабочего времени, общего объема производства на данном его участке и т. п. В таких случаях выбор весов-соизмерителей определяется взаимосвязью показателей-факторов, от которых зависит более сложный показатель.



Так, для измерения вклада предприятия (производственного объединения) в создание конечного результата используется показатель объема производства. Безразмерными являются отдельные продукты (изделия) при этом могут выступать нормативная стоимость, нормативная заработная плата, а также норматив чистой продукции. Норматив чистой продукции — это часть цены, включающая заработную плату, отчисления на социальное страхование и прибыль. Нормативы чистой продукции отражают величину вновь созданной стоимости, созданной собственным трудом работников предприятия (объединения) при производстве отдельных видов продукции.

Чтобы индекс отражал только изменение индексируемого объектного показателя, веса в его числителе и знаменателе фиксируются на уровне одного и того же периода. В практике экономической работы в индексах динамики объемных показателей веса обычно фиксируются на уровне базисного периода (см. формулу 15.7). Это, как будет показано далее, обеспечивает возможность построения систем взаимосвязанных индексов.

В отличие от индексов качественных показателей, которые исчисляются по сравнимому кругу единиц (сравнимой продукции), сводные индексы объемных показателей в целях полноты и точности должны охватывать весь круг единиц, произведенных (или проданных) в каждом периоде. В связи с этим возникает вопрос о том, какие значения весов следует брать для тех видов продукции, которые в одном из сравниваемых периодов не производились.

В практике советской статистики в таких случаях применяются два способа. При расчете индексов объема промышленной продукции новые ее виды, для которых нет цен базисного периода, оцениваются условно по ценам текущего периода. При расчете же индексов объема проданных товаров используется метод, основанный на условном предположении, что цены на новые товары изменились в той же степени, что и цены на сравнимый круг аналогичных товаров. Если, например, цены на сравнимый круг товаров снизились на 5% ( $I_p = 0.95$ ), а цена нового товара равна 19 руб., то его условная базисная цена составит 20 руб. ( $19 : 0.95$ ). Оба эти метода не безупречно теоретически, так как носят условный характер.

#### 15.4. РЯДЫ АГРЕГАТНЫХ ИНДЕКСОВ С ПОСТОЯННЫМИ И ПЕРЕМЕННЫМИ ВЕСАМИ

При изучении динамики экономических явлений строятся и исчисляются индексы за ряд последовательных периодов. Они образуют ряды либо базисных, либо цепных индексов. В ряду *базисных индексов* сравнение индексируемого показателя в каждом индексе производится с уровнем одного и того же периода, а в ряду *цепных индексов* индексируемый показатель сопоставляется с уровнем предыдущего периода.

В каждом отдельном индексе веса в его числителе и знаменателе обязательно фиксируются на одном и том же уровне.

Если же строится ряд индексов, то веса в нем могут быть либо *постоянными для всех индексов ряда, либо переменными.*

Ряд базисных индексов объема продукции:  $\frac{\sum Q_1 p_0}{\sum Q_0 p_0}$ ,  $\frac{\sum Q_2 p_0}{\sum Q_0 p_0}$ ,  $\frac{\sum Q_3 p_0}{\sum Q_0 p_0}$  и т. д. имеет постоянные веса ( $p_0$ ). Постоянные веса ( $p_0$ ) имеет и ряд цепных индексов:  $\frac{\sum Q_1 p_0}{\sum Q_0 p_0}$ ,  $\frac{\sum Q_2 p_0}{\sum Q_1 p_0}$ ,  $\frac{\sum Q_3 p_0}{\sum Q_2 p_0}$  и т. д.

Ряд цепных индексов цен:  $\frac{\sum p_1 Q_1}{\sum p_0 Q_1}$ ,  $\frac{\sum p_2 Q_2}{\sum p_1 Q_2}$ ,  $\frac{\sum p_3 Q_3}{\sum p_2 Q_3}$  и т. д. построен с переменными весами (в 1-м индексе —  $Q_1$ , во 2-м —  $Q_2$  и т. д.).

Для индексов динамики с постоянными весами сохраняет силу взаимосвязь между цепными и базисными темпами роста (индексами):

$$\frac{\sum Q_1 p_0}{\sum Q_0 p_0} \cdot \frac{\sum Q_2 p_0}{\sum Q_1 p_0} = \frac{\sum Q_2 p_0}{\sum Q_0 p_0} \quad \text{и т. п.} \quad (15.19)$$

Таким образом, использование постоянных весов в течение ряда лет позволяет переходить от цепных индексов к базисным, и наоборот. Поэтому ряды индексов объема продукции и объема проданных товаров строятся в нашей статистике с постоянными весами. Так, с 1982 г. в индексах объема продукции промышленности в качестве постоянных весов используются цены, зафиксированные на уровне, который был установлен на 1 января 1982 г. Такие цены, используемые в течение ряда лет, называются *сопоставимыми (фиксированными)*.

Использование в индексах объема продукции (товаров) сопоставимых цен позволяет путем простого суммирования получать итоги за несколько лет, например за пятилетку. Кроме того, при фиксированных весах индекс динамики равен произведению индекса планового задания и индекса выполнения плана:

$$\frac{\sum Q_1 p}{\sum Q_0 p} = \frac{\sum Q_{пл} p}{\sum Q_0 p} \cdot \frac{\sum Q_1 p}{\sum Q_{пл} p}, \quad (15.20)$$

где  $p$  — сопоставимые цены.

Сопоставимые цены не должны сильно отличаться от действующих (текущих) цен. Поэтому их периодически пересматривают, переходя к новым сопоставимым ценам (так, к ценам на 1.1.1982 г. перешли от цен на 1.1.1975 г.). Чтобы иметь возможность исчислять индексы объема продукции за длительные периоды, в течение которых применялись различные сопоставимые цены, продукции одного года оценивают как в прежних, так и в новых фиксированных ценах. Индекс за длительный период исчисляют цепным методом, т. е. путем перемножения индексов за отдельные отрезки этого периода, например:

$$\frac{\sum Q_{1975} p_{75}}{\sum Q_{1975} p_{75}} \cdot \frac{\sum Q_{1982} p_{82}}{\sum Q_{1975} p_{82}}$$

торыми допущениями можно рассматривать как базисный индекс в сопоставимых ценах  $\left(\frac{\sum Q_{\text{нр}}}{\sum Q_{\text{бр}}}\right)$ , хотя это связано с некоторой погрешностью.

Ряды индексов качественных показателей, которые экономически правильно взвешивать по весам текущего периода, строятся с переменными весами (см. выше ряд  $I_p$ ). Переход от цепных индексов к базисным также связан здесь с определенной погрешностью.

### 15.5. ПОСТРОЕНИЕ СВОДНЫХ ТЕРРИТОРИАЛЬНЫХ ИНДЕКСОВ

При построении *территориальных индексов*, т. е. при сравнении показателей в пространстве (межрайонные, межзаводские и др. сравнения), возникают вопросы о выборе *базы сравнения* и района (объекта), на уровне которого следует зафиксировать веса индекса. В каждом конкретном случае эти вопросы нужно решать исходя из задач исследования. Выбор базы сравнения зависит, в частности, от того, будут ли сравнения двусторонними (например, сравнение показателей двух соприкасающихся районов, заводов) или многосторонними (сравнение показателей нескольких районов, объектов).

При *двусторонних* сравнениях каждый район или объект с одинаковым основанием может быть принят как в качестве сравниваемого, так и в качестве базы сравнения (в отличие от индексов динамики, когда за базу сравнения принимается какой-либо прошлый период). В связи с этим возникает вопрос о фиксировании весов сводного индекса на уровне того или иного района (объекта). Пусть, например, нужно определить, в какой из двух областей и на сколько процентов ниже себестоимость единицы продукции и больше объем ее производства.

Таблица 15.3. Производство и себестоимость единицы продукции по областям

Вид продукции	Область				Расчет, млн. руб.					
	А		Б		$z_A Q_A$	$z_B Q_A$	$z_B Q_B$	$z_A Q_B$	$\bar{z}$ руб.	
	Производство, млн. т. $Q_A$	Себестоимость 1 т. руб. $z_A$	Производство, млн. т. $Q_B$	Себестоимость 1 т. руб. $z_B$						
Картофель	3	80	2	90	240	270	180	160	84	
Сахарная свекла	1	40	3	30	40	30	90	120	32,5	
Итого	—	—	—	—	280	300	270	280	—	

Если сравнивать область А с областью Б, зафиксировав в индексе себестоимости веса на уровне А, получим:  $I_z = \frac{\sum z_A Q_A}{\sum z_B Q_A} = \frac{280}{300} = 0,933$ , или 93,3 %, т. е. применительно к составу продукции в области А, себестоимость тонны продукции в А ниже, чем в Б, на 6,7 %. Однако, сопоставляя область Б и область А и фиксируя веса на уровне Б, найдем  $I_z = \frac{\sum z_B Q_B}{\sum z_A Q_B} = \frac{270}{280} = 0,964$ , т. е. применительно к составу продукции в области Б, себестоимость в этой области ниже, чем в А, на 3,6 %. Таким образом, в каждой области уровень себестоимости оказывается более низким, чем в другой, если веса фиксировать на уровне сравниваемой области.

Чтобы получить вполне определенный, однозначный результат, в зависимости от поставленной задачи возможны различные пути. Так, при двустороннем сравнении достаточно обоснованный и простой путь состоит в том, что в качестве весов принимаются объемы продукции в целом по обоим территориям ( $Q = Q_A + Q_B$ ):

$$I_z = \frac{\sum z_A Q_A}{\sum z_B Q_B} = \frac{80 \cdot 5 + 40 \cdot 4}{90 \cdot 5 + 30 \cdot 4} = \frac{560}{570} = 0,982, \text{ или } 98,2 \%,$$

т. е. в нашем случае уровень себестоимости в А ниже, чем в Б, на 1,8 %. Это подтверждает и обратный индекс:  $I_z = \frac{\sum z_B Q_B}{\sum z_A Q_A} = \frac{570}{560} = 1,018$ .

При *многосторонних* сравнениях, например, при сравнениях качественных показателей по нескольким областям, нужно соответственно расширить и границы территории, на уровне которой фиксируются веса.

Рассмотренный метод состоит в прямой стандартизации весов индекса. Возможен и косвенный метод стандартизации, учитывающий специфику соотношения весов на каждой из сравниваемых территорий.

$$I_z = \frac{\sum z_A Q_A}{\sum z_A Q_A} : \frac{\sum z_B Q_B}{\sum z_B Q_B},$$

где  $\bar{z} = \frac{z_A Q_A + z_B Q_B}{Q_A + Q_B}$  (см. табл. 15.3). В нашем примере

$$I_z = \frac{280}{84 \cdot 3 + 32,5 \cdot 1} : \frac{270}{84 \cdot 2 + 32,5 \cdot 3} = \frac{280}{284,5} : \frac{270}{265,5} = 0,968,$$

т. е. себестоимость 1 т в области А ниже, чем в Б, на 3,2 %.

Сумма экономии в А по сравнению с Б равна:  $(\sum z_A Q_A - \bar{z} Q_A) - (\sum z_B Q_B - \bar{z} Q_B) = (280 - 284,5) - (270 - 265,5) = -9,0$  млн. руб.

