КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ УПРАВЛЕНИЯ, ЭКОНОМИКИ И ФИНАНСОВ
Кафедра экономико-математического моделирования

Э.А. ПОЛОВКИНА, Е.А. ГРИГОРЬЕВА

СТАТИСТИКА
РАЗДЕЛ «ТЕОРИЯ СТАТИСТИКИ»

Учебное пособие

Казань – 2016

УДК 311
ББК С60

Принято на заседании кафедры экономико-математического моделирования
Протокол № 3 от 09 ноября 2016 года

Рецензенты:
кандидат экономических наук,
доцент кафедры экономико-математического моделирования КФУ
Е.И. Кадочникова

Половкина Э.А., Григорьева Е.А.
Статистика. Раздел «Теория статистики»: учебное пособие / Э.А. Половкина, Е.А. Григорьева. – Казань: Казан. ун-т, 2016. – 244 с.

Настоящее учебное пособие по дисциплине «Статистика» составлено по разделу «Теория статистики» в соответствии с Государственным образовательным стандартом, предусмотренным для экономических высших учебных заведений Российской Федерации. Пособие охватывает основной практический материал теории статистики и ориентирован на познание статистических методов исследования важнейших статистических показателей, используемых в учете, отчетности и анализе.

Цель данного пособия видится в оказании помощи студентам и всем интересующимся статистикой в исследовании актуальных социально-экономических проблем и практическом использовании при этом системы статистических показателей и методов анализа.

ПРЕДИСЛОВИЕ

В современном обществе эффективность управления экономикой в значительной степени зависит от качества информационной базы, основным компонентом которой является статистика. Статистическая грамотность – неотъемлемая часть экономического образования. Работая с цифрами, каждый экономист должен знать, как получены те или иные данные, какова их природа, насколько полны они и достоверны.

Статистика осуществляет наблюдение, сбор, научную обработку, обобщение и анализ информации обо всех явлениях и процессах общественной жизни. Поэтому в системе экономического образования особое место отводится изучению статистики, являющейся базовой научной дисциплиной, формирующей профессиональный уровень современного экономиста. В учебных программах вузов статистика объединяет комплекс учебных дисциплин, в составе которого следует выделить такие основные разделы, как:

1) «Общая теория статистики»;

2) «Социально-экономическая статистика»;

3) «Микроэкономическая статистика»;

4) «Многомерный статистический анализ»;

5) «Система национального счетоводства.

Каждый из данных разделов имеет свои конкретные цели и задачи.

«Общая теория статистики» призвана дать студентам представление о содержании статистики как научной дисциплины, познакомить с ее основными понятиями, методологией и методиками расчета важнейших статистических аналитических показателей. Ее задача – научить студентов решать проблему информационного обеспечения процесса принятия управленческих решений, начиная с подготовки информации, ее анализа и заканчивая количественной и качественной оценкой происшедших изменений в анализируемых процессах. В соответствии с этим данное учебное пособие охватывает самые общие начальные элементы статистической науки и рассматривает основные этапы статистического исследования (статистическое наблюдение, сводка, группировка, расчет обобщающих показателей – абсолютных, средних и относительных), методы изучения вариации и динамики, индексный метод анализа, основы регрессионного и корреляционного анализа, непараметрические методы анализа.

Статистические методы исследования находят широкое применение во многих отраслях знаний и человеческой деятельности. Региональная экономика, сравнительно молодая наука, в своем развитии и становлении изучает закономерности и особенности саморегулирования и государственного регулирования экономики в региональном разрезе с помощью статистики. В региональных исследованиях на основе данных, характеризующих результаты деятельности объектов региональной экономики, осуществляется оценка социально-экономического положения в регионах, изучается рыночная экономическая конъюнктура, инвестиционная привлекательность объектов, измеряется уровень территориальной дифференциации по основным показателям региональной экономики и т.п. Чем полнее и достовернее аналитическая информация, отражающая реальные процессы и явления в регионах, тем эффективнее осуществляется принятие управленческих решений.

В одном из журналов «Вопросы статистики» опубликована статья «Как я стал статистиком и почему не сожалею об этом» ведущего ученого – статистика, доктора экономических наук Санкт-Петербургского государственного аграрного университета М.М. Юзбашева, который пишет «… изучая статистику, тем более, участвуя в ее развитии, невозможно не гордиться тем, что соприкасаешься с такими истинами, которые едины и вечны на все времена и на всех галактиках! В самом деле: средняя величина – такая же мировая константа, как сумма, произведение, дифференциал!».

В статье есть стихотворение, посвященное статистике, шуточное по виду и серьезное по содержанию, которое М.М. Юзбашев читает своим студентам, заканчивая лекционный курс. Мы посчитали возможным привести некоторые строфы этого стихотворения, так как в нем хорошо показана познавательная значимость статистики.

«Статистика всем специалистам нужна:
В анализе ведь – королева она!
В пустые слова она учит не верить,
А все, что возможно, учесть и измерить!
Познать пожелаешь ли тип иль закон –
То в средней, как правило, выражен он:
Еще посчитай медиану и моду,
Лишь после вещай свои мысли народу!
Спеши динамический выровнять ряд –
Узнаешь, идешь ли вперед иль назад;
И чтоб болтуны не водили за нос,
Измерив тенденцию, сделай прогноз!
Запомни же твердо, себя зря не мучая:
Наш мир есть гибрид от Закона и Случая!
В грядущем у нас достоверностей нет;
Будь мужествен, знай: стохастичен весь свет!
Кто понял статистику, тот не внакладе,
Страшиться не нужно ему ничего:
Запас страховой и в душе и на складе
Спасет от случайности всякой его!»

ГЛАВА 1. ОСНОВНЫЕ ПОНЯТИЯ И КАТЕГОРИИ СТАТИСТИЧЕСКОЙ НАУКИ. ОРГАНИЗАЦИЯ СТАТИСТИКИ В РОССИЙСКОЙ ФЕДЕРАЦИИ

1.1. Предмет и основные категории статистики как науки

Слово *«статистика»* имеет латинское происхождение (от *status,* что означает «определенное положение вещей» – состояние). В средние века оно использовалось для характеристики политического состояния государства и употреблялось в значении слова «государствоведение». В науку этот термин введен в XVIII веке немецким ученым Готфридом Ахенвалем, выступившим книгу о государствоведении.

Собственно как наука статистика возникла только в XVII веке, когда правительства различных западноевропейских стран стали заниматься сбором разного рода информации о своих гражданах. Однако статистический учет существовал уже в глубокой древности, упоминания о статистических обследованиях встречаются и в библейские времена. «Исчислите все общество сынов Израилевых по родам их, по семействам их, по числу имен, всех мужского пола поголовно…» Четвертая книга Моисеева. Числа. Ветхий завет, гл. 1, 2.

Развитие статистики схоже с развитием языка, она постепенно развивалась там, где в ней возникала необходимость. Еще за 5 тыс. лет до н.э. проводились переписи населения в Китае, велся учет имущества граждан в Древнем Риме, использование средней было хорошо известно еще при жизни Пифагора. В средние века осуществлялось сравнение военного потенциала разных стран, численности их населения, домашнего имущества, земель.

У истоков статистической науки стояли две школы – немецкая описательная и английская школа политических арифметиков.

Представители описательной школы считали, что задачей статистики является описание достопримечательностей государства; территории, населения, климата, вероисповедания, ведения хозяйства и т.п. – только в словесной форме, без цифр и вне динамики, то есть без отражения особенностей развития государств в те или иные периоды, а только лишь на момент наблюдения. Видными представителями описательной школы были Г. Конринг (1606-1661), Г. Ахенваль (1719-1772), А. Бюшинг (1724-1793) и др.

Политические арифметики ставили целью изучать общественные явления с помощью числовых характеристик – меры веса и числа. Это был принципиально новый этап развития статистической науки по сравнению со школой государствоведения, так как от описания явлений и процессов статистика перешла к их измерению, исследованию, оценке и выработке вероятных гипотез будущего развития. Политические арифметики видели основное назначение статистики в изучении массовых общественных явлений, осознавали необходимость учета в статистическом исследовании требований закона больших чисел, поскольку закономерность может проявиться лишь при достаточно большом объеме анализируемой совокупности. Виднейшим представителем и основателем этого направления был В. Петти (1623-1687). Именно школа политических арифметиков стала основообразующей в развитии современной статистики.

В XIX веке получило развитие учение бельгийского статистика Адольфа Кетле (1796-1874), который первым применил современные методы сбора данных, его считают основоположником учения о средних величинах. Математическое направление в статистике развивалось в работах англичан – сэра ФрэнсисаГальтона (1822-1911) и Карла Пирсона (1857-1936), Рональда Фишера, которые внесли значительный вклад в развитие теории корреляции и оказали существенное воздействие на современную статистику.

Прогрессу статистической методологии способствовали труды российских статистиков – А.А. Чупрова (1874-1926), В.С. Немчинова (1894-1964), С.Г. Струмилина (1877-1974), В.Н. Старовского (1905-1975) и др.

Развитие статистической науки, расширение сферы практической статистической работы привели к изменению содержания самого понятия «статистика».

Статистическая наука – это отрасль знаний, изучающая явления общественной жизни с их количественной стороны в неразрывной связи с их качественным содержанием в конкретных условиях места и времени. Статистика – это особая научная дисциплина, которая в широком понимании разрабатывает методы сбора, систематизации, анализа, интерпретации и отображения результатов наблюдений массовых случайных явлений и процессов целью выявления существующих в них закономерностей.

Статистическая практика – это деятельность по сбору, накоплению, обработке и анализу цифрового материала, служащего для характеристики какой-либо области общественных явлений или территориального распределения какого-то показателя, публикуемые в периодической прессе, справочниках, сборниках. Статистика как отрасль практической деятельности имеет своей целью сбор, обработку, анализ и публикацию массовых данных о самых различных явлениях общественной жизни. Осуществляется сбор данных в каждом регионе и по стране в целом о численности и составе населения, ведется подсчет предприятий и организаций, собираются данные об объемах производства и объемах продаж и т.д. Эту деятельность на профессиональном уровне осуществляет Федеральная служба государственной статистики (Росстат) – и система ее учреждений, организованных, по административно-территориальному признаку. Так в Татарстане действует Территориальный орган Федеральной службы государственной статистике (Татарстанстат).

Статистические законы действуют в пределах времени и места, в которых они были обнаружены.

Если рассматривать статистику как инструмент изучения социально-экономических явлений и процессов, то *предмет статистики состоит в изучении размеров и количественных соотношений массовых общественных явлений в конкретных условиях места и времени,* а так же числовое выражение проявляющихся в них закономерностей.

Свой предмет статистика изучает при помощи определенных *категорий,* то есть понятий, которые отражают наиболее общие и существенные свойства, признаки, связи и отношения предметов и явлений объективного мира.

Что мы имеем в виду под закономерностью?

Закономерность, выявленная на основе массового наблюдения, то есть проявляющаяся лишь в большой массе явлений через преодоление свойственной ее единичным элементам случайности, называется статистической закономерностью.