Косвенная стандартизация весов в отличие от прямой, обеспечивая взаимосвязь:  $I_z \cdot I_q = I_{zq}$ , если  $I_q$  строится так, как показано ниже.

В сводных территориальных индексах объемных показателей в качестве весов могут быть приняты *средние уровни* соответствующих качественных показателей, вычисленные в целом по сравниваемым территориям. Так, в нашем примере  $I_Q = \frac{\sum Q_A \bar{a}}{\sum Q_B \bar{a}} = \frac{3 \cdot 84 + 1 \cdot 32,5}{2 \cdot 84 + 3 \cdot 32,5} = 1,072$ , т. е. объем продукции в области А на 7,2 % больше, чем в Б (при взвешивании по средней себестоимости).

#### 15.6. АГРЕГАТНЫЕ ИНДЕКСЫ — ОСНОВНАЯ ФОРМА СОБСТВЕННО ИНДЕКСОВ. СРЕДНИЕ ИНДЕКСЫ

Построение собственно индексов в соответствии с поставленными перед ними экономическими задачами непосредственно приводит, как было показано выше, к *индексам в агрегатной форме*. В числителе и знаменателе агрегатных индексов находятся суммы произведений индексированных показателей на связанные с ними показатели — веса, зафиксированные на одном и том же уровне. Экономическое содержание этих сумм произведений в агрегатном индексе выражено непосредственно, что делает ясным и отчетливым смысл как индекса в целом, так и разности между величинами, находящимися в его числителе и знаменателе. Поэтому в нашей статистике агрегатная форма собственно индекса рассматривается как основная. Она представляет собой исходную базу для построения сводного индекса в другой, производной форме — в виде *средней величины из индивидуальных (или групповых) индексов*. При этом вопрос о форме средней и выборе ее весов решается путем преобразования формулы агрегатного индекса.

Идея построения сводного индекса в виде средней величины из индивидуальных (групповых) индексов вполне естественна: ведь сводный индекс является общей мерой, характеризующей среднюю величину изменения индексированного показателя, и, конечно, его величина должна зависеть от величин индивидуальных индексов. А критерием правильности построения сводного индекса в форме средней величины (среднего индекса) является его тождественность агрегатному индексу.

Преобразование агрегатного индекса в средний из индивидуальных (групповых) индексов производится следующим образом: *либо в числителе, либо в знаменателе агрегатного индекса индексированный показатель заменяется его выражением через соответствующий индивидуальный индекс*. Если такую замену сделать в числителе, то агрегатный индекс будет преобразован в средний арифметический, если же в знаменателе — то в средний гармонический из индивидуальных индексов. Практически замена обычно делается там, где в агрегатном индексе находится условная (расчетная, а не реальная) величина.

Возьмем, например, агрегатный индекс производительности труда (формула 15.186). Так как в этом индексе условные затраты времени ( $\sum i_q Q_1$ ) находятся в числителе, то из формулы индивидуального индекса производительности труда, выраженного, как и агрегатный, через трудоемкость единицы продукции ( $i_q = t_0 : t_1$ ), найдем базисную трудоемкость:  $t_0 = i_q t_1$ . Подставив в числитель агрегатного индекса вместо  $t_0$  выражение  $i_q t_1$ , получим:

$$I_q = \frac{\sum i_q Q_1}{\sum t_1 Q_1} = \frac{\sum i_q t_1 Q_1}{\sum t_1 Q_1} = \frac{\sum i_q T_1}{\sum T_1} \quad (15.21)$$

(с учетом того, что  $t_1 Q_1 = T_1$ ).

Допустим, что при производстве одного вида продукции производительность труда повысилась на 8 %, а при производстве другого — на 5 %, причем в отчетном периоде общие затраты рабочего времени ( $T_1$ ) на первый продукт составили 30, а на второй — 20 тыс. чел.-ч. Используя формулу (15.21), найдем:

$$I_q = \frac{\sum i_q T_1}{\sum T_1} = \frac{108 \cdot 30 + 105 \cdot 20}{50} = 106,8 \%$$

К расчету сводного индекса как среднего из индивидуальных (или групповых) приходится прибегать в тех случаях, когда отсутствуют данные, необходимые для исчисления индекса непосредственно в агрегатной форме, но известны индивидуальные (либо групповые) индексы и соответствующие веса. Так, в государственной розничной торговле учет товарооборота ведется в денежном выражении по группам товаров, данные же о натуральном количестве проданных товаров каждого отдельного вида обычно отсутствуют. В то же время расчетным путем могут быть получены групповые индексы цен по товарным группам. Это дает возможность исчислять средний индекс цен. Поскольку числитель агрегатного индекса ( $\sum p_1 Q$ ) известен, то преобразование (замена) производится в знаменателе, что приводит к среднему гармоническому индексу:  $i_p = p_1 : p_0$ , откуда  $p_0 = p_1 : i_p$  и

$$I_p = \frac{\sum p_1 Q_1}{\sum p_0 Q_1} = \frac{\sum p_1 Q_1}{\sum \frac{p_1 Q_1}{i_p}} \quad (15.22)$$

Расчет по этой формуле рассмотрим на следующем примере (табл. 15.4).

В целом (в среднем) по всем трем группам товаров цены снижались на 5,8 %, за счет чего покупатели сэкономили 1450 тыс. руб. (по группе А экономия 900 тыс. руб. (17 100 — 18 000), по группе Б — 750, по группе В — дополнительные затраты 200 тыс. руб. (2200 — 2000)).



Таблица 15.4. Товарооборот и изменение цен по группам товаров

Группа товаров	Товарооборот за отчетный период, тыс. руб.	Изменение цен в отчетном периоде по сравнению с базисным, %	Индекс цен (коэффициент)	Стоимость товаров отчетного периода по базисным ценам, тыс. руб.
	$p_1 Q_1$	$I_p - 100$	$I_p$	$p_1 Q_1 = p_1 Q_0 : I_p$
А	17 100	-5	0,95	18 000
Б	4 250	-15	0,85	5 000
В	2 200	+10	1,10	2 000
Итого	23 550	—	—	25 000

$$I_p = \sum p_1 Q_1 : \frac{p_1 Q_1}{I_p} = \frac{23550}{25000} = 0,942.$$

Нужно иметь в виду, что агрегатный индекс может быть преобразован и рассчитан как средний из индивидуальных индексов только при совпадении перечня видов продукции или товаров (их ассортимента) в отчетном и базисном периодах, т. е. когда агрегатный индекс построен по *сравнимому кругу единиц* (агрегатные индексы качественных показателей и агрегатные индексы объемных показателей при условии сравнимого ассортимента).

### 15.7. ИНДЕКСНЫЙ МЕТОД АНАЛИЗА ВЛИЯНИЯ ФАКТОРОВ

Социально-экономические явления и процессы связаны между собой, и это находит свое выражение во взаимосвязи соответствующих показателей. Одна из форм взаимосвязи между экономическими показателями состоит в том, что многие из них могут быть выражены в виде произведения нескольких других показателей. Так, общая выработка какого-либо вида продукции может быть представлена как произведение выработки продукции за единицу времени, т. е. уровня производительности труда, и общих затрат времени на эту продукцию:  $Q = qT$ ; товарооборот по данному товару равен произведению цены товара на проданное его количество и т. п.

Показатели-сомножители в подобных случаях выступают как *факторные показатели (факторы)*, от величины которых функционально зависит результирующий показатель. Так, изменение объема продукции является результатом изменения либо затрат времени ( $T$ ), либо производительности труда ( $q$ ), либо, наконец, результатом и того и другого.

В связи с этим при анализе динамики и выполнения плана возникает задача измерить роль отдельных факторов в изме-

нении данного явления, т. е. показать, как изменился результирующий показатель за счет изменения каждого фактора в отдельности. Решение этой задачи имеет существенное практическое значение. С народнохозяйственных позиций совсем не безразлично, например, за счет чего и насколько увеличилось производство продукции: за счет роста ли производительности труда, т. е. интенсивного фактора, или же в результате увеличения числа работников, т. е. экстенсивного фактора. Ведь рост численности работников имеет определенные пределы, а возможности повышения производительности труда в условиях научно-технического прогресса поистине безграничны. И именно этот интенсивный фактор должен сейчас являться основным источником роста производства.

Для оценки влияния отдельных факторов может быть использован индексный метод. При этом индексы решают *аналитическую* задачу и строятся с использованием специфики индексного метода — взвешивания при элиминировании изменения весов. Индексный метод факторного анализа применяется в тех случаях, когда между результирующим и факторным показателями существует функциональная связь. Поэтому прежде всего на основе экономического анализа нужно убедиться в наличии такой связи и установить ее форму. Форма связи между показателями должна отражать объективную зависимость одного явления от других и реальный характер влияния каждого фактора.

Мы рассмотрим *факторный индексный анализ* при двух наиболее распространенных в экономике *формах связи*: 1) когда результирующий показатель есть функция произведения двух или нескольких факторных показателей, т. е. когда мы имеем дело с одним видом продукции или материала ( $Q = qT$ ;  $M = mQ$  и т. п.), 2) когда результирующий показатель представляет собой сумму произведений показателей-факторов ( $\Sigma zQ$ ;  $\Sigma pQ$ ), т. е. имеется несколько видов продукции, товаров и т. п. Отдельно (см. раздел 15.8) будет рассмотрен случай, когда результирующий показатель является средним уровнем качественного показателя ( $\bar{q} = \frac{\Sigma qT}{\Sigma T}$ ;  $\bar{z} = \frac{\Sigma zQ}{\Sigma Q}$  при одинаковой

(однородной) продукции, производимой на разных участках производства). Во всех случаях после установления формы связи задача анализа влияния факторов решается путем построения системы взаимосвязанных индексов: в первом случае — индивидуальных, во втором — сводных, в последнем случае — особых индексов, которые будут рассмотрены ниже.

Рассмотрим анализ влияния факторов при одном виде единицы совокупности. Система взаимосвязанных индексов должна выразить индекс результирующего показателя в виде *произведения индексов факторных показателей (факторных индексов)*. Однако при этом и индивидуальные индексы нужно строить с использованием специфики индексного метода — взвешивания

и фиксирования весов. Веса необходимы здесь для того, чтобы выявить влияние факторов не только в относительном, но и в абсолютном выражении (математически они сокращаются).

В зависимости от того, на уровне какого периода фиксируются веса, возможны два метода построения системы взаимосвязанных индивидуальных индексов. *Первый метод* исходит из изучения влияния данного фактора при условии, что все другие факторы сохраняются на уровне базисного периода. Назовем его *методом выявления изолированного влияния факторов*. Поэтому веса во всех факторных индексах фиксируются на базисном уровне. Так, для взаимосвязи  $Q = qT$  будем иметь такую систему индексов:

$$\frac{Q_1}{Q_0} = \frac{q_1 T_0}{q_0 T_0} \cdot \frac{T_1 q_0}{T_0 q_0} \quad (15.23a)$$

*Второй метод* исходит из взаимосвязанного изучения влияния факторов. Он вытекает из того, что сущность факторов и характер их отношения одного к другому неодинаковы: наличие одних факторов есть необходимое условие проявления и возможности измерения других. Так, процесс труда, затраты рабочего времени — условия, необходимые для проявления того или иного уровня производительности труда и, следовательно, ее измерения. Поэтому сначала выявляется влияние объемного (исходного) фактора ( $T$ ) при фиксировании качественного фактора (весов) на уровне базисного периода, а затем — влияние качественного (производного) фактора ( $q$ ) при фиксировании весов на уровне отчетного периода:

$$\frac{Q_1}{Q_0} = \frac{T_1 q_0}{T_0 q_0} \cdot \frac{q_1 T_1}{q_0 T_1} \quad (15.23b)$$

Этот метод называется *цепным*. В случае трех и более факторов их следует располагать в определенной последовательности, вытекающей из их сущности, взаимосвязи и порядка расчета. При этом произведение любого числа факторов (начиная с начала или с конца последовательности) должно представлять собой более сложный факторный показатель, имеющий реальный экономический смысл. Например,  $y = v \cdot xz$ .

$$\left[ \begin{array}{c} \text{Объем} \\ \text{продукции} \end{array} \right] = \left[ \begin{array}{c} \text{Объем} \\ \text{продукции} \\ \text{Количество} \\ \text{затраченных} \\ \text{человеко-часов} \\ \text{выработка продукции} \\ \text{за 1 чел.-ч (качественный фактор) } q \end{array} \right] \cdot \left[ \begin{array}{c} \text{Количество} \\ \text{затраченных} \\ \text{человеко-часов} \\ \text{Количество} \\ \text{затраченных} \\ \text{человеко-дней} \\ \text{средняя продолжительность} \\ \text{рабочего дня в часах} \\ \text{(промежуточный фактор) } x \end{array} \right] \times$$

результативный показатель  $y$

$$(3) \times \left[ \begin{array}{c} \text{Количество} \\ \text{затраченных} \\ \text{человеко-дней} \\ \text{объемный фактор } z \end{array} \right]$$

где  $vx$  — выработка продукции за 1 чел.-день,  $xz$  — количество затраченных человеко-часов.

Часто (в том числе и в нашем примере) первым фактором-множителем берется тот, числителем расчетной формулы которого является результирующий показатель; следующим фактором-множителем берется тот, числителем расчетной формулы которого является знаменатель первого фактора и т. д. Последним множителем будет объемный фактор.

Фиксирование весов при построении системы индексов вторым методом производится так: в индексе объемного фактора (последнего в цепи множителей) веса (т. е. все другие факторы) фиксируются на уровне базисного периода; в каждом следующем факторном индексе (сходя от последнего фактора к первому) уже исследованные (стоящие справа) факторы фиксируются на уровне отчетного периода, а еще не исследованные (стоящие слева) — на уровне базисного периода. Тогда в индексе качественного фактора, стоящего первым в цепи, все веса будут зафиксированы на уровне отчетного периода. Так, в нашем примере будем иметь такую систему:

$$\frac{y_1}{y_0} = \frac{v_1 x_1 z_1}{v_0 x_1 z_1} \cdot \frac{v_0 x_1 z_1}{v_0 x_0 z_1} \cdot \frac{v_0 x_0 z_1}{v_0 x_0 z_0} \quad (15.24)$$

Знаменатель каждого индекса является здесь числителем последующего индекса, как и в цепи факторов. В результате факторные индексы оказываются как бы звеньями цепи множителей, в связи с чем этот метод получил название цепного метода. Обоснованием такого порядка фиксирования весов является градиация факторов, их переход от низшего порядка (объемный фактор) ко все более высокому (качественный фактор — первый в цепи). Рассмотренный порядок фиксирования весов обеспечивает возможность перехода в дальнейшем от оценки влияния изменения факторов на относительное изменение результирующего показателя к измерению их влияния на абсолютное его изменение.

Однако при обоих методах построения системы взаимосвязанных индексов и при методе выявления изолированного, абсолютного влияния факторов, и при методе взаимосвязанного, совместного их влияния (цепной метод) — относительный прирост результирующего показателя не совпадает по величине с суммой относительных приростов факторов. Так, если общие затраты времени возрастут на 5%, а производительность труда — на 10%, то объем продукции увеличится не на 15%, а на 15,5%, т. е.  $i_Q = i_v \cdot i_x = 1,05 \cdot 1,1 = 1,155$ , или 115,5%. Это

вызвано тем, что взаимосвязь между темпами прироста имеет более сложный вид, чем между темпами роста (индексами).

Покажем это для случая двух факторов:  $i_q = i_r \cdot i_q$ . Выразим каждый индекс (темп роста) через соответствующий темп прироста, обозначив последний так:

$$\frac{q_1 - q_0}{q_0} = \frac{\Delta Q}{Q_0} = \Delta'Q$$

и т. п. Тогда получим:  $1 + \Delta'Q = (1 + \Delta'T)(1 + \Delta'q)$ . Раскроем скобки и упростим:

$$\Delta'Q = \Delta'T + \Delta'q + \Delta'T \cdot \Delta'q, \quad (15.25a)$$

т. е. темп прироста результирующего показателя равен сумме темпов прироста факторов плюс произведение темпов их прироста.

Это обстоятельство несколько лишает наглядности интерпретацию относительного влияния факторов, так как, кроме двух исходных экономических факторов и темпов их прироста, появляется третье слагаемое  $(\Delta'T \cdot \Delta'q)$ , которому также нужно дать экономическую интерпретацию.

Многие экономисты и статистики рассматривают произведение темпов прироста факторов как эффект их взаимодействия в процессе изменения, т. е. как результат того, что факторы изменяются одновременно и взаимосвязанно, за счет чего и возникает дополнительный эффект — эффект их совместного действия. Есть еще и другие точки зрения по этому вопросу.

Графически это можно изобразить так. Пусть  $T_0 = 100$  чел.,  $T_1 = 120$  чел.,  $i_r = 1,2$ , или 120%;  $q_0 = 20$  т,  $q_1 = 28$  т,  $i_q = 1,4$ , или 140%. Будем откладывать на одной оси прямоугольной системы координат уровни одного фактора-сумматора (яля его индекс), а на другой оси — уровни (или индексы) второго фактора (рис. 15.1). Тогда произведение факторов, т. е. результирующий показатель ( $Q$ ) будет изображаться площадью соответствующего прямоугольника или квадрата, а прирост результирующего показателя изобразится как сумма трех площадей, расположенных на рисунке сверху и справа. Площадь  $T_0 \Delta q$  (или  $\Delta'q$ ) изображает прирост  $Q$  за счет изолированного изменения  $q$ , площадь  $q_0 \Delta T$  (или  $\Delta'T$ ) — за счет изолированного изменения  $T$ , а площадь  $\Delta T \Delta q$  (или  $\Delta'T \Delta'q$ ) —

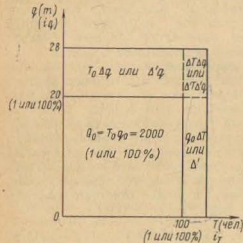


Рис. 15.1. Разложение прироста объема продукции по факторам

за счет взаимодействия факторов, т. е. их взаимосвязанного изменения.

Этот график позволяет наглядно представить влияние изменения факторов на результирующий показатель и в абсолютном выражении, т. е. абсолютный прирост результирующего показателя за счет изменения каждого фактора в отдельности. Решение этой задачи называется разложением абсолютного прироста по факторам.

Математически это выглядит так. Обозначим, как и ранее,  $Q_1 - Q_0 = \Delta Q$ ,  $q_1 - q_0 = \Delta q$ ,  $T_1 - T_0 = \Delta T$ . Тогда  $\Delta Q = Q_1 - Q_0 = q_1 T_1 - q_0 T_0 = (q_0 + \Delta q)(T_0 + \Delta T) - q_0 T_0$ , откуда

$$\Delta Q = T_0 \Delta q + q_0 \Delta T + \Delta q \Delta T, \quad (15.25b)$$

что аналогично полученной выше формуле (15.25a). Здесь  $T_0 \Delta q$  — это прирост  $Q$  за счет изолированного изменения фактора  $q$ ;  $q_0 \Delta T$  — прирост  $Q$  за счет изолированного изменения  $T$ , а  $\Delta q \Delta T$  — эффект взаимодействия факторов, т. е. дополнительный прирост  $Q$  за счет взаимосвязанного их изменения.

Сопоставляя эти результаты с формулой (15.23a), нетрудно убедиться, что приросты  $Q$  за счет отдельных факторов представляют собой разность между числителями и знаменателями соответствующих факторных индексов, построенных первым методом (изолированного влияния факторов). Действительно, обозначив  $T_0 \Delta q$  через  $\Delta Q_q$ ,  $q_0 \Delta T$  — через  $\Delta Q_T$ , получим:

$$\Delta Q_q = T_0 \Delta q = T_0 (q_1 - q_0) = T_0 q_1 - T_0 q_0; \quad (15.26a, б)$$

$$\Delta Q_T = q_0 \Delta T = q_0 (T_1 - T_0) = q_0 T_1 - q_0 T_0.$$

Это справедливо при любом числе факторов. Что же касается эффекта взаимодействия, то он может быть вычислен как разность между общим приростом  $Q$  и суммой приростов за счет отдельных факторов, т. е. как остаточная величина:

$$\Delta Q_{\text{взим}} = \Delta Q - \sum \Delta Q_{\text{факт.}}$$

В нашем примере (использованном при построении графика)

$$\begin{aligned} \Delta Q_T &= q_0 (T_1 - T_0) = 20 \cdot (120 - 100) = 400 \text{ т}; & \Delta Q_q &= T_0 (q_1 - q_0) = \\ &= 100 \cdot (28 - 20) = 800 \text{ т}; & \Delta Q_{\text{взим}} &= \Delta T \Delta q = (120 - 100) (28 - 20) = \\ &= 160 \text{ т (или } \Delta Q_{\text{взим}} = \Delta Q - \sum \Delta Q_{\text{факт.}} = (Q_1 - Q_0) - (\Delta Q_T + \Delta Q_q) = \\ &= (3360 - 2000) - (400 + 800) = 160 \text{ т).} \end{aligned}$$

Таким образом, за счет обособленно рассматриваемого изменения  $T$  объем продукции увеличился на 400 т, за счет обособленного роста производительности труда — на 800 т, а за счет взаимодействия факторов дополнительно было получено еще 160 т продукции.

Рассмотрим теперь, как решается задача разложения прироста по факторам при втором (цепном) методе построения системы взаимосвязанных индексов.



При использовании цепного метода построения системы индексов, как и при первом методе, приросты результативного показателя за счет изменения отдельных факторов также исчисляются как *разность между числителем и знаменателем соответствующих факторных индексов* (см. формулы 15.236 и 15.242). Однако в результате принятой при цепном методе системы взвешивания сумма приростов результативного показателя за счет отдельных факторов равна при этом общему его приросту. Таким образом, общий прирост разлагается на столько же частей (слагаемых), на сколько факторов-составителей разложен результативный показатель, т. е. сколько исследуется отдельных факторов. Тот член разложения, который при выявлении изолированного влияния факторов рассматривается как результат их взаимодействия (см. рис. 15.1 и формулу 15.25а, б), присоединяется теперь к приросту за счет качественного фактора, а если таких факторов несколько, то распределяется между ними.