Свойство статистических закономерностей проявляться лишь *в массе* явлений при обобщении данных по достаточно большому числу единиц, находит свое отражение в законе больших чисел, сущность которого состоит в том, что по мере увеличения числа наблюдений влияние случайных факторов взаимопогашается и на поверхность выступает действие основных факторов, которые и определяют закономерность. Например, характеристика экологической ситуации предполагает изучение закономерности динамики выбросов загрязняющих веществ в атмосферный воздух регионов от динамики физического объема валового регионального продукта.

Познание закономерностей возможно только в том случае, если изучаются не отдельные явления, а совокупности явлений. То есть объектом статистического изучения является статистическая совокупность – множество единиц изучаемого явления, объединенных качественной однородностью, определенной целостностью, взаимозависимостью состояний отдельных единиц и наличием вариации. Таковы, например, совокупность домохозяйств, совокупность предприятий и фирм, совокупность нефтяных месторождений, совокупность регионов и т.п.

Совокупность называется *однородной,* если один или несколько изучаемых существенных признаков являются *общими для всех единиц*. Например, принадлежность предприятий к одной и той же отрасли – заводы металлургического комплекса или регионы, относящиеся к одной природно-климатической зоне.

Совокупность, в которую входят явления разного типа, считается разнородной. Совокупность может быть однородна в одном отношении и разнородна в другом. Регионы, включенные в одну группу по природно-климатическим характеристикам, различаются по уровню социально-экономического развития. Заводы, входящие в металлургический комплекс России, различаются по своей специализации – выделяются группы заводов по производству труб, или по производству листового проката и т.п. В каждом отдельном случае однородность совокупности устанавливается путем проведения качественного анализа, выяснения содержания изучаемого общественного явления.

Статистическая совокупность состоит из единиц совокупности. *Единицы статистической* совокупности представляют собой качественно *однородные первичные элементы* этой совокупности. Каждая единица совокупности представляет собой частный случай проявления изучаемой закономерности.

Решение вопроса о единице и границах изучаемой совокупности определяется целью исследования. Это связано со сложной природой социально-экономических явлений. В каждом отдельном явлении одновременно реализуются различные процессы. Например, при изучении совокупности работников, каждый работник может рассматриваться как член определенной социально-профессиональной группы, как работник предприятия, как житель города, поселка и т.д., то есть единица совокупности – это предел дробления объекта исследования, при котором сохраняются все свойства изучаемого процесса.

Единицы совокупности обладают определенными свойствами, качествами, которые принято называть признаками. *Признак – качественная особенность единицы совокупности.* Например, признаки человека: возраст, пол, образование, вес, семейное положение и т.д. Признаки предприятия: форма собственности, отрасль, численность работников, величина уставного фонда и т.д. Статистика изучает явления через их признаки: чем более однородна совокупность, тем больше общих признаков имеют ее единицы, тем меньше варьируют ее значения.

По характеру отображения свойств единиц изучаемой совокупности признаки делятся на:

- *признаки,* имеющие непосредственное *количественное* выражение, например, площадь территории, численность жителей города и т.д. Они могут быть дискретно или непрерывно варьируемыми. Дискретно варьируемые признаки – это признаки, отдельные значения которых отличаются друг от друга на некоторую конечную величину (обычно целое число). Так, дискретные признаки мы используем, когда проводится группировка, например, магазинов по числу в них отделов или касс. В магазинах может быть один, два, три и т.д. отдела, но не может быть полтора или два с половиной отдела. Существует множество признаков, значения которых отличаются друг от друга на сколько угодно малую величину и могут принимать любые значения на некотором интервале. Такие признаки называют непрерывно варьирующими или непрерывными признаками. К ним относятся индексы экономического состояния, среднедушевые доходы, весовые и объемные характеристики товаров;

- *признаки, не имеющие* непосредственного *количественного выражения*. В этом случае отдельные единицы совокупности различаются своим содержанием, например, отраслевая специализация предприятий и организаций; деление природных ресурсов по их происхождению: минеральные, водные, земельные или деление населения по полу – мужчины и женщины и т.д. Такие признаки обычно называют *атрибутивными* (в философии «атрибут» – неотъемлемое свойство предмета). В случае, когда имеются противоположные по значению варианты признака, говорят об альтернативном признаке (да, нет). Например, продукция может быть годной или бракованной (не годной); каждое лицо может состоять в браке или нет и т.д.

Особенностью статистического исследования является то, что в нем изучаются только *варьирующие* признаки, то есть признаки, принимающие различные значения (для атрибутивных, альтернативных признаков) или имеющие различные количественные уровни у отдельных единиц совокупности.

Поскольку статистика, как уже сказано, изучает количественную сторону массовых явлений, то возникает необходимость в обобщающих характеристиках статистической совокупности. Эту роль выполняет статистический показатель, являющийся количественной характеристикой какого-то свойства совокупности.

Статистический показатель – это количественная оценка свойства изучаемого явления. Статистические показатели можно подразделить на два основных вида. Первый вид – это учетно-оценочные показатели, которые показывают размеры, объемы, уровни изучаемого явления. Второй вид показателей – аналитические, которые показывают, как развивается изучаемое явление, из каких частей состоит целое, то есть в каком соотношении находятся части целого между собой и как распространяется явление в пространстве. К аналитическим относят относительные и абсолютные величины, показатели вариации и т. д.

1.2. Методы статистики

Свой предмет статистика изучает при помощи своих, специфических методов. Общей основой разработки и применения статистической методики является диалектический метод познания, согласно которому общественные явления и процессы рассматриваются в развитии.

В соответствии с этим статистика изучает явления:

- в их взаимосвязи;
- в движении и изменении;
- выделяя их в типы и формы;
- в соответствии с соблюдением принципа места и времени.

В процессе развития, наряду с количественными изменениями, в изучаемом предмете происходят коренные качественные изменения. Поэтому необходимо располагать методами, позволяющими изучать количественные изменения в явлениях, оценивать существенность или несущественность наблюдаемых различий, улавливать переход количественных изменений в качественные.

Любое статистическое исследование проходит три стадии: сбор информации, обработка собранной информации, анализ результатов обработки.

Методы статистики – это целая совокупность приемов, пользуясь которыми статистика исследует свой предмет.

На первой стадии исследования статистика использует *метод статистического наблюдения*, который заключается в сборе первичного статистического материала, в научно-организованной регистрации всех существенных факторов, относящихся к рассматриваемому объекту. Требования массовости единиц наблюдения обуславливается тем, что изучаемые статистической закономерности проявляются в достаточно большом массиве данных на основе действия закона больших чисел.

На второй стадии исследования применяется *метод группировок*, который дает возможность все собранные в результате массового статистического наблюдения факты подвергать систематизации и классификации.

На третьей стадии исследования статистика использует ряд методов. *Метод обобщающих показателей* позволяет характеризовать изучаемые явления и процессы при помощи статистических величин – абсолютных, относительных и средних. На этом этапе статистического исследования выявляются *взаимосвязи* и масштабы явлений, изучается *степень вариации* показателей, определяются *закономерности* их развития, даются прогнозные оценки. *Индексный метод* позволяет изучить и количественно измерить факторы, обусловившие изменение того или иного явления и процесса. *Балансовый метод* дает представление о сложившихся пропорциях в различных областях экономики.

К какой бы области не относился предмет статистики (территория, население, промышленность, финансы) методы ее везде применимы. Статистика выявляет типичное и закономерное. Иначе говоря, методы статистики обусловлены спецификой ее предмета.

Поскольку количественная сторона общественных явлений в их неразрывной связи с качественной стороной изучается другими общественными науками, то статистика теснейшим образом связана с ними. Для измерения количественной стороны явлений, разработки соответствующих показателей необходимо знать особенности, отличающие одни явления от других, свойственные им специфические закономерности развития, то есть, говоря проще, статистика должна владеть предметной областью, для которой разрабатываются показатели. Например, для разработки показателей, характеризующих размещение производительных сил необходимо знание экономической географии, для финансовых показателей – финансов и т.д.

Среди общественных наук, определяющей по отношению к другим, является экономическая теория, раскрывающая механизмы законов, регулирующих изменения развития общества. Статистика при определении количественной стороны массовых общественных явлений базируется на экономической теории. В зависимости от общественной формации, политического строя значительно меняется содержание показателей статистики. Так до 60-х годов прошедшего столетия в отечественной статистике не существовал показатель прибыли, поскольку его не было в политической экономии социализма того времени.

Статистика представляет собой чрезвычайно мощное орудие познания, но подобно всякому инструменту, она может быть использована как для увеличения возможностей повышения жизненного уровня людей, так и для искажения реальной действительности в результате злоупотреблений. Наверное, это и дало основание английскому политику Б. Дизраэли (1804-1881) заявить: «Есть ложь, есть наглая ложь, а есть статистика».

Связь между статистикой и другими общественными науками не односторонняя. Статистика – важнейшее звено обобщения результатов массовой практики. Никакие вопросы, касающиеся экономического строя современного государства и его развития, не могут быть решены только путем теоретических умозаключений, они, безусловно, должны базироваться на проверенных и научно-обоснованных фактах. Не зря тот же Б. Дизраэли сказал: «В жизни, как правило, преуспевает тот, кто располагает лучшей информацией»!

Закономерности массовых явлений социально-экономической жизни нередко обнаруживаются раньше статистикой, а потом уже становятся объектом изучения конкретной науки. Статистика нередко выявляет такие закономерности общественной жизни, которые на первый взгляд противоречат логике вещей, принятой теоретической концепции и поэтому становятся предметом теоретического исследования, после того как они были обнаружены эмпирически.

Статистика имеет большое познавательное значение, она содержательно и количественно освещает изучаемые явления и процессы, служит надежным способом оценки проверить различные теоретические предположения, доказать или опровергнуть какие-либо расхожие утверждения.

1.3. Современная организация государственной статистики Российской Федерации и ее основные задачи

Федеральная служба государственной статистики (Росстат) является уполномоченным федеральным органом исполнительной власти, осуществляющим функции по формированию официальной статистической информации о социальном, экономическом, демографическом и экологическом положении страны. В своей деятельности Федеральная служба статистики руководствуется Конституцией Российской Федерации, федеральными законами и другими нормативными актами.

Федеральная служба государственной статистики Российской Федерации, ее органы в республиках, краях, областях, автономных областях и округах, в городах Москве и Санкт-Петербурге, других городах и районах, а также подведомственные ей организации составляют единую систему государственной статистики страны.

Формы и методы сбора и обработки статистических данных, методология расчета статистических показателей, установленные Росстатом, являются статистическими стандартами Российской Федерации.

В соответствии с положением основными задачами Федеральной службы госстатистики России являются:

- предоставление официальной статистической информации Президенту, Правительству, федеральному собранию Российской Федерации, федеральным органам исполнительной власти, общественности;

- разработка научно-обоснованной статистической методологии, соответствующей международным стандартам;

- координация статистической деятельности в государстве;

- разработка экономико-статистической информации, ее анализ, составление национальных счетов, проведение необходимых балансовых расчетов.