Так, для случая  $Q = qT$  на основе формулы (15.236) получим:

$$\Delta Q_T = T_1 q_0 - T_0 q_0 = (T_1 - T_0) q_0 = q_0 \Delta T; \quad (15.27a)$$

$$\Delta Q_q = q_1 T_1 - q_0 T_1 = (q_1 - q_0) T_1 = T_1 \Delta q. \quad (15.27b)$$

Итак, при цепном методе (применительно к одному виду единиц совокупности и двум факторам) абсолютный прирост результативного показателя за счет объемного фактора ( $T$ ) равен приросту самого объемного фактора, умноженному на базисный уровень качественного фактора ( $q_0$ ); прирост же за счет качественного фактора равен приросту этого качественного фактора, умноженному на уровень объемного фактора в отчетном периоде.

В нашем примере (см. с. 346)  $\Delta Q_T = (120 - 100) \cdot 20 = 400$  т, т. е. столько же, сколько и при методе изолированного влияния;  $\Delta Q_q = (28 - 20) \cdot 120 = 960$  т. Общий прирост продукции равен 1360 т (400 + 960). Следовательно, из 1360 т общего прироста продукции лишь 960 т были получены за счет роста производительности труда — интенсивного фактора. Это составляет только 71 % общего прироста продукции, принятого за 100 % ( $960 : 1360 = 0,706$ ). Остальные 29 % общего прироста были получены за счет экстенсивного фактора — увеличения численности рабочих.

Для рассмотренных метода анализа влияния отдельных факторов в общем случае приводят к разным результатам. В литературе предлагаются и другие методы разложения прироста по факторам, при которых эффект взаимодействия ( $\Delta T \Delta q$ ) тем или иным способом распределяется между факторами: либо поровну, либо пропорционально темпам их прироста, либо иным путем. Однако наиболее разработанным теоретически (несмотря на некоторые допущения и условности)

и наиболее широко применяемым в практике является цепной метод.

В целях наглядности представим схему анализа влияния двух факторов цепным методом, которая отражает порядок фиксирования весов в факторных индексах (но не действительную последовательность изменения факторов, изменяющихся обычно одновременно) и позволяет легко получить формулу (15.27):

$$Q_0 = T_0 q_0 \xrightarrow{T} Q' = T_1 q_0 \xrightarrow{q} Q_1 = T_1 q_1.$$

Стрелки условно изображают изменение фактора, обозначенного над ними. Разность между величиной после изменения фактора и до его изменения равна приросту объема продукции за счет этого фактора, а частное от деления этих величин — соответствующему факторному индексу системы (15.236).

Рассмотренный выше способ расчетов при разложении прироста по факторам цепным методом условно может быть назван *разностным*, так как в формуле (15.27) используются приросты, т. е. разности уровней факторов. При цепном методе возможен и иной порядок расчетов  $\Delta Q_T$  и  $\Delta Q_q$ . Разделив обе части обоих равенств (15.27) на  $Q_0$ , получим:

$$\frac{\Delta Q_T}{Q_0} = \frac{T_1 q_0 - T_0 q_0}{Q_0} = \frac{T_1 q_0 - T_0 q_0}{T_0 q_0} = i_T - 1;$$

$$\frac{\Delta Q_q}{Q_0} = \frac{q_1 T_1 - q_0 T_1}{Q_0} = \frac{q_1 T_1 - q_0 T_1}{q_0 T_0} = i_q - i_T.$$

Если для простоты обозначений принять  $\frac{\Delta Q_T}{Q_0} = \Delta' Q_T$  и  $\frac{\Delta Q_q}{Q_0} = \Delta' Q_q$ , будем иметь:

$$\Delta' Q_T = i_T - 1; \quad (15.28a)$$

$$\Delta' Q_q = i_q - i_T. \quad (15.28b)$$

Следовательно, в условиях цепного метода при двух факторах относительный прирост результативного показателя за счет объемного фактора равен относительному приросту самого объемного фактора. Иными словами, на сколько процентов увеличивается объемный фактор, на столько же процентов за счет этого увеличивается и результативный показатель по сравнению с базисным уровнем. А относительный прирост результативного показателя за счет качественного фактора равен разности между индексами числителя и знаменателя расчетной формулы этого качественного фактора (в примере расчетная формула  $q = \frac{Q}{T}$ ).

Это справедливо не только для двух, но и для любого числа факторов при условии, что результативный показатель является объемным, а разложение производится цепным методом. Так,

поскольку  $m = \frac{M}{Q}$ , а  $M = mQ$  — показатель объемный, будем иметь:

$$\Delta' M_Q = i_Q - 1 \quad \text{и} \quad \Delta' M_m = i_m - i_Q.$$

Если в формуле (15.24)  $y$  и  $z$  — объемные показатели, причем  $v = \frac{y}{w}$ ,  $x = \frac{w}{z}$ , то  $\Delta' y_z = i_z - 1$ ,  $\Delta' y_x = i_w - i_z$ ,  $\Delta' y_v = i_y - i_w$ .

Этот способ расчетов условно можно назвать индексным. При этом для перехода от относительных приростов к абсолютным нужно относительные приросты умножить на базисный уровень результативного показателя. Пусть, например, объем продукции в 1980 г. составлял 1000 т, а в 1985 г. — 1265 т, причем общие затраты рабочего времени увеличились на 2 % ( $i_T = 1,02$ ). Найдем  $i_Q = 1265 : 1000 = 1,265$  и по формуле (15.28 а, б) получим:  $\Delta' Q_T = 0,02$ ;  $\Delta' Q_Q = 0,265 - 0,02 = 0,245$ . Умножив эти относительные приросты на  $Q_0 = 1000$ , найдем, что за счет увеличения затрат времени выпуск продукции возрос на 20 т (или на 2 %), а за счет роста производительности труда — на 245 т (или на 24,5 %). Отметим, что производительность труда при этом повысилась только на 24 % ( $i_Q = i_Q : i_T = 1,265 : 1,02 = 1,240$ ). Таким образом, за счет роста производительности труда было получено 92,5 % общего прироста продукции ( $245 : 265 = 0,9245$ ), что соответствует заданиям 11-й пятилетки. Этот расчет можно сделать и с помощью индексов:  $\Delta Q_Q : \Delta Q = (i_Q - i_T) : (i_Q - 1) = (1,265 - 1,02) : (1,265 - 1) = 0,925$ .

Рассмотрим влияние факторов при нескольких видах единиц совокупности (применительно к сводным агрегатным индексам). Системы взаимосвязанных сводных индексов также могут быть построены двумя методами: методом выявления изолированного влияния изменения факторов и цепным методом. Фиксирование весов в факторных индексах производится при этом так же, как и в индивидуальных индексах: при методе изолированного влияния веса во всех факторных индексах фиксируются на базисном уровне, а при цепном методе в индексе объемного фактора — на базисном, в индексе качественного фактора — на отчетном уровне.

Однако при первом методе сводный индекс результативного показателя (в отличие от индивидуального индекса) в общем случае не равен произведению индексов факторов. Это объясняется тем, что сводные индексы факторов являются средними величинами из индивидуальных индексов, которые обычно варьируют, так что в процессе изменения факторов между их индивидуальными индексами могут сложиться эмпирические корреляционные связи, что усложняет взаимодействие факторов. Поэтому сводный индекс результативного показателя равен произведению не только факторных индексов, но и еще одного индекса-множителя — индекса ковариации (взаимо-

связанной, совместной вариации). Например, сводный индекс затрат на производство (общей себестоимости —  $zQ$ ) может быть разложен так:

$$\frac{\Sigma z_1 Q_1}{\Sigma z_0 Q_0} = \frac{\Sigma z_1 Q_0}{\Sigma z_0 Q_0} \cdot \frac{\Sigma z_0 Q_1}{\Sigma z_0 Q_0} \cdot \left( \frac{\Sigma z_1 Q_1}{\Sigma z_0 Q_1} : \frac{\Sigma z_1 Q_0}{\Sigma z_0 Q_0} \right). \quad (15.29a)$$

Первый и второй индексы в правой части этого равенства характеризуют степень влияния обособленного изменения факторов  $z$  и  $Q$  на результативный показатель. Третий же индекс — индекс ковариации (заключенный в скобки) — измеряет степень дополнительного влияния взаимосвязанной вариации индивидуальных индексов факторов.

В теории индексов доказывается, что индекс ковариации может быть выражен так:  $I_{cov} = 1 + r_{zQ} r_{zQ} r_{zQ}$ , где  $r_{zQ}$  — коэффициент корреляции между  $i_z$  и  $i_Q$ , а  $r_{zQ}$  и  $r_{zQ}$  — коэффициенты вариации  $i_z$  и  $i_Q$ . Отметим также, что индекс этого вида уже рассматривался нами выше (см. формулу 15.17б, в). Как видно из формулы,  $I_{cov} = 1$ , если хотя бы один из индивидуальных индексов не варьирует.

При использовании цепного метода индекс результативного показателя равен произведению факторных индексов:

$$\frac{\Sigma z_1 Q_1}{\Sigma z_0 Q_0} = \frac{\Sigma z_1 Q_1}{\Sigma z_0 Q_1} \cdot \frac{\Sigma z_0 Q_1}{\Sigma z_0 Q_0}. \quad (15.29b)$$

Влияние ковариации включено здесь в индекс качественного фактора ( $I_z$ ).

Разложение абсолютного прироста результативного показателя по факторам производится аналогично тому, как это делается при одном виде единиц, но на базе сводных индексов. При обоих методах анализа влияние каждого фактора в абсолютном выражении рассчитывается как разность между числителем и знаменателем соответствующего факторного индекса. Рассмотрим это на примере, когда какой-либо материал расходуется на два вида продукции (см. табл. 15.5).

Следовательно, за счет изолированного роста объема продукции общий расход материала увеличился на 260 т. Если бы изменение удельного расхода происходило при базисном объеме продукции, то общий расход материала уменьшился бы на 120 т. За счет взаимодействия  $Q$  и  $m$  расход материала дополнительно уменьшился на 18 т. Общее увеличение расхода составило 122 т ( $260 - 120 - 18$ ).

В последней графе таблицы показано изменение общего расхода материала за счет изменения удельного расхода и взаимодействия факторов (138 т). Это совпадает с его изменением за счет удельного расхода при использовании цепного метода разложения прироста по факторам:  $\Sigma m_1 Q_1 - \Sigma m_0 Q_1$ .

По данным каждой строки таблицы видно, как сложились эти результаты из результатов по отдельным видам продукции, в частности из экономии за счет снижения удельного расхода по продукту А и дополнительного расхода по продукту Б.