Основные функции Федеральной службы госстатистики России состоят в том, что она:

- организует проведение государственных статистических наблюдений по разработанным им или согласованным с ним программам, формам и методикам;

- обеспечивает функционирование ЕГРПО (Единого государственного регистра предприятий и организаций);

- обеспечивает сбор, обработку, хранение и защиту статистической информации, соблюдение государственной и коммерческой тайны, необходимую конфиденциальность данных (конфиденциальный – секретный, доверительный);

- сопоставляет основные социально-экономические показатели России с аналогичными показателями других стран, совместно с Центробанком составляет платежный баланс страны;

- проводит единую техническую политику в области сбора, обработки и передачи статистической информации, в разработке и формировании федеральных программ по вопросам, порученным Федеральной службе.

В современных условиях необходим новый подход к реформам в статистике. Одно из основных направлений – разработка методологии и организации получения информации о теневой экономике. Очень важное значение в настоящее время приобретает разворачивание системы мониторингов – специально организованных систематических наблюдений, особенно организованных систематических наблюдений, особенно в социальной сфере (изучение уровня жизни населения). Также важна программа компьютеризации статистики на основе создания информационно-телекоммуникационной системы, которая строится на основе вводимой в эксплуатацию информационно-вычислительной сети, включающей локальные вычислительные сети во всех органах государственной статистики федерального и регионального уровней.

Интеграция России в международную систему учета обуславливает необходимость внедрения Системы национального счетоводства. В этой связи важно развитие профессиональных контактов с международными статистическими службами ООН.

Статистическая комиссия ООН осуществляет разработку методологии статистических работ, сопоставимости показателей, подготавливает рекомендации для статистического бюро Секретариата ООН, координирует статистическую работу специализированных органов ООН, осуществляет консультации по вопросам сбора, накопления, разработки, анализа и распространения статистической информации. Так, при переписи населения Российской Федерации в 2010 году проводились предварительные консультации органов отечественной статистики со статистической комиссией ООН по организационным и программно-методологическим вопросам.

Деятельность статистических служб стран – членов Содружества независимых государств координируется Статистическим комитетом СНГ, созданным, прежде всего для разработки и осуществления на основе взаимных консультаций единой статистической методологии.

ГЛАВА 2. ИСТОЧНИКИ СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ

2.1. Понятие о статистическом наблюдении

Принятие правильного решения (по сути дела, по любому вопросу) предполагает наличие исчерпывающей и достоверной информации об изучаемом объекте. Формирование информационной базы начинается со сбора статистической информации.

Если при сборе статистических данных допущена ошибка или материал оказался недоброкачественным, это повлияет на правильность и достоверность как теоретических, так и практических выводов. Поэтому статистическое наблюдение от начальной до завершающей стадии должно быть тщательно продуманным и четко организованным.

Статистическое наблюдение – это представляет собой научно-организованный по единой программе учет и сбор фактов, характеризующих явления и процессы общественной жизни. Однако не всякий сбор сведений является статистическим наблюдением. О статистическом наблюдении можно говорить лишь тогда, когда, во-первых, обеспечивается регистрация устанавливаемых фактов в специальных учетных документах (формулярах) и, во-вторых, изучаются статистические закономерности, то есть такие, которые проявляются только в массовом процессе, в большом числе единиц определенной совокупности. Поэтому статистическое наблюдение должно быть *планомерным, массовым и систематическим*.

К статистическому наблюдению предъявляются требования:

- полноты и практической ценности статистических данных;
- достоверности и точности данных;
- их единообразия и сопоставимости.

Наблюдаемые явления должны иметь научную и практическую ценность, выражать определенные социально-экономические типы явлений. Необходимо обеспечивать сбор полных данных, относящихся к рассматриваемому вопросу. Поскольку общественные явления находятся в постоянном изменении, развитии, то при отсутствии полных данных, анализ и выводы могут быть ошибочными. Для обеспечения достоверности статистических данных необходима тщательная проверка качества собираемых фактов, логический и счетный контроль данных. Сопоставимость данных означает, что за анализируемый период фиксируемые показатели должны быть в одинаковой размерности и неизменны по своему содержанию, как во времени, так и у различных единиц совокупности.

2.2. Программно-методологические и организационные вопросы статистического наблюдения

Любое статистическое исследование необходимо начинать с точной формулировки его цели и конкретных задач, а тем самым и тех сведений, которые могут быть получены в процессе наблюдения. После этого определяются объект и единица наблюдения, разрабатывается программа, выбираются вид и способ наблюдения.

Цель наблюдения – это основной результат статистического исследования. Так, при переписи населения ставится цель выявления численности населения, проживающего на определенной территории, поло – возрастной структуры населения, сложившейся на данный момент времени и т.д.

Статистическое наблюдение проводится по плану и программе составленным заранее.

Программа наблюдения – это перечень признаков, регистрируемых в процессе наблюдения. С разработкой программы наблюдения связано определение цели, объекта и единицы наблюдения.

Объект наблюдения – совокупность социально-экономических явлений и процессов, которые подлежат исследованию, или точные границы, в пределах которых будут регистрироваться статистические сведения. Например, при переписи населения необходимо установить, какое именно население подлежит регистрации – наличное, то есть фактически находящееся в данной местности в момент переписи, или постоянное, то есть живущее в данной местности постоянно. Если ставится задача охарактеризовать развитие торговли в регионе, необходимо определиться, товарооборот розничной или оптовой торговли мы будем фиксировать и т.п.

В ряде случаев для отграничения объекта наблюдения пользуются тем или иным цензом. Ценз есть ограничительный признак, которому должны удовлетворять все единицы изучаемой совокупности. Например, в совокупность обследуемых предприятий торговли будут включены только те, у которых товарооборот в месяц составляет не менее определенной величины, допустим, 300 тысяч рублей.

Единица наблюдения – это та первичная единица, от которой должны быть получены необходимые сведения. Единицу наблюдения следует отличать от единицы совокупности.

Единица совокупности – это первичный элемент объекта, являющийся носителем признаков и являющейся единицей счета.

Иначе говоря, единица совокупности – это то, что подвергается обследованию, а единица наблюдения – это источник получаемых сведений. Например, при проведении переписи оборудования в промышленности единицей наблюдения является промышленное предприятие, а единицей совокупности – промышленное оборудование.

К программе наблюдения предъявляется ряд требований. Важнейшие из них:

- программа должна содержать существенные признаки, непосредственно характеризующие изучаемые явление, его тип, основные черты, свойства;

- в программу не следует включать второстепенные вопросы, чтобы не затруднять работу по сбору, обработке и анализу информации;

- при разработке программы необходимо стремиться к полноте собираемых сведений;

- в программу должны включаться только такие вопросы, на которые действительно можно получить объективные и достаточно точные ответы;

- формулировка вопросов должна быть такой, чтобы понималась единообразно;

- иногда в программу следует включить контрольные вопросы.

Для получения достоверных данных составляется *инструкция.* Это совокупность разъяснений и указаний по программе наблюдения. Она должна быть простой, краткой и ясной. Инструкция может быть в виде отдельной брошюры или в виде пояснений к вопросам.

Для записи ответов на вопросы программы разрабатывается *формуляр* наблюдения. Это особый бланк, в котором содержится перечень вопросов программы. Обязательными элементами формуляра являются титульная часть, которая включает наименование наблюдения, наименование организации проводящей исследование, кем и когда утвержден формуляр и номер его, присвоенный в общей системе формуляров. Примером формуляра может быть любой бланк отчетности.

Иногда в формуляре после вопроса сразу же даются некоторые варианты ответа на него. Это называется *статистическим подсказом.*

При организации статистического наблюдения необходимо решить вопрос о времени данного наблюдения, включая выбор сезона, установление срока (периода) наблюдения, а в некоторых случаях и так называемого критического момента.

Период (срок) наблюдения – это время, в течение которого осуществляется регистрация единиц наблюдения по установленной программе. Он, как правило, обозначается указанием даты (иногда и часа) начала и окончания наблюдения. В некоторых случаях указывается число дней, в течение которых оно должно быть проведено. Для некоторых наблюдений указывается не позднее какого числа данные должны быть представлены по назначению.

Критическим моментом статистического наблюдения (как правило, переписи) называется момент времени, по состоянию на который производится регистрация сведений.

Организационный план статистического наблюдения – это документ, в котором содержится перечень подготовительных работ и проведения статистического наблюдения с указанием конкретных сроков их проведения.

В организационном плане указываются:

- объекты наблюдения (его определение, описание и отличительные признаки);
- цели и задачи наблюдения;
- органы наблюдения, осуществляющие подготовку и проведение наблюдения и несущие ответственность за эту работу;
- время и сроки наблюдения;
- подготовительные работы: подбор и подготовка кадров, инструктаж, подготовка статистического инструментария (бланков, инструкций и т.д.), в некоторых случаях пропаганда через печать, радио, телевидение;
- порядок проведения наблюдения;
- порядок приема и сдачи материалов наблюдения;
- порядок получения и представления предварительных и окончательных итогов и другие работы.

Организационные планы составляются разными звеньями системы статистических учреждений от высших до низших.

Помимо постоянных органов, осуществляющих статистическое наблюдение, иногда создаются временные органы для проведения, как правило, крупных обследований.

2.3. Формы, виды и способы наблюдения

В статистической практике используются *две* организационные *формы* наблюдения – отчетность и специальное статистическое обследование.

Отчетность – это такая организационная форма, при которой единицы наблюдения представляют сведения о своей деятельности в виде формуляров регламентированного образца.

Особенность отчетности состоит в том, что она обязательна, документально обоснована и юридически подтверждена подписью руководителя. Каждый отчет содержит определенные реквизиты: номер или индекс формы, название отчета, отчетный период или на какую дату составляется отчет, название предприятия, где расположено. Отчетность может быть общегосударственной и ведомственной. Общегосударственная отчетность поступает в вышестоящие органы и в органы государственной власти. Она необходима для целей обобщения, контроля, анализа и прогнозирования. Ведомственная – используется в министерствах и ведомствах для оперативных нужд. Отчетность утверждается Росстатом.

С*пециального организованное статистическое наблюдение* – это наблюдение, организованное с некоторой определенной целью для получения сведений, которых нет в отчетности, или с целью проверки данных отчетности. Это переписи населения, скота, оборудования, всевозможные единовременные обследования.

Перепись – форма наблюдения, при которой осуществляется учет численности и характеристика состава изучаемого явления путем записи в статистических формулярах данных об обследуемых единицах статистической совокупности.

По характеру регистрации фактов различают *текущее или систематическое наблюдение* и *прерывное.*

При систематическом наблюдении регистрация фактов происходит по мере их совершения, например, учет производства продукции. Прерывное наблюдение может быть периодическим, т.е. через определенные промежутки времени, например, перепись населения или регистрация цен на рынке.

По источнику сведений различают непосредственное наблюдение, документальное и опрос.

- *Непосредственный учет фактов* – это когда сведения получают путем личного учета единиц совокупности: пересчета, взвешивания, измерения и т. д., что требует значительных затрат труда.

- *Документальный* сбор статистической информации базируется на систематических записях в первичных документах, подтверждающих тот или иной факт, например, счета клиентов, свидетельства о рождении.

- *Опрос* предполагает запись сведений со слов опрашиваемого.

С точки зрения полноты охвата фактов статистическое наблюдение может быть сплошным и не сплошным.

Сплошное наблюдение представляет собой полный учет всех единиц изучаемой совокупности.

Не сплошное наблюдение организуют как учет части единиц совокупности, на основе которой можно получить обобщающую характеристику всей совокупности. При этом подлежащая изучению статистическая совокупность, из которой производится отбор части единиц, называется *генеральной совокупностью*. К видам не сплошного наблюдения относятся: способ основного массива, выборочные наблюдения, монографические описания.

Способ основного массива заключается в том, что сбор данных осуществляется только по тем единицам совокупности, у которых величины изучаемого признака является преобладающей, например, чтобы изучить структуру грузооборота морских перевозок, достаточно исследовать крупнейшие морские торговые порты (Мурманск, Находка, Новороссийск).

Монографическое обследование представляет собой детальное изучение и описание отдельных, характерных в каком-либо отношении единиц совокупности. Например, при изучении и распространении передового опыта, новой прогрессивной организации производства и сбыта продукции, углубленно исследуется отдельное предприятие.

Выборочное наблюдение получило наибольшее признание и распространение в статистической практике. При выборочном наблюдении характеристика всей совокупности фактов дается по некоторой их части, отобранной в случайном порядке. Так, контроль качества продукции или изучение спроса на продукцию осуществляются при использовании данного вида наблюдений. Для обеспечения репрезентативности (представительности) выборки необходимо соблюдение принципа случайности отбора единиц. Это предполагает, что на включение или исключение объекта из выборки, не может повлиять какой-либо иной фактор, кроме случая.

Сведения об изучаемом явлении могут быть собраны различными способами: *экспедиционным, анкетным или корреспондентским*.

При *экспедиционном способе* специально подготовленный регистратор опрашивает людей и с их слов заполняет бланк обследования. Работа регистратора гарантирует единообразное понимание вопросов и максимальную правильность ответов. При *анкетном способе* определенному кругу лиц вручают специальные анкеты, заполнение которых носит добровольный характер и осуществляется анонимно. Это снижает полноту и достоверность полученной информации. При *корреспондентском* – бланки обследования и указания по их заполнению рассылаются по адресам. Респондент заполняет их и отсылает обратно, при этом проверить достоверность и качество полученной информации практически невозможно.

2.4. Ошибки статистического наблюдения

Точность и достоверность собираемой статистической информации – важнейшая задача статистического наблюдения. Материалы, собранные в результате наблюдения, подвергаются всесторонней проверке и контролю.

В зависимости от характера и степени влияния на конечные результаты, а также, исходя из источников возникновения ошибок, их принято подразделять на ошибки регистрации и ошибки репрезентативности.

Ошибки регистрации возникают вследствие неправильной записи. Они могут быть случайными и систематическими. Случайные ошибки регистрации возникают в результате случайных причин: оговорки, случайной перестановки цифр и т.д. В силу действия закона больших чисел в достаточно большой совокупности они взаимопогашаются.

Систематические ошибки регистрации возникают вследствие определенных причин и приводят к искажениям. Например, округление возраста до нуля или до пяти. Особенно большой вред наносят преднамеренные искажения.

Преднамеренные ошибки завышают или занижают конкретные значения признака, показателя. Известно, что в последние годы в Российской Федерации наблюдается массовое сокрытие предприятиями и фирмами прибыли или объектов прибыли от налогообложения. Программа статистического наблюдения предусматривает проверку расчетов прибыли налоговой инспекцией на каждом предприятии. За злостные ошибки к предприятиям или лицам применяются экономические и административные меры.

Ошибки репрезентативности свойственны только выборочному наблюдению. Они показывают, в какой степени выборочная совокупность представляет (репрезентирует) генеральную совокупность. Если выборка организована с нарушениями, то такая ошибка существенно искажает результаты наблюдения.

В соответствии с особенностями ошибок статистического наблюдения для проверки его результатов могут применяться логический или счетный контроль собранной информации.

ГЛАВА 3. СВОДКА И ГРУППИРОВКА МАТЕРИАЛОВ

СТАТИСТИЧЕСКИХ ДАННЫХ

3.1. Сущность и задачи сводки и группировки

Статистическая информация, полученная в результате статистического наблюдения, представляет собой большое количество первичных сведений об отдельных единицах объекта исследования. Задача статистики – привести эти материалы в систематизированный вид и на этой основе дать сводную характеристику всей изучаемой совокупности, выявить закономерность.

Сводка – это операция по обработке конкретных единичных фактов, образующих совокупность и собранных в результате наблюдения. В результате сводки множество индивидуальных показателей относящихся к каждой единице объекта наблюдения, превращаются в систему статистических таблиц и итогов, проявляются типические черты и закономерности изучаемого явления в целом. По глубине и точности обработки различают сводку простую и сложную.

Простая сводка – это операция по подсчету общих итогов, т.е. по совокупности единиц наблюдения.

Сложная сводка – это комплекс операций, включающих группировку единиц наблюдения, подсчет итогов по каждой группе и по объекту в целом, оформление результатов в виде статистических таблиц.

Проведение сводки включает следующие этапы:

- выбор группировочного признака;
- определение порядка формирования группы;
- разработка системы показателей для характеристики групп и объекта в целом;
- разработка макетов таблиц для представления результатов сводки.

По форме обработки сводка бывает:

- *централизованная* (весь первичный материал поступает в одну вышестоящую организацию, например, в Росстат, и там полностью обрабатывается);

- *децентрализованная* (обработка собранного материала идет по восходящей линии, т.е. материал подвергается сводке и группировке на каждой ступени).

На практике обычно сочетают обе формы организации сводки. Так, например, при переписи населения предварительные итоги получают в порядке децентрализованной сводки, а сводные окончательные – в результате централизованной разработки бланков переписи.

По технике выполнения сводка бывает *механизированной* и *ручной.*

Группировкой называется расчленение изучаемой совокупности на однородные группы по определенным существенным признакам. Признак, лежащий в основе группировки, называется *группировочным.* Группировочный признак может иметь *количественное выражение* и может быть *атрибутивным (качественным).* Первые имеют числовое выражение (объем товарооборота, возраст человека, доход семьи и т.д.), а вторые отражают состояние единицы совокупности (пол, семейное положение, отраслевая принадлежность предприятия, профессия рабочего, форма собственности и т.д.).

На основе метода группировок решаются центральные задачи исследования, обеспечивается правильное применение других методов статистического и статистико-математического анализа.

Работа по составлению группировок сложная и трудная. Приемы группировок разнообразны, что обусловлено разнообразием группировочных признаков и различными задачами исследования. К основным задачам, решаемым с помощью группировок, относятся:

- выделение социально - экономических типов;
- изучение структуры совокупности, структурных сдвигов в ней;
- выявление связи между явлениями и взаимозависимости.

От группировок следует отличать классификации. *Классификация* – это систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основе их сходства и различия. В основе классификации всегда атрибутивный (качественный) признак. Классификации стандартны, устойчивы и неизменны в течение длительного периода времени. Например, классификация затрат на производство по экономическим элементам, классификация основных фондов и т.д. Классификация это своеобразный статистический стандарт, установленный на определенный промежуток времени. Например, ОКВЭД – общероссийский классификатор видов экономической деятельности, классификация основных фондов по натурально-вещественному составу и т.д.

3.2. Виды группировок

В зависимости от задач, решаемых с помощью группировок, выделяют 3 вида группировок: *типологические, структурные и аналитические.*

Типологическая группировка решает задачу выявления социально-экономических типов. При построении группировки этого вида основное внимание должно быть уделено идентификации типов и выбору группировочного признака. Исходят при этом из сущности изучаемого явления. В основе такой группировки лежит атрибутивный признак.

Таблица 3.1

Распределение предприятий и организаций Российской Федерации
по формам собственности на 1 января 2014 г.

Форма собственности	Количество предприятий и организаций
А	1
Государственная	155064
Муниципальная	231039
Частная	2725901
Общественных и религиозных организаций	236755
Прочие формы собственности	245078
Итого по РФ	3593837

Структурная группировка решает задачу изучения состава отдельных типических групп по какому-то признаку. В основе такой группировки может быть как количественный, так и атрибутивный признак. Например, в группировке постоянного населения по возрастным группам количественный признак, а в группировке распределение населения на городское и сельское – атрибутивный признак.

Таблица 3.2

Распределение населения Российской Федерации
на городское и сельское на 1 января 2015 года

Группы населения	Численность в % к итогу
А	1
Городское	74
Сельское	26
Итого	100

Аналитическая группировка позволяет выявить взаимосвязи между явлениями и их признаками, т.е. выявить влияние одних признаков (*факторных)* на другие *(результативные).* Взаимосвязь проявляется в том, что с возрастанием факторного признака возрастает или убывает значение результативного признака. В основе аналитической группировки всегда лежит *факторный* признак, а каждая группа характеризуется *средними* величинами результативного признака.

Например, зависимость объема розничного товарооборота от величины торговой площади магазина. Здесь факторный (группировочный) признак – торговая площадь, а результативный – средний на 1 магазин объем товарооборота или зависимость объема произведенной продукции (результативный признак) от величины основных производственных фондов (факторный признак).

Таблица 3.3

Зависимость объема произведенной продукции

от величины основных производственных фондов

Группы предприятий по стоимости основных производственных фондов, млн. руб.	Количество предприятий	Объем произведенной продукции в среднем на 1 предприятие, млн. руб.
А	1	2
2-3	20	2,9
1	2	3
3-4	14	3,1
4-5	16	5,1
5-6	12	7,2
6-7	10	7,3
7 и более	8	8,0
Итого	80	4,9

По сложности группировка бывает простой и сложной (комбинированной).

В *простой* группировке в основании один признак (таблицы 3.1; 3.2; 3.3), а в *сложной* – два и более в сочетании (в комбинации). В этом случае сначала группы образуются по одному (основному) признаку, а затем каждая из них делится на подгруппы по второму признаку и т.д. Это позволяет изучить совокупность сразу по нескольким признакам одновременно (таблица 3.4).

Таблица 3.4

Группировка сельхозпредприятий по наличию сельхозугодий
и основных фондов на 1 работника по области за отчетный период

№ групп	Группы сельхозпредприятий		Количество сельхоз. предприятий	Получено продукции на 100 га. сельхозугодий, тыс. руб.
	по наличию сельхозугодий на 1 работника, га.	По размеру основных фондов на 1 работника, тыс. руб. (фондовооруженность)		
1	2	3	4	5
1.	7,1-11,0	до 2,0 2,0-2,5 2,5 и более	36 25 19	15,5 18,3 19,3
	Всего по группе		80	17,1
2.	11,1-15,0	до 2,0 2,0-2,5 2,5 и более	7 38 24	14,8 14,9 17,6
	Всего по группе		69	15,8
3.	15,1-19,0	до 2,0 2,0-2,5 2,5 и более	3 11 7	12,8 12,9 12,9
	Всего по группе		21	12,9
	Итого		170	15,8

Группировка показывает, что с ростом фондовооруженности труда увеличивается выпуск продукции на 100 га сельхозугодий, а с увеличением наличия сельхозугодий на 1 работника выпуск продукции на 100 га сельхозугодий уменьшается.