Таблица 15.5. Выявление изолированного влияния и взаимодействия факторов на общий расход материала

Вид продукции	Выпуск, тыс. шт.		Удельный расход материала, кг. шт.		Общий расход материала, т		$m_1 Q_1$	$m_1 Q_0$	Общие изменения расхода материала $\Delta M = M_1 - M_0$	в том числе в результате				Изменения за счет изолированного влияния $Q$ и $m$
	$Q_0$	$Q_1$	$m_0$	$m_1$	$M_0$	$M_1$				изолированного изменения $Q$ и $m$		в про- цессе их взаимо- действия $Q$ и $m$		
										объема продук- ции	удель- ного расхода мате- риала			
А	1	2	3	4	5	6	7	8	9	10	11	12	13	
А	35	40	40	36	1400	1440	1600	1260	40	200	-140	-20	-160	
Б	20	22	30	31	600	682	660	620	82	60	20	2	22	
Итого	—	—	—	—	2000	2122	2260	1880	122	260	-120	-18	-138	

## 15.8. ИНДЕКСНЫЙ МЕТОД АНАЛИЗА ДИНАМИКИ (ВЫПОЛНЕНИЯ ПЛАНА) СРЕДНЕГО УРОВНЯ КАЧЕСТВЕННОГО ПОКАЗАТЕЛЯ

На разных участках производства (предприятиях, объединениях и других хозяйственных системах или в разных регионах) часто производится одноименная продукция (полностью одинаковая либо такая, которую можно, хотя бы с оговорками, суммировать). Уровень различных *качественных* показателей, относящихся к этой продукции (производительность труда, себестоимость и трудоемкость единицы изделия и т. п.), обычно бывает при этом неодинаков, а его вариация по участкам достигает значительных размеров. Средний же уровень *качественного* показателя, исчисленный в целом по всем таким участкам, зависит как от уровней на отдельных участках, так и от удельных весов этих участков в общих затратах труда, в общем объеме продукции и т. п. Так, средняя себестоимость единицы продукции данного вида зависит от уровней себестоимости на отдельных участках производства и от удельных весов продукции, произведенной на этих участках, в общем ее объеме:

$$\bar{z} = \frac{\sum zQ}{\sum Q} = \sum z d_Q, \quad (15.30a)$$

где  $d_Q = Q : \sum Q$  — удельный вес продукции участка в общем ее объеме.

Аналогично обстоит дело со средним удельным расходом какого-либо материала, затрачиваемого на одноименную продукцию:

$$\bar{m} = \frac{\sum M}{\sum Q} = \frac{\sum mQ}{\sum Q} = \sum m d_Q. \quad (15.30b)$$

Средний уровень производительности труда зависит от ее уровней на отдельных участках и от удельных весов этих участков в общих затратах рабочего времени на данную продукцию (или в общей численности рабочих, работающих):

$$\bar{q} = \frac{\sum Q}{\sum T} = \frac{\sum qT}{\sum T} = \sum q d_T, \quad (15.30b)$$

где  $d_T = T : \sum T$ .

Следовательно, динамика, изменение среднего уровня *качественного* показателя обусловлено действием двух факторов: 1) изменением *уровней* *качественного* показателя на отдельных участках и 2) изменением *удельных* *весов* этих участков в *общем* *объеме* *показателя*, с помощью которого вычисляется *средний* *уровень*, т. е. изменением *структуры*, *структурными* *сдвигами*.

В связи с этим при анализе динамики (или степени выполнения плана) среднего уровня *качественного* показателя возникает вопрос о том, в какой мере изменение среднего уровня было вызвано действием каждого из этих двух факторов в от-



дельности. Например, на сколько тонн (штук и т. п.) и на сколько процентов изменился средний уровень производительности труда за счет изменения ее уровня на отдельных участках и на сколько — за счет изменения удельных весов участков в общем объеме затрат труда, т. е. за счет структурных сдвигов.

Ответы на подобные вопросы имеют большое познавательное и практическое значение. В наших планах для повышения эффективности производства предусматривается использование обоих этих факторов — внутрипроизводственного и структурного. Известно, например, что добыча угля открытым способом (в разрезах, карьерах) обходится дешевле, чем подземным способом (в шахтах), а производительность труда при открытом способе в несколько раз выше, чем в шахтах. Именно поэтому в 11-й пятилетке добычу угля наиболее эффективным открытым способом предусмотрено развивать опережающими темпами, увеличивая тем самым ее удельный вес в общей добыче.

Анализ динамики среднего уровня качественного показателя до известной степени аналогичен анализу влияния факторов на изменение результативного показателя, представляющего сумму произведений двух факторов. Однако при этом есть и существенные различия: вместо качественного и объемного факторов мы имеем теперь дело с уровнями качественного показателя на разных участках производства и с удельными весами этих участков в общем объеме признака, используемого при взвешивании, а вместо единиц совокупности разного вида — совокупности участков производства, на которых производится одноименная продукция, и т. п.

Анализ динамики (или степени выполнения плана) среднего уровня производится путем построения системы взаимосвязанных индексов. При этом индекс среднего уровня, характеризующий его динамику за счет изменения всех факторов в целом, разлагается на произведение аналитических индексов-множителей, каждый из которых отражает изменение только одного фактора и тем самым влияние этого изменения на динамику среднего уровня. Система взаимосвязанных индексов может быть построена (как и при анализе влияния факторов) либо методом выделения изолированного влияния факторов, либо цепным методом.

При выявлении изолированного влияния факторов система взаимосвязанных индексов среднего уровня производительности труда будет такова:

$$\bar{q}_1 = \frac{\sum q_1 T_0}{\sum q_0 T_0} = \left( \frac{\sum q_0 T_1}{\sum T_1} : \frac{\sum q_0 T_0}{\sum T_0} \right) \cdot \left( \frac{\sum q_1 T_1}{\sum q_0 T_1} : \frac{\sum q_1 T_0}{\sum q_0 T_2} \right) \quad (15.31a)$$

или после замены  $T: \sum T = d_T$ :

$$\frac{\sum q_1 d_{T_1}}{\sum q_0 d_{T_0}} = \frac{\sum q_1 d_{T_0}}{\sum q_0 d_{T_0}} \cdot \left( \frac{\sum q_1 d_{T_1}}{\sum q_0 d_{T_1}} : \frac{\sum q_1 d_{T_0}}{\sum q_0 d_{T_0}} \right) \quad (15.31b)$$

Индекс в левых частях этих равенств представляет отношение среднего уровня производительности труда в отчетном периоде к среднему ее уровню в базисном периоде, т. е.  $\frac{q_1}{q_0}$  является индексом среднего уровня

производительности труда. Он характеризует изменение  $\bar{q}$  в целом за счет изменения всех факторов и называется индексом переменного состава, ибо удельные веса ( $d_T = T: \sum T$ ), выражающие состав, т. е. структуру, в его числителе и знаменателе неодинаковы.

Первый индекс-множитель показывает, как изменился бы средний уровень производительности труда только за счет изменения ее уровней на отдельных участках при сохранении удельных весов участков в общих затратах труда на уровне базисного периода. Этот индекс называется индексом фиксированного (постоянного) состава. Он может быть представлен как средняя арифметическая взвешенная из индексов на отдельных участках ( $i_{q_1}$ ).

Второй индекс-множитель показывает, как изменился бы средний уровень производительности труда только за счет структурных сдвигов, т. е. только за счет изменения удельных весов участков в общих затратах труда при сохранении базисных ее уровней на отдельных участках. Этот индекс называется индексом влияния структурных сдвигов на средний уровень производительности труда.

Наконец, третий индекс-множитель показывает, как изменился средний уровень производительности труда за счет взаимосвязанного изменения ее уровней на отдельных участках ( $i_{q_1}$ ) и их удельных весов в общих затратах времени ( $d_T$ ). Он может быть представлен и как индекс ковариации:  $1 + i_{q_1} d_T$ .

Метод выявления изолированного влияния мало еще применяется в практических исследованиях. Чаще используются второй метод — цепной. При цепном методе анализа динамики среднего уровня его общее изменение за счет всех факторов в целом характеризуется отношением среднего уровня качественного показателя за отчетный период к среднему уровню за базисный период (как и в случае метода выявления изолированного влияния). Это отношение, называемое индексом переменного состава<sup>1</sup>, разлагается на произведение двух индексов-множителей, выполняющих аналитические функции: индекс фиксированного (постоянного) состава и индекс влияния структурных сдвигов на динамику среднего уровня. При этом используются две различные системы взвешивания аналитических индексов. В одной из них индекс фиксированного состава, в котором изменяются только уровни качественного показателя ( $q$ ) на отдельных участках, строится с весами базисного периода (как и при методе изолированного влияния), а индекс влияния структурных сдвигов — с весами отчетного периода (весами в этом индексе являются уровни качественного показателя ( $q_1$ )). Применительно к динамике

<sup>1</sup> Поскольку в этом индексе нет одной из специфических особенностей индексного метода — фиксирования весов, — этот показатель является индексом лишь в широком понимании этого термина, и его можно было бы обозначить  $i_{\bar{q}}$ .

среднего уровня производительности труда система индексов выглядит так (как и выше,  $d_T = T : \Sigma T$ ):

$$\frac{\bar{q}_1}{\bar{q}_0} = \frac{\sum q_1 T_0}{\sum q_0 T_0} \cdot \left( \frac{\sum q_1 T_1}{\sum T_1} : \frac{\sum q_1 T_0}{\sum T_0} \right) = \frac{\sum q_1 d_{T_0}}{\sum q_0 d_{T_0}} \cdot \frac{\sum q_1 d_{T_1}}{\sum q_1 d_{T_0}},$$

Индекс фиксированного состава
Индекс влияния структурных сдвигов на  $\bar{q}$ 
 $I_{d_T}$ 
 $\bar{I}_q$ 
 $I_{d_T}$

(15.32а, б)

В целях наглядности даем схему анализа, отражающую порядок фиксирования весов при построении этой системы индексов (но не действительную последовательность изменения факторов), используя различную символику (обозначения над стрелками указывают анализируемый фактор):

$$\begin{aligned} \frac{\sum q_0 T_0}{\sum T_0} &\xrightarrow{q} \frac{\sum q_1 T_0}{\sum T_0} \xrightarrow{d_T} \frac{\sum q_1 T_1}{\sum T_1}; \\ \sum q_0 d_{T_0} &\xrightarrow{q} \sum q_1 d_{T_0} \xrightarrow{d_T} \sum q_1 d_{T_1}; \\ \bar{q}_0 &\xrightarrow{q} \bar{q} \xrightarrow{d_T} \bar{q}_1. \end{aligned}$$

Каждый индекс системы может быть построен путем деления величины, полученной после изменения фактора, на величину, которая была до этого изменения. Здесь  $\bar{q} = \sum q_1 T_0 : \sum T_0 = \sum q_1 d_{T_0}$  — условный (расчетный) средний уровень производительности труда, который был бы при отчетных ее уровнях на каждом участке, но при базисных удельных весах участков в общих затратах труда.

Систему индексов можно представить так:

$$\frac{\bar{q}_1}{\bar{q}_0} = \frac{\bar{q}'}{\bar{q}_0} \cdot \frac{\bar{q}_1}{\bar{q}'} \quad \text{или} \quad I_q = \bar{I}_q \cdot I_{d_T}. \quad (15.32в)$$

Индекс фиксированного состава в этой системе ( $\bar{I}_q$ ) показывает, как изменился бы средний уровень производительности труда только за счет изменения ее уровней на отдельных участках, т. е. при сохранении базисных удельных весов участков в общем объеме затрат труда ( $d_{T_0}$ ). Этот индекс может быть преобразован в средний арифметический из индексов на отдельных участках ( $\bar{I}_q = \frac{\sum I_q q_0}{\sum q_0}$ ), т. е. он измеряет среднее изменение производительности труда на отдельных участках и относительное влияние этого фактора на средний уровень производительности труда.