Комбинационные группировки очень важны в экономических исследованиях. Однако в них по мере увеличения группировочных признаков растет число групп, и таблица становится малообозримой. Многофакторный анализ способен дать лишь система простых и комбинированных группировок.

3.3. Принципы построения статистических группировок

При проведении любой группировки сначала *определяется группировочный признак,* т.е. по которому расчленяется совокупность на группы.

Следующим этапом группировки является *определение числа групп*.

В группировках с атрибутивным признаком в основании число групп зависит от количества типов, а интервал соответствует переходу явления из одного качества в другое (таблица 3.1).

В группировках с количественным признаком в основании число групп рекомендуется брать с таким расчетом, чтобы в каждую группу попало достаточно большое число единиц совокупности. Интервалы таких группировок могут быть *равными (*таблицы 3.3; и 3.4) и *неравными* (таблицы 3.5; 3.6). Неравные интервалы, в свою очередь, могут быть *возрастающими* (таблица 3.5) и *убывающими* (таблица 3.6).

Таблица 3.5

Зависимость урожайности зерновых культур от количества внесенных удобрений по фермерским хозяйствам региона

Количество внесенных удобрений, в процентах от нормы.	Число фермерских хозяйств	Средняя урожайность, ц/га
А	1	2
До 30	3	20,2
30-50	5	27,6
50-80	8	32,4
80 и более	14	37,3
Всего	30	33,0

Таблица 3.6

Распределение семейных пар с детьми по возрасту женщины

Возраст женщины, лет	Количество семей
А	1
До 30	189
30-40	267
40-45	44
Итого	500

В группировках с равным интервалом число групп можно рассчитать математическим путем. С использованием, например, формулы Стерджесса: $n = 1 + 3.322 \lg N$, где n – число групп, а N – число единиц совокупности. Согласно этой формуле выбор числа групп зависит от объема совокупности. Недостаток этой формулы состоит в том, что ее применение дает хорошие результаты, если совокупность состоит из большого числа единиц и если распределение единиц по группировочному признаку близко нормальному. Поэтому есть ряд других формул, но каждая имеет свои недостатки.

Если размах вариации группировочного признака (разность между максимальным и минимальным его значениями в совокупности) велик и значения признака изменяются (варьируют) неравномерно, то надо использовать группировку с неравным интервалом.

Применение неравных интервалов обусловлено тем, что в первых группах небольшая разница в показателях имеет большое значение, а в последних группах эта разница не существенна. Возрастающий интервал может возрастать в арифметической прогрессии, а может – в геометрической. Использование неравного интервала более обосновано, но представляет большую трудность. Такую группировку можно составить только на основе знания исходного материала, его анализа и личного опыта специалиста. Главное условие и в этом случае, чтобы не было «пустых» или малочисленных групп.

Когда определено число групп, то следует определить интервалы группировки. Величина интервала – это разность между верхними или нижними его границами. Величину равновеликого интервала можно определить по формуле:

$$i = \frac{x_{\max} - x_{\min}}{n}$$

, где $x_{\max}$ – максимальное значение группировочного признака в совокупности; $x_{\min}$ – минимальное.

При образовании интервалов необходимо точно обозначить границы групп. По непрерывно варьирующим признакам образуют непрерывный интервал, т.е. такой, в котором верхняя граница предыдущего интервала равна нижней границе последующего интервала (таблицы 3.3; 3.5; 3.6). По прерывно варьирующим признакам образуют прерывный интервал, где верхняя граница предыдущего интервала не равна нижней границе последующего интервала (таблица 3.4; по первому признаку).

Интервалы группировки могут быть закрытыми и открытыми. Закрытые интервалы имеют обе (верхнюю и нижнюю) границы (таблица 3.4; по основному признаку). Открытые – только одну из них (таблицы 3.3; 3.4 по второму признаку; 3.5; 3.6).

При непрерывном интервале встает вопрос, в какую группу, например в таблице 3.6; включить семьи, где возраст женщин 40 лет? Надо включать по принципу «до». Следовательно, семьи, где возраст женщины 40 лет, попадут в третью группу.

Величину интервала в расчетах можно, а порою и нужно, округлять. При этом следует помнить, что по непрерывно варьирующим признакам, если величина интервала получилась целой без округления или округление произведено в меньшую сторону, последнюю группу необходимо делать с открытым интервалом, иначе максимальное значение (по принципу «до») не войдет в группу.

Пример 3.1.

Таблица 3.7

Предположим, имеются следующие данные по 30 заводам:

№ п/п	Стоимость основных фондов, млн. руб.	Объем продукции, млн. руб.	Численность персонала, чел.
1	2	3	4
1	3,0	2,3	359
2	2,2	2,8	330
3	4,4	4,6	410
4	3,3	3,0	500
5	2,5	2,7	385
6	2,1	2,0	300
7	5,0	6,8	600
8	5,1	6,2	400
9	2,7	2,9	450
10	6,8	7,6	490
11	3,1	3,1	380
12	3,5	4,0	450
13	2,0	2,5	250
14	2,4	2,0	395
15	7,9	8,4	560
16	4,1	4,9	320
17	5,5	8,8	608
18	6,3	7,0	433
19	5,4	8,2	500
20	2,6	3,2	332
21	4,7	5,8	520
22	5,6	6,8	390
23	2,3	3,9	454
24	2,8	3,9	520
25	2,9	3,5	440
26	5,3	6,3	500
27	4,2	4,9	450
28	4,0	5,5	407
29	7,2	7,6	444
30	4,8	5,2	460

Необходимо произвести группировку заводов по стоимости основных фондов. Для этого надо рассчитать число групп и величину равновеликого интервала. Результаты группировки изложить в табличной форме. Каждую группу и совокупность заводов в целом охарактеризовать:

1) количеством заводов;

2) объемом продукции всего и в среднем на один завод;

3) численностью персонала всего и в среднем на один завод.

В основе искомой группировки будет факторный (независимый) признак – стоимость основных фондов (x) в млн. руб. Определяем количество групп:

$n = 1 + 3{,}322\ \lg 30 = 1 + 3{,}322 \cdot 1{,}477 = 5{,}9 \approx 6$ групп.

Определим величину равновеликого интервала группировки:

$i = \dfrac{x_{\max} - x_{\min}}{n} = \dfrac{7{,}9 - 2{,}0}{6} = 0{,}98$ (1 млн. руб.).

Следовательно, в подлежащем макета таблицы будет группировка заводов по стоимости промышленно-производственных основных фондов, в 1 млн. руб.

Запишем группировку с прерывным интервалом:

2,0-2,9

3,0-3,9

4,0-4,9

5,0-5,9

6,0-6,9

7,0 и более.

Возможные варианты записи группировки с непрерывным интервалом:

2-3	До 3	2-3	До 3
3-4	3-4	3-4	3-4
4-5	4-5	4-5	4-5
5-6	5-6	5-6	5-6
6-7	6-7	6-7	6-7
7-8	7 и более	7 и более	7-8

После того как разработано подлежащее таблицы, надо определить сказуемое. Результативным (зависимым) признаком (y) будет объем продукции. Кроме того, надо подсчитать численность единиц совокупности (заводов) по каждой группе и в целом, получится общее число заводов – 30 (контрольная цифра). Кроме этого, в макет таблицы необходимо включить численность персонала. Следовательно, макет таблицы будет иметь следующий вид:

Таблица 3.8

Группировка заводов по стоимости основных фондов

Группы заводов по стоимости основных фондов, млн. руб.	Количество заводов	Объем продукции, млн. руб.		Численность персонала, чел.	
		всего	в среднем на один завод	всего	в среднем на один завод
А	1	2	3	4	5
2-3					
3-4					
4-5					
5-6					
6-7					
7 и более					
Итого					

Составляем разработочную таблицу:

№ завода по исходным данным	Стоимость основных фондов, млн. руб.	Объем продукции, млн. руб.	Численность персонала, чел.
I группа 2-3			
2	2,2	2,8	330
5	2,5	2,7	385
6	2,1	2,0	300
9	2,7	2,9	450
13	2,0	2,5	250
14	2,4	2,0	395
20	2,6	3,2	332
23	2,3	3,9	454
24	2,8	3,9	520
25	2,9	3,5	440
Итого	24,6	29,4	3856
II группа 3-4			
1	3,0	2,3	359
4	3,3	3,0	500
11	3,1	3,1	380
12	3,5	4,0	450
Итого	12,9	12,4	1689
III группа 4-5			
3	4,4	4,6	410
16	4,1	4,9	320
21	4,7	5,8	520
27	4,2	4,9	450
28	4,0	5,5	407
30	4,8	5,2	460
Итого	22,2	30,9	2567
IV группа 5-6			
7	5,0	6,8	600
8	5,1	6,2	400
17	5,5	8,8	608
19	5,4	8,2	500
22	5,6	6,8	390
26	5,3	6,3	500
Итого	31,9	43,1	2998

V группа 6-7

10	6,8	7,6	490
18	6,3	7,0	433
Итого	13,1	14,6	923

VI группа 7 и более

15	7,9	8,4	560
29	7,2	7,6	444
Итого	15,1	16,0	1004
Всего	1198	146,4	13037

На основании итоговых строк разработочной таблицы заполняем макет таблицы.

Таблица 3.9

Группировка заводов по стоимости основных фондов

Группы заводов по стоимости основных фондов, млн. руб.	Количество заводов	Объем продукции, млн. руб.		Численность персонала, чел.	
		всего	в среднем на 1 завод	Всего	в среднем на 1 завод
А	1	2	3	4	5
2-3	10	29,4	2,9	3856	386
3-4	4	12,4	3,1	1689	422
4-5	6	30,9	5,1	2567	428
5-6	6	43,1	7,2	2998	450
6-7	2	14,6	7,3	923	461
7 и более	2	16,0	8,0	1004	502
Итого	30	146,4	4,9	13037	435

Как видно из таблицы, с увеличением стоимости основных фондов объем продукции возрастает. Средний объем продукции одного завода составил 4,9 млн. руб. Численность персонала с увеличением стоимости основных фондов также увеличивается и составляет в среднем 435 человек.

Группировка, в которой известна только численность групп или удельный вес группы в общем итоге, называется *рядом распределения*. Ряды распределения, как и всякие другие группировки, могут быть по количественному и атрибутивному признакам. Ряды распределения по атрибутивному признаку называются *атрибутивными*, а построенные по количественному признаку – *вариационными*. Вариационные ряды распределения состоят из двух элементов: вариантов и частот. *Вариантами* называют числовые значения количественного признака. Частотой являются численность каждого варианта (повторяемость). Сумма всех частот является объемом совокупности. Частоты, выраженные в виде относительных величин, называются *частостями.*

Ряд распределения по количественному признаку может быть дискретным и интервальным. В *дискретных* рядах распределения варианты признака имеют значения целых чисел, т.е. между ними не может быть никаких промежуточных значений. Например, распределение рабочих по тарифному разряду, когда группировочный признак в каждой группе конкретное число: 2 разряд, 3, 4, 5, 6. В *интервальных* же рядах распределения группировочный признак может принимать любые значения в некотором промежутке (таблицы 3.3-3.6).

3.4. Вторичная группировка

Для достижения сопоставимости данных различных группировок используют вторичную группировку.

Вторичная группировка или перегруппировка – это образование новых групп на основе уже сгруппированных данных. Вторичную группировку можно осуществить двумя способами: *укрупнения интервала и долевой перегруппировки. Укрупнение интервала* заключается в объединении двух, трех и т.д. прежних интервалов. Этот способ чрезвычайно прост, но не всегда приемлем. При *долевой перегруппировке* за каждой группой закрепляется определенная доля единиц совокупности. Долевая перегруппировка может быть по величине прежнего интервала и по удельному весу групп в общей численности. Долевая перегруппировка *по величине прежнего интервала* самый распространенный способ перегруппировки.

Пример 3.2.

Предположим, необходимо сопоставить данные группировок магазинов по товарообороту на 10 кв. м. торговой площади магазина двух областей.

Таблица 3.10

Область «А»

№№ п/п	Группы магазинов по величине товарооборота на 10 кв. м. торговой площади, тыс. руб.	Количество магазинов	
		единиц	% к итогу
1	2	3	4
1	до 10	812	2,2
2	10-40	3037	8,1
3	40-100	10597	28,2
4	100-180	13355	35,6
5	180-300	8048	21,5
6	300-400	1205	3,2
7	400 и более	469	1,2
	Итого	37523	100,0

Таблица 3.11

Область «Б»

№ п/п	Группы магазинов по величине товарооборота на 10 кв. м. торговой площади, тыс. руб.	Количество магазинов	
		единиц	% к итогу
1	2	3	4
1	до50	350	0,9
2	50-100	598	1,6
3	100-200	10843	28,5
4	200-400	24260	63,9
5	400 и более	1944	5,1
	Итого	37995	100,0

В таком виде они несопоставимы. Так как группировка по области «A» методологически более верная, осуществляем перегруппировку данных по области «A» в группировку области «$Б$» способом по величине прежнего интервала. Для группировки области «A» новая величина интервала первой группы 50. Чтобы получить интервал 50 из прежних интервалов необходимо суммировать интервалы первой и второй групп, а от третьего интервала взять $\frac{1}{6}$ часть (оставшиеся 10 единиц нового интервала / величину интервала третьей прежней группы – 60). Остальные $\frac{5}{6}$ интервала пойдут во вторую новую группу, где величина интервала также 50 $\left(\frac{5}{6}\cdot 60 = 50\right)$. Точно в таком же соотношении распределяем по новым группам количество магазинов:

Первая группа $\left(812 + 3037 + \frac{1}{6}\cdot 10597 = 5615\right)$, вторая группа $\left(\frac{5}{6}\cdot 10597 = 8831\right)$.

В третьей группе новой группировки величина интервала 100. Чтобы получить 100 из прежних интервалов, суммируем интервал четвертой группы 80 и $\frac{1}{6}$ интервала пятой группы $\left(\frac{20}{120}\right)$ и т.д.

Таблица 3.12

Перегруппировка данных области «А»

№№ п/п	Группы магазинов по величине товарооборота на 10 кв. м. торговой площади, тыс. руб.	Расчет интервала	Расчет количества магазинов
1	2	3	4
1	до 50	$10+30+\frac{1}{6}\cdot 60=50$	$812+3037+\frac{1}{6}\cdot 10597=5615$
2	50-100	$\frac{5}{6}\cdot 60=50$	$\frac{5}{6}\cdot 10597=8831$
3	100-200	$80+\frac{20}{120}\cdot 120=100$	$13355+\frac{20}{120}\cdot 8048=14696$
4	200-400	$\frac{100}{120}\cdot 120+100=200$	$\frac{100}{120}\cdot 8048+1205=7912$
5	400 и более	без изменения	469
	Итого		37523

Таблица 3.13

Окончательный вид новой группировки по области «А»:

№№ п/п	Группы магазинов по величине товарооборота на 10 кв. м. торговой площади, тыс. руб.	Количество магазинов	
		единиц	% к итогу
1	2	3	4
1	до 50	5615	15,0
2	50-100	8831	23,5
3	100-200	14696	39,2
4	200-400	7912	21,1
5	400 и более	469	1,2
	Итого	37523	100,0

Теперь можно сопоставить данные этих группировок и сделать вывод, что в области «А» наибольший удельный вес 39,2 % составляют магазины с величиной товарооборота на 10 кв. м. торговой площади магазина от 100 до 200 тыс. руб., а в области «Б» соответственно 63,9 % магазинов имеют от 200 до 400 тыс. руб. на 10 кв. м. торговой площади магазина.

Аналогично делают перегруппировку и в тех случаях, когда группировка выполнена с нарушениями принципов группировки и все данные пересчитывают в вновь составленную группировку.

ГЛАВА 4. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ И ГРАФИКИ

4.1. Сущность и виды статистических таблиц

Таблица представляет собой наиболее рациональную форму изложения результатов сводки и группировки.

Если из таблицы изъять все слова и цифры, то получится графлёная сетка, так называемый *скелет таблицы*. Вертикальные столбцы называют графами, горизонтальные – строками. Если записать заголовки таблицы, граф и строк, получится *макет таблицы*. *Полная таблица* – это макет, заполненный результатами сводки и группировки.

Таблица имеет подлежащее и сказуемое. *Подлежащее* таблицы показывает, о чем идет речь в таблице, т.е. это объект, который характеризуется цифрами. Это может быть одна или несколько совокупностей перечня или сгруппированные по каким-либо признакам территориальные единицы.

Сказуемое таблицы показывает, каким признаком или признаками характеризуется подлежащее, т.е. это система показателей, которыми характеризуется объект изучения.

Обычно подлежащее таблицы располагается слева сверху вниз (содержание строк), а сказуемое сверху слева направо (содержание граф). Однако в отдельных случаях для более полного и лучшего способа прочтения и анализа исходной информации расположение подлежащего и сказуемого может меняться местами.

В зависимости от структуры подлежащего таблицы, от группировки единиц в нем различают 3 вида таблиц: простые, групповые и комбинационные.

Простые таблицы (перечневые) в подлежащем содержат лишь перечень каких-либо объектов или их единиц, территориальных единиц или временных единиц (таблица 4.1).

Таблица 4.1

Пассажирооборот отдельных видов транспорта общего пользования в Республике Татарстан

Вид транспорта	Пассажирооборот, тыс. пассажиро-километров
А	1
Все виды транспорта в том числе:	12136
Автобусный	6651
Троллейбусный	416
Трамвайный	1061
Железнодорожный	3200
Воздушный	739
Внутренний водный	69

Групповые таблицы в подлежащем имеют группировку по одному признаку (таблицы 3.1; 3.2; 3.4-3.6).

Комбинационные таблицы содержат в подлежащем группировку по двум и более признакам (таблица 3.3). Комбинационная таблица, например, по двум признакам не может быть заменена двумя групповыми таблицами, т.к. они не дают описания явления во взаимосвязи. В комбинационной таблице нельзя произвольно менять место признака в комбинации. Признаки нужно располагать либо по важности, либо по последовательности изучения.

Сказуемое статистической таблицы может быть с простой и сложной разработкой. *При простой разработке сказуемого* каждый признак подлежащего характеризуется одним или несколькими, но изолированными друг от друга признаками. Например, распределение населения по полу, на городское и сельское:

Области	Все население	в том числе:			
		городское	сельское	мужчины	женщины

Здесь население областей характеризуется двумя изолированными друг от друга признаками.

При комбинации этих признаков (*сложная разработка сказуемого*) заголовки граф будут выглядеть по-другому. Здесь признаки будут связаны между собой:

<table>
<tr><td rowspan="3">Области</td><td rowspan="3">Все население</td><td colspan="4">в том числе:</td></tr>
<tr><td colspan="2">городское</td><td colspan="2">сельское</td></tr>
<tr><td>мужчины</td><td>женщины</td><td>мужчины</td><td>женщины</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Вид таблицы не зависит от способа разработки сказуемого.

4.2. Правила построения, оформления, переноса таблиц и записи цифр в них

Статистические таблицы должны быть грамотно выполнены.

Основные *правила построения* таблиц:

1. Таблица должна быть компактной и содержать только те данные, которые непосредственно отражают исследуемое явление и необходимы для достижения цели исследования.

2. Цифровой материал необходимо излагать таким образом, чтобы при анализе таблицы сущность явления раскрывалась чтением строк слева направо и сверху вниз.

3. Заголовок таблицы и названия граф и строк должны быть четкими, краткими, лаконичными и представлять собой законченное целое и вписываться ограниченно в содержание текста.

4. Таблица должна содержать итоговую строку (или графу). Существуют различные способы соединения слагаемых граф (строк) с их итогом:

- строка «итого» или «всего» завершает таблицу (таблицы 2.1-2.6);

- итоговая строка располагается первой строкой таблицы и соединяется с совокупностью ее слагаемых словами «в том числе» или «из них» (таблица 4.1).

5. Если названия отдельных граф повторяются между собой, содержат

повторяющиеся термины или несут единую смысловую нагрузку, то необходимо им присвоить объединяющий заголовок.

6. Графа (строка), характеризующая численность единиц совокупности по изучаемому признаку должна быть первой графой (строкой) сказуемого.

7. Графы и строки полезно нумеровать, если их много.

Правила оформления таблиц:

1. Заголовок таблицы должен отражать объект, признак, время и место совершения события, а также единицы измерения, если они одинаковые для всей таблицы, оговорку о границах, если данные не сопоставимы по территории.

2. Заголовки граф и строк должны быть краткими и записаны по возможности полными словами, а также содержать единицы измерения, если они разные.

3. Графы подлежащего и строки нумеруются заглавными буквами, а графы сказуемого и содержание строк арабскими цифрами.

4. Страны, области, края, города располагаются в алфавитном порядке или по значимости.

5. Если данные приводятся за многие годы, они должны быть расположены в хронологическом порядке.

6. Для удобства пользования данными таблицы сначала следует приводить данные в абсолютных цифрах, а потом соответствующие им относительные.

7. Если в таблице приводятся проценты (коэффициенты) к какому-то предшествующему году, то этот год должен быть показан в таблице.

8. В случае необходимости дополнительной информации – разъяснений к таблице, могут даваться примечания.

Правила записи цифр в таблице:

1. Цифры записываются на пересечении граф и строк.

2. Если одно из численных выражений данного признака равно нулю

или явление отсутствует, то ставится прочерк «-».

3. Если численное значение признака неизвестно, то ставится многоточие «…» или пишется «нет сведений».