Индекс влияния структурных сдвигов показывает, как изменился бы средний уровень производительности труда за счет структурных сдвигов, т. е. только за счет изменения удельных весов участков в общих затратах труда, применительно к отчетному уровню производительности труда на каждом участке.

Следует, однако, отметить, что само по себе изменение удельных весов участков в общих затратах труда — отнюдь не достаточное условие, чтобы индекс структурных сдвигов отразил влияние структурного фактора на средний уровень. В данном случае для этого необходимо также, чтобы в отчетном периоде была вариация уровней производительности труда по участкам и коэффициент корреляции между  $q_1$  и  $i_T$  не оказался равным нулю. Это следует из того, что индекс структурных сдвигов здесь (15.32) может быть представлен как  $1 + r_{q_1 T_0} q_1 i_T$ .

При другой системе взвешивания индексе производительности труда фиксированного состава строится с весами отчетного периода ( $T_1$  или  $d_{T_1}$ ), а индекс влияния структурных сдвигов — с весами базисного периода ( $q_0$ ):

$$\frac{\bar{q}_1}{\bar{q}_0} = \frac{\sum q_1 T_1}{\sum q_0 T_1} \cdot \left( \frac{\sum q_0 T_1}{\sum T_1} : \frac{\sum q_0 T_0}{\sum T_0} \right) = \frac{\sum q_1 d_{T_1}}{\sum q_0 d_{T_1}} \cdot \frac{\sum q_0 d_{T_1}}{\sum q_0 d_{T_0}},$$

(15.33а, б)

Обозначив  $\sum q_0 T_1 : \sum T_1$  через  $\bar{q}'$ , систему индексов можно представить и так:

$$\frac{\bar{q}_1}{\bar{q}_0} = \frac{\bar{q}'}{\bar{q}_0} \cdot \frac{\bar{q}_1}{\bar{q}'} \quad \text{или} \quad I_q = I_{d_T} \cdot \bar{I}_q. \quad (15.33в)$$

Соответственно здесь изменяется и статистико-экономический смысл этих двух индексов: индекс фиксированного состава характеризует изменение среднего уровня производительности труда за счет ее изменения на отдельных участках применительно к их удельным весам в общих затратах труда в отчетном периоде, а индекс влияния структурных сдвигов — их влияние применительно к базисным уровням производительности труда на отдельных участках.

Схема анализа здесь такова:

$$\begin{aligned} \frac{\sum q_0 T_0}{\sum T_0} &\xrightarrow{d_T} \frac{\sum q_0 T_1}{\sum T_1} \xrightarrow{q} \frac{\sum q_1 T_1}{\sum T_1}; \\ \sum q_0 d_{T_0} &\xrightarrow{d_T} \sum q_0 d_{T_1} \xrightarrow{q} \sum q_1 d_{T_1}; \\ \bar{q}_0 &\xrightarrow{d_T} \bar{q}' \xrightarrow{q} \bar{q}_1. \end{aligned}$$

В статистической литературе данная система взвешивания аналитических индексов рекомендуется чаще, чем первая. Так,

при анализе динамики средней себестоимости единицы продукции обычно рекомендуется такая система индексов:

$$\frac{\bar{z}_1}{\bar{z}_0} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \left( \frac{\sum z_0 q_1}{\sum z_0 q_0} \cdot \frac{\sum z_0 q_0}{\sum z_0 q_1} \right) = \frac{\sum z_1 d_{q_1}}{\sum z_0 d_{q_1}} \cdot \frac{\sum z_0 d_{q_1}}{\sum z_0 d_{q_0}}, \quad (15.34a, б)$$

где  $d_q = Q : \Sigma Q$  — удельные веса участков в общем объеме отчетной продукции;  $I_z$  — индекс средней себестоимости переменного состава;  $I_z$  — индекс себестоимости фиксированного состава;  $I_{d_q}$  — индекс влияния структурных сдвигов на среднюю себестоимость, измеряющий влияние на нее изменений удельных весов участков в общем объеме продукции.

Аналогично строятся аналитические индексы и других качественных показателей (урожайности, удельного расхода и т. д.): в индексах фиксированного состава веса фиксируются на отчетном уровне, а индекс структурных сдвигов строится с весами базисного периода.

Как при методе выявления изолированного влияния факторов, так и при обоих вариантах цепного метода каждый аналитический индекс-множитель измеряет степень влияния соответствующего фактора на средний уровень качественного показателя, т. е. характеризует это влияние в относительном выражении. Чтобы найти абсолютный прирост среднего уровня за счет того или иного фактора, целесообразно использовать индексы, числители и знаменатели которых выражены в виде средних уровней, т. е. для производительности труда индексы типа (15.31б; 15.32б, в; 15.33б, в). Разность между числителем и знаменателем каждого из этих индексов и представляет собой абсолютный прирост среднего уровня за счет соответствующего фактора.

В системе (15.31б) абсолютный прирост за счет взаимодействия  $i_q$  и  $i_T$  равен  $(\sum q_1 d_{T_1} - \sum q_0 d_{T_1}) - (\sum q_1 d_{T_0} - \sum q_0 d_{T_0})$ .

Рассмотрим индексный метод анализа среднего уровня производительности труда на следующем примере (табл. 15.6).

В шахтах производительность труда повысилась на 6,7 % ( $i_q = q_1 : q_0 = 0,64 : 0,6$ ), в разрезах — на 10 %, а средний ее уровень возрос на 962 т, или на 48,1 %:

$$I_q \text{ перем. сост.} = \frac{\sum q_1}{\sum T_1} : \frac{\sum q_0}{\sum T_0} = \frac{\bar{q}_1}{\bar{q}_0} = \frac{2,962}{2,0} = 1,481, \text{ или } 148,1 \, \%. \quad (15.34a, б)$$

Такой парадоксальный результат объясняется тем, что, кроме роста производительности труда на отдельных участках на средний ее уровень влиял и другой фактор — структурные сдвиги в распределении рабочих по участкам производства. А здесь произошли коренные изменения: удельный вес разрезов, где производительность труда гораздо выше, увеличился

с 40 до 60 %. Естественно, что это должно было повысить средний уровень производительности труда.

Если бы удельные веса шахт и разрезов сохранялись на базисном уровне, то за счет роста производительности труда на

Таблица 15.6. Добыча угля, число рабочих и производительность труда в шахтах и разрезах

Участки произ-водства (способ добычи)	Базисный год				Отчетный год			
	Добыча угля, тыс. т $Q_0$	Среднесписочное число рабочих		Добыча на одного рабочего, тыс. т $Q_1$	Добыча угля, тыс. т $Q_1$	Среднесписочное число рабочих		Добыча на одного рабочего, тыс. т $Q_1$
		человек $T_0$	в % к $T_0$ $d_{T_0}$			человек $T_1$	в % к $T_0$ $d_{T_1}$	
Шахты	360	600	60	0,6	320,0	500	40	0,64
Разрезы	1640	400	40	4,1	3382,5	750	60	4,51
Итого	2000	1000	100	2,0	3702,5	1250	100	2,962

отдельных участках ее средний уровень повысился бы только на 9,4 %:

$$\bar{I}_q \text{ фикс. сост.} = \frac{\sum q_1 T_0}{\sum T_0} = \frac{0,64 \cdot 600 + 4,51 \cdot 400}{2000} = \frac{2188}{2000} = 1,094$$

$$(\text{или } \bar{I}_q \text{ фикс. сост.} = \frac{\sum q_1 d_{T_0}}{\sum q_0 d_{T_0}} = \frac{0,64 \cdot 0,6 + 4,51 \cdot 0,4}{2,0} = \frac{2,188}{2,0} = 1,094).$$

В абсолютном выражении это составило 188 т ( $\Delta \bar{q}_q = \sum q_1 d_{T_0} - \sum q_0 d_{T_0} = 2,188 - 2,0 = 0,188$ ).

За счет структурных сдвигов в распределении рабочих по участкам средний уровень производительности труда повысился дополнительно на 35,4 %, или на 774 т:

$$I_{d_T} = \frac{\sum q_1 T_1}{\sum T_1} : \frac{\sum q_0 T_0}{\sum T_0} = \frac{3702,5}{1250} : \frac{2188}{1000} = 2,962 : 2,188 = 1,354;$$

$$\Delta \bar{q}_{d_T} = \bar{q}_1 - \bar{q}_0 = 2,962 - 2,188 = 0,774 \text{ тыс. т.}$$

Мы использовали 1-й вариант взвешивания факторных индексов при цепном методе анализа (формула 15.32). При 2-м варианте индекс фиксированного состава строится с отчетными весами (формула 15.33):

$$\bar{I}_q \text{ фикс. сост.} = \frac{\sum q_1 T_1}{\sum q_0 T_1} = \frac{3702,5}{0,6 \cdot 500 + 4,1 \cdot 750} = \frac{3702,5}{3375,0} = 1,097$$

$$\text{или } \bar{I}_q \text{ фикс. сост.} = \frac{\sum q_1 d_{T_1}}{\sum q_0 d_{T_1}} = \frac{2,962}{0,6 \cdot 0,4 + 4,1 \cdot 0,6} = \frac{2,962}{2,7} = 1,097$$



Следовательно, применительно к распределению рабочих по участкам в отчетном периоде средний уровень производительности труда за счет ее роста на отдельных участках повысился на 9,7 %, или на 262 т (2,962—2,700).

Индекс влияния структурных сдвигов найдем так:

$$I_{d\tau} = I_{d\text{ перем. сост.}} : I_{d\text{ фикс. сост.}} = 1,481 : 1,097 = 1,350,$$

т. е. за счет структурных сдвигов применительно к базисным уровням производительности труда на отдельных участках средний ее уровень повысился на 35 %, или на 700 т:

$$\Delta Q_{d\tau} = \sum q_0 d_{\tau 1} - \sum q_0 d_{\tau 0} = 2,7 - 2,0 = 0,7 \text{ тыс. т.}$$

Факторные индексы, полученные при разных системах взвешивания, оказались близки друг к другу, однако абсолютные приросты среднего уровня производительности труда за счет отдельных факторов при разных вариантах взвешивания более существенно отличаются (например,  $\Delta Q_q$  в одном случае равно 188 т, а в другом—262 т, или почти в 1,4 раза больше).