4. Если пересечение строки и графы не дает осмысленного содержания, то ставится знак «X».

5. По возможности цифры целесообразно округлять. Округление в пределах одной и той же графы или строки должно быть одинаковым. Если, например, округление до десятых долей, то все числа записываются до десятых (в целых числах пишется ноль десятых).

6. Если численное значение признака мало и не может быть записано в принятом округлении (например, до десятых долей), то оно записывается в этом случае как «0,0».

Правила переноса таблиц:

1. Заголовок таблицы нельзя отделять от таблицы.

2. Итоговую строку (графу) нельзя отделять от таблицы (должна быть записана хотя бы одна рядом стоящая строка (графа).

3. При переносе заголовки граф (строк) не переписываются, а только нумеруются (следовательно, в этом случае необходима нумерация при любом количестве граф и строк).

4.3. Понятие о статистическом графике

Статистический график – это графическое изображение, на котором с помощью условных геометрических образов или знаков описываются статистические совокупности. График выделяет, подчеркивает в изучаемой совокупности особенности; представляет ее свойства в наиболее обобщенном виде, концентрирует внимание на главном, позволяет делать достоверные выводы. Графический метод изучения статистических совокупностей применяется как в решении сложных задач математико-статистического анализа, так и для решения элементарных задач.

Если рассматривать статистический график в двумерном изображении, то в нем можно выделить следующие элементы:

- *поле графика* – эта пространство, в котором размещаются геометрические знаки, образующие график. Оно должно иметь определенные размеры и пропорции сторон. Размер графика должен соответствовать его назначению (выставка, лекция, доклад, отчет, книга, статья). Пропорции графика определяются эстетическими соображениями, т.е. должна быть геометрическая гармония. Этому требованию лучше соответствует график с неравными сторонами поля. В практике применяют форматы с соотношением сторон от 1:1,33 до 1:1,5. Удобны графики со сторонами поля $1:\sqrt{2}$. В некоторых случаях удобна квадратная форма поля.

- *Геометрические знаки* – это знаки, с помощью которых формируется понятие об отображаемых на графиках явлениях. Знаки-символы образуют основу графика. Это могут быть точки, отрезки прямых, круги, секторы, геометрические фигуры, силуэты. Правильный выбор символов для графика имеет важное значение. Знак направляет внимание читателя на нужный аспект. Выбор знака зависит от того, какой аспект явления необходимо подчеркнуть.

- *Пространственные ориентиры* определяют размещение знаков на поле графика. Они зависят от принятой системы координат. Если график строится в прямоугольных координатах, то на осях координат располагаются характеристики статистических признаков, а в поле графика соответственно располагаются знаки графика. В графиках, отображающих динамику, на оси абсцисс слева на право располагают отрезки времени, а на оси ординат – величины признака в возрастающем порядке.

В картограммах и картодиаграммах, как средство пространственной ориентации, применяют криволинейные контурные линии, которые выступают либо как географические ориентиры, либо как административные границы.

- *Масштабные ориентиры* дают геометрическим знакам количественную определенность. Они изображаются либо масштабными знаками в виде кругов, прямоугольников, квадратов, либо масштабной шкалой. Масштабная шкала применяется в координатных статистических графиках. В этом случае на осях координат делаются отметки, и каждый отрезок соответствует либо величине признака, либо периоду времени и т.д. Масштабные знаки применяются преимущественно для статистических карт. Обычно они выносятся за поле графика, а величина явления определяется сравнением с графическим знаком-эталоном.

- *Экспликация графика* – это словесное объяснение содержания графика и смыслового значения каждого знака на графике. Без экспликации график нельзя ни прочитать, ни понять. Словесное содержание графика подчиняется определенным правилам. Прежде всего, экспликация определяет конкретную область действительности, к которой относится график или задача, которая решается с помощью графика. Она включает заголовок, отражающий содержание графика с указанием места, времени и единиц измерения, надписи вдоль масштабных шкал, обозначение цифрами интервалов времени и т.д.

Прежде чем приступить к построению графика необходимо внимательно изучить данные и решить, какой тип графика наиболее отвечает поставленной цели. *Композиция графика* – творческий процесс, ограниченный лишь положениями общего характера. Цель построения графика сводится к тому, чтобы придать ему такую форму, которая наглядно представляла бы изучаемую совокупность, выявляя наиболее интересные и важные особенности. Для этого график должен точно отражать факты в сжатом, логическом и простом изображении.

Главное же требование к графику – это хорошая читаемость графика, простота и изящество (никаких деталей, отвлекающих от основного содержания графика).

Большая роль в композиции принадлежит эстетической стороне оформления графика. Это относится к оформлению масштабной шкалы (она не должна быть очень редкой, но и не очень густой), и к толщине линий, и к размерности знаков-символов (между шириной и высотой столбиков и расстоянием между ними) и т.д.

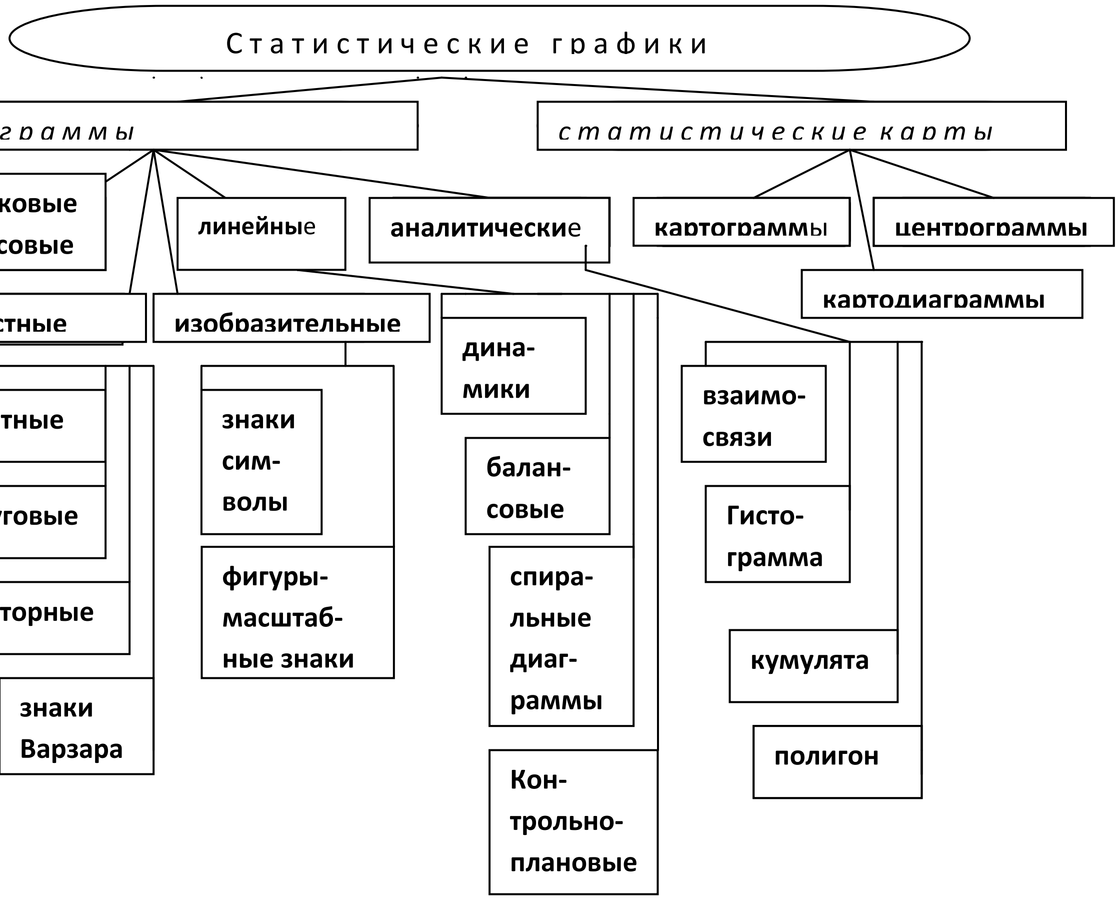
Очень важна в графике смысловая нагрузка. При большой смысловой нагрузке (большое количество информации) ослабляется выразительность графика, график делается громоздким, трудночитаемым. Излишнее упрощение обедняет график, делает его малополезным.

Цвет в статистических графиках используется как средство различия. Однако необходимо избегать пестроты. Лучше использовать спокойные тона. Различный цвет можно применять и для обозначения линий, и для обозначения структуры, а также интенсивности явления. Иногда цвет используют и как фон. Цветные графики более выразительны.

4.4. Классификация статистических графиков

Статистические графики, как по тематике, так и по композиции очень многообразны. В связи с этим существует множество различных классификаций графиков.

Классификация статистических графиков *по способу построения.*



4.5. Методика построения статистических графиков

4.5.1. Столбиковые и полосовые диаграммы

Столбиковые и полосовые диаграммы используются для *сопоставления* явлений *в пространстве*, для характеристики *структуры* явлений и их *динамики.* Они отображают величину явления *высотой столбика* или *длиной полосы.*

В столбиковых и полосовых диаграммах величина признака может быть выражена абсолютными, относительными и средними величинами. Используются столбиковые и полосовые диаграммы для сравнения одного или нескольких показателей в комбинации по различным объектам, для характеристики динамики явлений, для сравнения структуры явлений (рис.4.1-4.6).

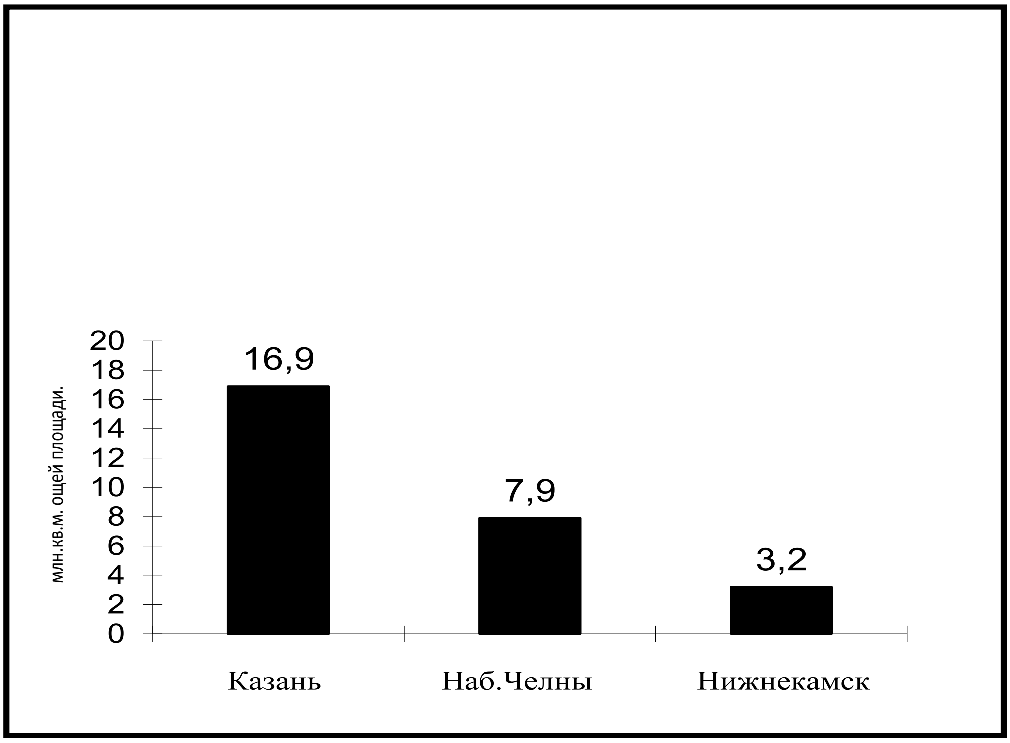


Рис. 4.1. Жилищный фонд Казани, Набережных Челнов и Нижнекамска на 1 января 2015 года