Индексный анализ динамики среднего уровня позволяет дополнить и углубить разложение прироста результативного показателя по факторам. В нашем случае прирост продукции за счет увеличения числа рабочих и за счет роста среднего уровня производительности труда определяется так:

$$\Delta Q_{\Sigma T} = (\Sigma T_1 - \Sigma T_0) \bar{q}_0;$$

$$\Delta Q_{\bar{q}} = (\bar{q}_1 - \bar{q}_0) \Sigma T_1.$$

По данным табл. 15.6 получим:

$$\Delta Q_{\Sigma T} = (1250 - 1000) \cdot 2 = 500 \text{ тыс. т.};$$

$$\Delta Q_{\bar{q}} = (2,962 - 2,0) \cdot 1250 = 1202,5 \text{ тыс. т.}$$

Последнюю величину можно также разложить по факторам. При 1-м варианте цепного метода прирост продукции за счет роста производительности труда на отдельных участках ( $\Delta Q_q$ ) и за счет структурных сдвигов ( $\Delta Q_{d\tau}$ ) выразится так:

$$\Delta Q_q = (\bar{q}' - \bar{q}_0) \Sigma T_1; \quad \Delta Q_{d\tau} = (\bar{q}_1 - \bar{q}') \Sigma T_1. \quad (15.35a, б)$$

В нашем примере  $\Delta Q_q = (2,188 - 2,0) \cdot 1250 = 235 \text{ тыс. т.}$

$$\Delta Q_{d\tau} = (2,962 - 2,188) \cdot 1250 = 967,5 \text{ тыс. т.}$$

При 2-м варианте цепного метода соответствующие формулы таковы:

$$\Delta Q_q = (\bar{q}_1 - \bar{q}') \Sigma T_1; \quad \Delta Q_{d\tau} = (\bar{q}' - \bar{q}_0) \Sigma T_1. \quad (15.36a, б)$$

В нашем примере получим:  $\Delta Q_q = (2,962 - 2,7) \cdot 1250 = 327,5 \text{ тыс. т.}$ ;  $\Delta Q_{d\tau} = (2,7 - 2,0) \cdot 1250 = 875 \text{ тыс. т.}$

Общий прирост продукции за счет этих двух факторов в обоих вариантах цепного метода равен приросту продукции за счет среднего уровня производительности труда (1202,5 тыс. т). Однако прирост за счет каждого отдельного фактора в одном варианте отличается от прироста в другом варианте на 92,5 тыс. т, что объясняется разной системой взвешивания аналитических индексов.

Величина 92,5 может быть получена так:  $(\bar{q}_1 - \bar{q}' - \bar{q}' + \bar{q}_0) \Sigma T_1 = (2,962 - 2,7 - 2,188 + 2,0) \cdot 1250 = 92,5$ . Это эффект совместной вариации  $i_{\tau}$  и  $i_q$  (15,316 и текст петитом на с. 358).

В этих расчетах мы использовали средние данные (суммарные и средние показатели). Иные результаты получаются при использовании данных отдельно по шахтам и разрезам и последующем их суммировании.

Сравнивая эти результаты с полученными ранее при сводном анализе, видим, что прирост продукции за счет изменения числа рабочих при раздельном анализе (по участкам) на 875 тыс. т больше, чем при сводном анализе (1375—500=875). На такую же величину прирост продукции за счет изменения производительности труда при раздельном анализе меньше, чем прирост за счет среднего уровня производительности труда при сводном анализе (1202,5—327,5=875). Это объясняется тем, что при анализе по отдельным участкам нельзя выявить влияние структурных сдвигов — фактора, который проявляется только для всей совокупности участков производства. Влияние структурного фактора, выявляемое при 1-м варианте цепного метода (967,5), при раздельном анализе распределяется между объемным и качественным факторами, а при 2-м варианте цепного метода все выявленное влияние структурных сдвигов (875,0) относится к объемному фактору. Поэтому при 2-м варианте влияние роста производительности труда на отдельных участках совпадает с влиянием этого фактора при раздельном анализе (327,5).

Возникает вопрос: нельзя ли распределить по участкам прирост продукции за счет структурных сдвигов, т. е. определить вклад каждого участка в прирост продукции за счет структурного фактора? Применительно ко 2-му варианту цепного метода в литературе обычно рекомендуют для решения этой задачи такая формула:  $\Delta Q_{d\tau} = Q_0 (i_{\tau} - i_{\tau'})$ . Применительно к нашему примеру это дает для шахт —  $\Delta Q_{d\tau} = 360 \cdot (5/6 - 5/4) = -150 \text{ т}$ , для разрезов —  $\Delta Q_{d\tau} = 1640 \cdot (1,875 - 1,25) = 1025 \text{ т}$ . В целом это совпадает

Таблица 15.7. Разложение прироста продукции по факторам в шахтах и разрезах

Участки производства	$\Delta Q$		всего
	$\Delta Q_{\tau} =$ $=(\frac{T_1}{T_0} - 1) Q_0$	$\Delta Q_q =$ $=(\frac{q_1}{q_0} - 1) T_1$	
Шахты	-60	20,0	-40,0
Разрезы	1435	307,5	1742,5
Итого	1375	327,5	1702,5

связываем структурного фактора при 2-м варианте цепного метода (875 тыс. т — см. выше). Однако величина по отдельным участкам, на наш взгляд, трудно (или даже невозможно) экономически и статистически рационально интерпретировать. Действительно, если есть всего два участка, то удельный вес каждого из них может измениться только вместе с изменением удельного веса другого, т. е. изменение удельных весов двух участков — это единый и одновременно происходящий процесс и его результаты на обоих участках должны быть либо благоприятными (увеличение), либо неблагоприятными (снижение среднего уровня производительности труда и объема продукции). Естественно, например, ожидать, что если увеличивается удельный вес участка, где в базисном периоде<sup>1</sup> производительность труда была более высокой, то это должно привести к росту среднего уровня производительности труда и увеличению за счет этого объема продукции. Следовательно, и уменьшение удельного веса другого участка, где производительность труда была ниже, тоже должно способствовать росту среднего уровня производительности труда и объема продукции. Между тем по рассмотренной методике в шахтах за счет структурного фактора объем продукции уменьшился.

Если участка не два, а больше, то согласно этой методике оказывается, что увеличение веса какого-либо участка всегда приводит к росту среднего уровня производительности труда и объема продукции на этом участке, а уменьшение удельного веса — к снижению этих показателей, причем совершенно независимо от того, выше или ниже средний уровень производительности труда в базисном периоде на этом участке.

Нам представляется, что при любом количестве участков значительно легче поддаются интерпретации результаты, получаемые по такой формуле:  $\Delta Q_{\Sigma} = (I_T - I_T) (q_0 - q_0) T_0$ . Так, в нашем примере получим: для шахт —  $\Delta Q_{\Sigma} = (5/6 - 5/4) \cdot (0,6 - 2,0) \cdot 600 = 350$  т; для разрезов —  $\Delta Q_{\Sigma} = (1,875 - 1,25) \cdot (4,1 - 2,0) \cdot 400 = 525$  т.

Эффект роста производительности труда можно выразить и измерить либо *приростом продукции*, как это было показано выше, либо *экономией рабочего времени* (рабочей силы, труда). При анализе среднего уровня производительности труда экономии (дополнительные затраты) рабочего времени за счет изменения производительности труда можно определить по таким формулам: для отдельного участка —

$$\Delta T_q = T_1 - I_q Q_1 = T_1 (1 - i_q) = -\Delta Q_q : q_0; \quad (15.37a, 6, v)$$

в целом по всем участкам —

$$\Delta T_{\Sigma} = \Sigma T_1 - \Sigma I_q Q_1 = \Sigma T_1 (1 - i_q) = \Delta Q_{\Sigma} : \bar{q}_0. \quad (15.38a, 6, v)$$

Так, используя полученный выше результат ( $\Delta Q_{\Sigma} = 1202,5$  тыс. т), найдем, что за счет роста среднего уровня производительности труда в целом по всем участкам был сэкономлен труд более 600 человек:

$$\Delta T_{\Sigma} = -\Delta Q_{\Sigma} : \bar{q}_0 = -1202,5 : 2 = -601 \text{ чел.}$$

или

$$\Delta T_{\Sigma} = \Sigma T_1 (1 - i_q) = 1250 \cdot (1 - 1,481) = -601 \text{ чел.}$$

Из последней формулы следует: на сколько процентов повысился средний уровень производительности труда, столько же

<sup>1</sup> В базисном периоде потому, что при 2-м варианте цепного метода индекс структурных сдвигов взвешивается по базисным весам ( $q_0$ ).

процентов от общих затрат труда в отчетном периоде было за счет этого сэкономлено. Это справедливо также для тех случаев, когда рост производительности труда измеряется с помощью аналитического индекса с базисными весами ( $I_{q \text{ фикс. вес}}$  в (15.31; 15.32) и  $I_{q \text{ стр. ед}}$  в (15.33)). Кроме того, это верно и для разноименной продукции, когда используются формулы (15.186) и (15.21).

Советская статистика при построении индексов исходит из экономического содержания реальных процессов и тех конкретных познавательных задач, которые должны быть решены с помощью индексов. Каждый индекс в нашей статистике имеет экономическое обоснование и определенный экономический смысл.

Иначе подходит к построению индексов буржуазная статистика, которая рассматривает эту проблему с *формальных логических и математических позиций*. Так, американский статистик И. Фишер в работе «Построение индексов» писал, что ставит задачу «...найти такую математическую формулу, которая совершенно не зависела бы от природы того материала, к которому мы ее будем применять»<sup>1</sup>. Понятие индекса сводится к чисто «счетному модели». Считая, что сводные индексы должны иметь те же математические свойства, что и индивидуальные, Фишер «скрещивает» два разновзвешенных агрегатных индекса, вычисляя из них среднюю геометрическую:

$$I_p = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \cdot \frac{\Sigma p_1 Q_1}{\Sigma p_0 Q_1}}.$$

Эту формулу, лишенную ясного и конкретного экономического смысла, Фишер считает «идеальной», так как только она удовлетворяет ряду формальных логических и математических требований. Лишены экономического обоснования и многие другие индексные построения буржуазных теоретиков.

Серьезными недостатками страдает и практика исчисления индексов в странах капитала. Главный из них — отсутствие полных и точных данных, необходимых для расчета индексов. Это вынуждает буржуазных статистиков прибегать к разного рода условностям: вместо полного перечня продуктов или товаров брать лишь набор некоторых их представителей, вместо прямых данных использовать различные косвенные и суррогатные, а также многократно корректировать и исправлять как первичные данные, так и индексы. В результате последние часто дают лишь весьма приблизительную и ориентировочную оценку реальной динамики. В то же время многочисленные условности при расчете открывают широкий простор для тенденциозного искажения действительного положения вещей в тех случаях, когда это связано с классовыми интересами монополистической буржуазии.

<sup>1</sup> Фишер И. Построение индексов. М., 1928, с. 181.

## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Абсолютное значение 1% прироста 273  
 Абсолютные величины 114  
 — индивидуальные 114  
 — суммарные 115  
 Автокорреляция 307  
 Автоматизированный банк данных (АБД) 31  
 Автоматизированная система государственной статистики (АСТС) 30  
 Вариация 7, 151  
 Вексовая зависимость населения 197 г. 38, 44, 192  
 Выборка 169—194  
 — взаимопроницаемая 189  
 — комбинированная 187  
 — малая 189  
 — механическая 185  
 — многоступенчатая 187  
 — многофазная 188  
 — определение объема 177  
 — принципы отбора 184  
 — репрезентативность 179  
 — серийная (гнездовая) 186  
 — собственно-случайная 184  
 — типическая 184, 186  
 Выборочное обследование бюджетов семей рабочих, служащих и колхозников 191  
 — социологическое 194  
 Городская (районная) информационно-вычислительная станция (ГР) ИВС 29  
 Графики 107—114  
 — гистограммы 111  
 — диаграммы 109  
 — изображение ряда динамики 273—275  
 — тенденции 290  
 — картограммы и картодиаграммы 109, 130  
 — кривая концентрации 113  
 —кумулятивная 112  
 — нормального распределения 162  
 — распределения 161  
 — поле корреляции 222  
 — полигон распределения 111  
 — разложение абсолютного прироста по факторам 347  
 — уравнивание регрессии 219  
 — эмпирическая линия регрессии 207  
 Группировка 74—94  
 — аналитическая 87, 206  
 — в буржуазной статистике 11—12, 76  
 — вторичная 89  
 — интервалы 83  
 — замкнутые 87  
 — открытые 87

— равные и неравные 86  
 — комбинационная 79  
 — определенная 74  
 — переизбыток 89  
 — простая 78  
 — структурная 85  
 — типологическая 80  
 Группировки и РКА 252—259  
 Графоаналитический способ 256  
 Дисперсия 152  
 — альтернативного признака 155  
 — внутригрупповая 158  
 — выборочной совокупности 171  
 — межгрупповая 157  
 — общая 157  
 — остаточная 233  
 — правило сложения дисперсий 157  
 — средняя из групповых 158  
 — факторная 232  
 Достоверность данных наблюдения 51  
 — контроль 57  
 — виды 57, 58  
 — методы проверки 55  
 Единица наблюдения 42  
 — совокупности 42  
 — отчетная 43  
 Единицы измерения 115  
 — натуральные 115  
 — стоимостные 116  
 Закон больших чисел 17, 132  
 Индексы 321—364  
 — агрегатные 329—336  
 — взаимосвязь индексов 324, 332, 354  
 — влияния ассортиментных сдвигов 332  
 — влияния структурных сдвигов 355, 357  
 — выбор весов 327, 335  
 — выполнения плана 324, 332  
 — детерминации при аналитической группировке 209  
 — при регрессионно-корреляционном анализе 246  
 — динамики 324, 332  
 — индивидуальные 323  
 — ковариации 351  
 — корреляции 246  
 — совокупные 246  
 — частные 250  
 — переменного состава 355  
 — сводные 323  
 — качественных показателей 329  
 — объемных показателей 335  
 — планового задания 324, 332  
 — производительности труда 333—334  
 — для односторонней продукции 333, 354—357  
 — для разносторонней продукции 334

ряды индексов с переменными весами 336  
 — с постоянными весами 336  
 — сезонности 302  
 — себестоимости единицы продукции (сводные) 330  
 — символика (условные обозначения) 322—323  
 — средние из индивидуальных (групповых) индексов 340—342  
 — арифметические 340  
 — гармонические 340  
 — трудоемкости единицы продукции (сводные) 331  
 — условных расходов сырья, топлива (сводные) 331  
 — фиксированного (постоянного) состава 355  
 — экономический эффект от изменения качественных показателей 329  
 Классификация (в статистике) 82  
 Корреляционная матрица 240  
 — связь 197  
 Корреляционное отношение 209  
 Корреляция логична 236  
 — множественная 199, 239  
 — парная 199  
 Коэффициент автокорреляции 308  
 — асимметрии 167  
 — вариации 154  
 — взаимной сопряженности 212  
 — детерминации 236  
 — доверия 175, 238  
 Коэффициент корреляции 236  
 — между отклонениями от трендов 313  
 — рангов 202  
 — частный 250  
 — опережения 305  
 — ускорения (замедления) роста 289  
 — Фехнера 202  
 — эксцесса 168  
 — эластичности 237, 248  
 Критерий Дарбина—Уотсона 311  
 — согласия Колмогорова 165  
 — Пирсона 164  
 — Романовского 165  
 — t-Стьюдента 302, 180, 237, 250  
 — Фишера 210, 217, 238, 301  
 Лаг временной 315  
 Левин В. И. о статистике 9, 11, 43, 67, 73, 77, 108, 148  
 Математическая статистика 17  
 Медиана 144  
 Метод наименьших квадратов 323, 241, 297  
 — сравнения параллельных рядов 199  
 Мода 147

Моделирование дискретно-непрерывное 258  
 Моменты (распределения) 137, 167—169  
 Мультиколлинеарность 240  
 Наблюдение статистическое 36  
 — виды 38  
 — время 47  
 — выборочное 39, 169  
 — объект 42  
 — критический момент 47  
 — программа 43  
 — способы проведения 41  
 — сплошное 38  
 — формуляр 45  
 — формы 48  
 Нормальное распределение 162  
 Обследование специальное статистическое 50  
 Общесоюзная автоматизированная система сбора и обработки информации для учета, планирования и управления народным хозяйством (ОГАС) 30  
 Общесоюзные классификаторы 95  
 Относительные величины 117  
 — выполнения плана 119  
 — динамики (темпы роста) 118  
 — интенсивности 124  
 — координации 123  
 — планового задания 119  
 — сравнения 125  
 — структуры 121  
 Отчетность статистическая 48  
 — виды 49  
 — контроль достоверности 58  
 — проверка качества 55  
 Ошибка выборки (репрезентативности) 172  
 — виды 172  
 — относительная 176  
 — предельная 174  
 — средняя (стандартная) 174  
 Ошибки 172  
 — при выборке взаимопроницаемой 190  
 — комбинированной 187  
 — малой 180  
 — механической 186  
 — многоступенчатой 188  
 — многофазной 189  
 — серийной 187  
 — собственно-случайной 184  
 — типической 186  
 Партийность статистики 13  
 Переписной лист 45  
 Плотность распределения 93, 112  
 Показатели статистические 68—72  
 — вариации 150—158  
 — вторичные 69  
 — индивидуальные 70



— качественные (интенсивные) 322	тренд 291
— объемные (экстенсивные) 322	укрупнение интервалов 292
— отчетные 68	экстраполяция 316
— зрительные 69	Сводка
— плазовые 68	— машинизированная 94
— статистические 69	— программа 73
— сводные 70	— ручная 94
Признак статистический 8, 17, 42, 43, 63	Связь криволинейная 198
— альтернативный 155	— корреляционная 197
— атрибутивный 64	— обратная 198
— варьирующий 63	— прямая 198
— группировочный 75	— прямолинейная 198
— количественный 64	— стохастическая 197
— качественный 63	— функциональная 196
— результативный 196, 217	Совокупность выборочная 171
— существенный 63	— генеральная 171
— факторный 196, 217	— статистическая 7—8, 61
Прирост абсолютный 269	Способ моментов 137, 155
— базисный 269	Среднее квадратическое отклонение 152
— средний 282, 284	— линейное отклонение 151
— цепной 269	Средняя величина 132—133
Прогнозирование 263, 316	— арифметическая 133
Промилле 118	— взвешенная 134
Регрессионно-корреляционный анализ (РКА) 214—264	— гармоническая 139
— построение модели 217	— геометрическая 143, 285
— применение 259—264	— определяющий показатель 141, 284
— требования к информации 216—217	— простая (невзвешенная) 134
— этапы 215	— степенная 143
— шаговый анализ 251	— хронологическая 281
Регрессия множественная (многофакторная модель) 239	Таблица статистическая 99
— парная (однофакторная модель) 218	— групповая 101
Ряд вариационный 92	— комбинационная 102
— распределения 92	макет 100
— варианты 92	оформление 105
Ряды динамики (временные) 265	подлежащие 100
— анализ ряда 268	— простая 101
— нескольких рядов 304—307	сказуемое 100
— автокорреляция уровней 307	Темп прироста 272
— остатков 311	— — средний 285
— авторегрессия 307	Темп роста 118, 272
— виды 267	— — базисный 271
— интервальные 266	— — средний 285, 287, 292
— моментные 266	— — цепной 271
— выбор формы тренда 295, 300	Удельный вес 121
— выравнивание (аналитическое) 295	Уравнение регрессии 214
— по параболе 300	Уровень ряда динамики 265
— по показательной кривой 299, 321	— — интервального 277
— по прямой 297	— — среднего 277
интерполяция 319	— — моментного 278—281
корреляция 307—316	— — средних величин 281
регрессия 307—316	— — относительных величин 281
сглаживание с помощью скользящей средней 293	Учет 21
случайная компонента уровня (остаток) 295—298, 310	— бухгалтерский 25
смысл рядов 273	— оперативно-технический 25
	— статистический 26
	Центральное статистическое управление (ЦСУ СССР) 27
	Частоты 92, 134, 143

## ОГЛАВЛЕНИЕ

Предисловие	5
<b>РАЗДЕЛ I</b> Статистика как наука. Организация статистики	7
ГЛАВА 1 Статистика как наука. Предмет и метод статистики	7
ГЛАВА 2 Организация статистики	21
<b>РАЗДЕЛ II</b> Сбор статистической информации и обеспечение качества ее материалов	36
ГЛАВА 3 Статистическое наблюдение	36
ГЛАВА 4 Обеспечение качества материалов статистического наблюдения	51
<b>РАЗДЕЛ III</b> Статистическая совокупность и ее характеристика	61
ГЛАВА 5 Совокупность как объект статистического изучения	61
ГЛАВА 6 Сводка и группировка	72
ГЛАВА 7 Способы наглядного представления статистических данных	99
ГЛАВА 8 Абсолютные и относительные величины	114
ГЛАВА 9 Средние величины	132
ГЛАВА 10 Показатели вариации и анализ вариационных рядов	150
ГЛАВА 11 Выборочное наблюдение	169
<b>РАЗДЕЛ IV</b> Статистическое изучение взаимосвязей	195
ГЛАВА 12 Взаимосвязи социально-экономических явлений и методы их статистического изучения	195
ГЛАВА 13 Анализ регрессии и корреляций	214
<b>РАЗДЕЛ V</b> Анализ динамики	265
ГЛАВА 14 Анализ рядов динамики	265
ГЛАВА 15 Индексный анализ	321
Предметный указатель	364

Арон Яковлевич Боярский,  
Лидия Леонтьевна Викторова,  
Абрам Монсеевич Гольдберг и др.

#### ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Зав. редакцией Р. А. Казьмина  
Редактор Л. В. Сергеева  
Мл. редакторы В. Г. Крылова, А. В. Щурова  
Техн. редактор Л. Г. Челышева  
Корректоры Г. В. Хлопцева, Г. А. Башарина,  
Т. М. Иванова  
Худож. редактор М. К. Гуров  
Переплет художника М. Л. Мучниковой

ИБ № 1637

Сдано в набор 04.06.85. Подписано в печать 15.11.85. А15318.  
Формат 60x90<sup>1/4</sup>. Бумага кал.-журн. Гарнитура «Литературная».  
Печать высокая. Усл. п. л. 23,0. Усл. кр.-отт. 23,0. Уч.-изд. л.  
24,86. Тираж 20 000 экз. Заказ 1096. Цена 1 р. 10 к.

Издательство «Финансы и статистика», 101000, Москва,  
ул. Чернышевского, 7

Ленинградская типография № 4 ордена Трудового Красного  
Знамени Ленинградского объединения «Техническая книга»  
им. Евгении Соколовой Союзполиграфпрома при Государствен-  
ном комитете СССР по делам издательств, полиграфии и книж-  
ной торговли. 191126, Ленинград, Социалистическая ул., 14.