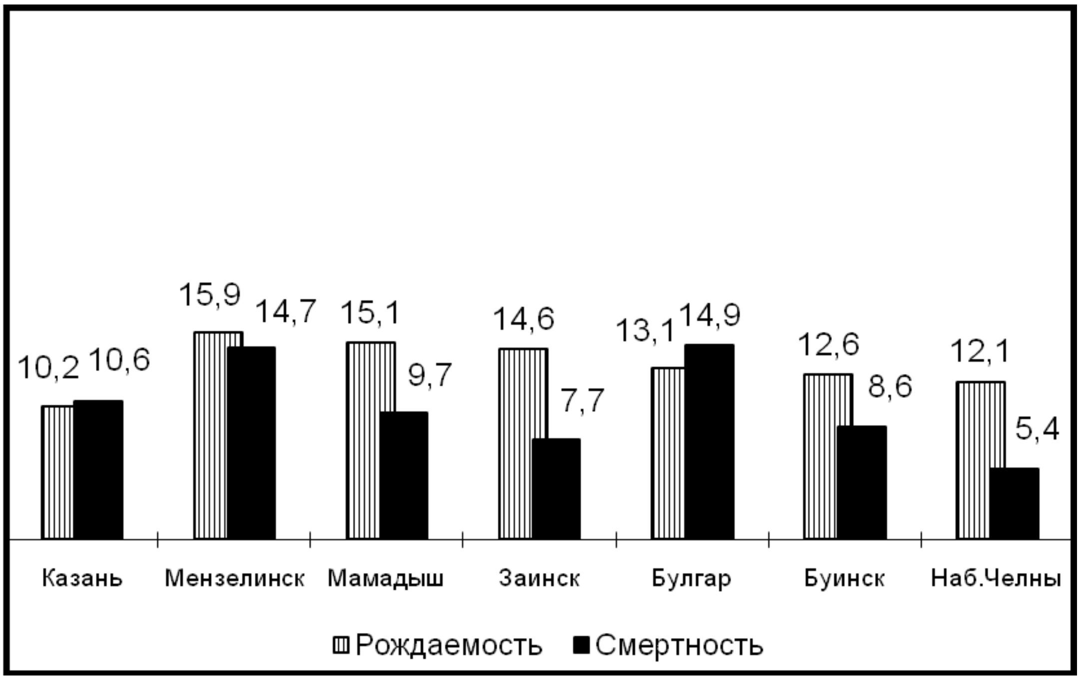


Рис. 4.2. Показатели рождаемости и смертности по некоторым городам РТ, число родившихся и умерших на 1000 человек населения

Рис. 4.3. Плотность населения в экономических районах РТ (чел. кв. км.)

Возрастная пирамида населения (рис. 4.4) с помощью полосовой диаграммы характеризует структуру населения по полу и возрасту.

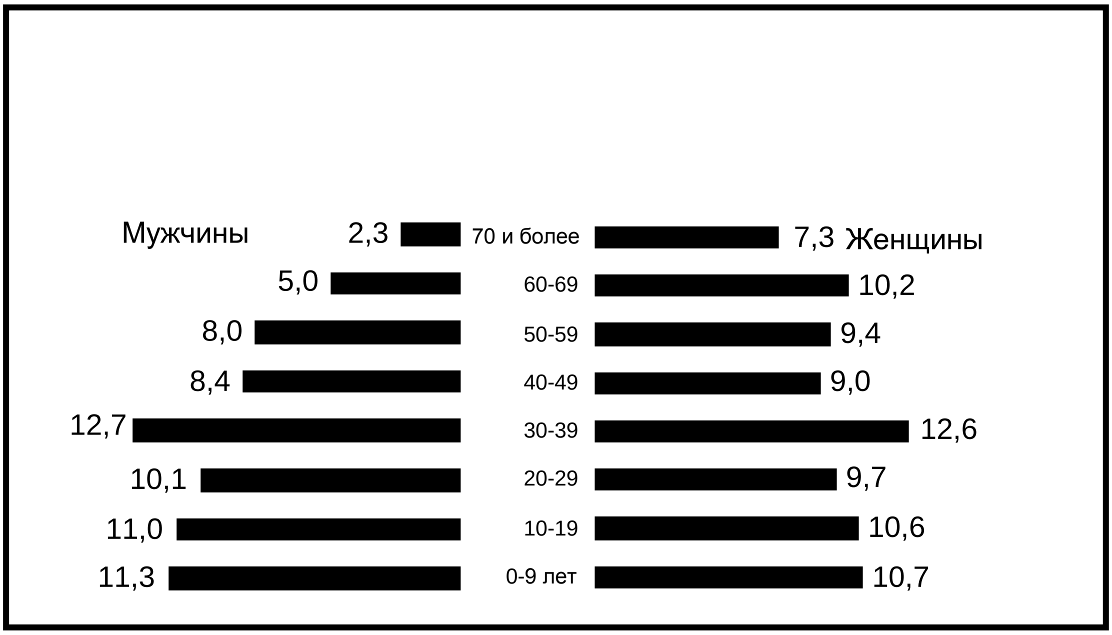


Рис. 4.4. Структура населения РФ по полу и возрасту

Структуру совокупности в сравнении можно отразить столбиковой диаграммой структуры (рис. 4.5), а структуру совокупности в динамике – полосовой диаграммой структуры (рис. 4.6).

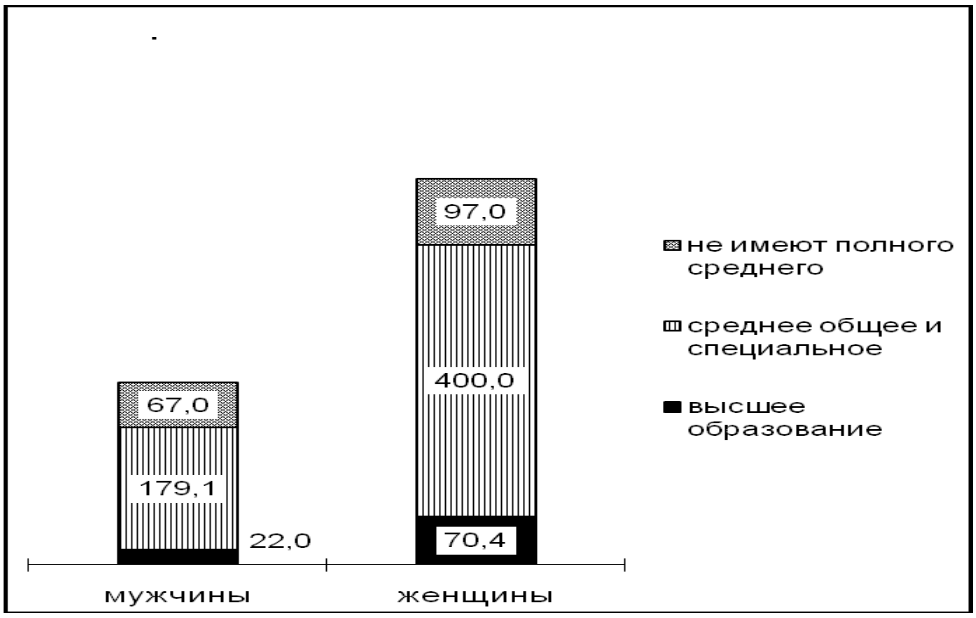


Рис. 4.5. Распределение безработных в РФ по полу и образованию, тыс. чел.

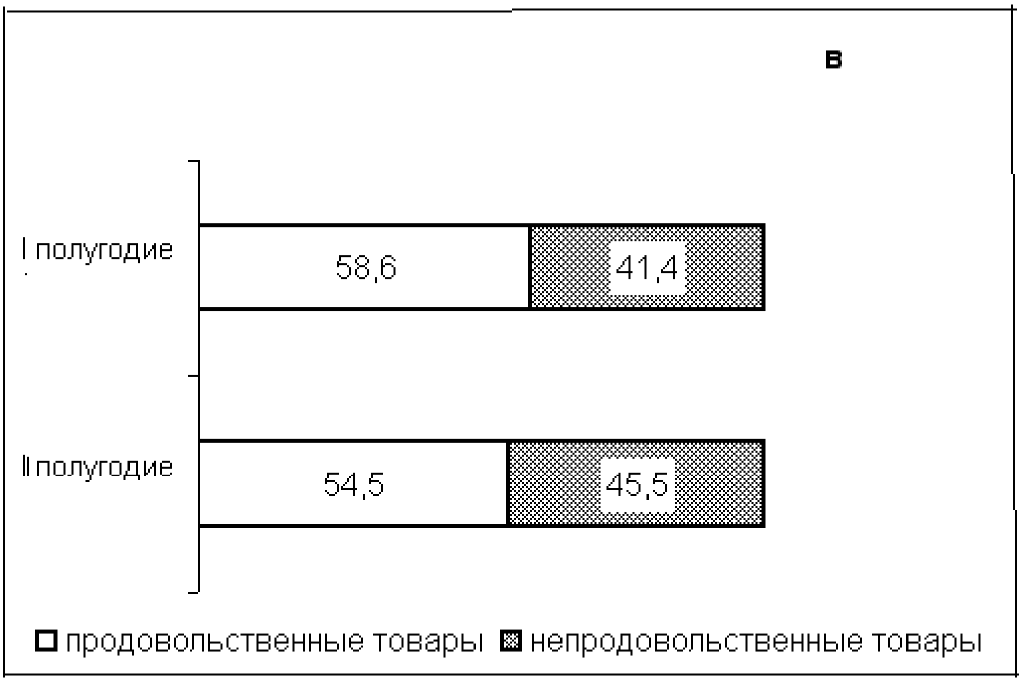


Рис. 4.6. Структура розничного товарооборота в РТ, %

4.5.2. Плоскостные диаграммы

Плоскостные диаграммы выражают величину отображаемого явления *размером площади* соответствующей геометрической фигуры (квадрата, прямоугольника, треугольника, круга и т.д.). С помощью плоскостных диаграмм можно производить сравнение явлений в пространстве и в динамике, характеризовать структуру явления.

Построение квадратных плоскостных диаграмм предусматривает расчет *стороны квадрата*. Она равна *корню квадратному из величины отображаемого признака* (рис. 4.7).

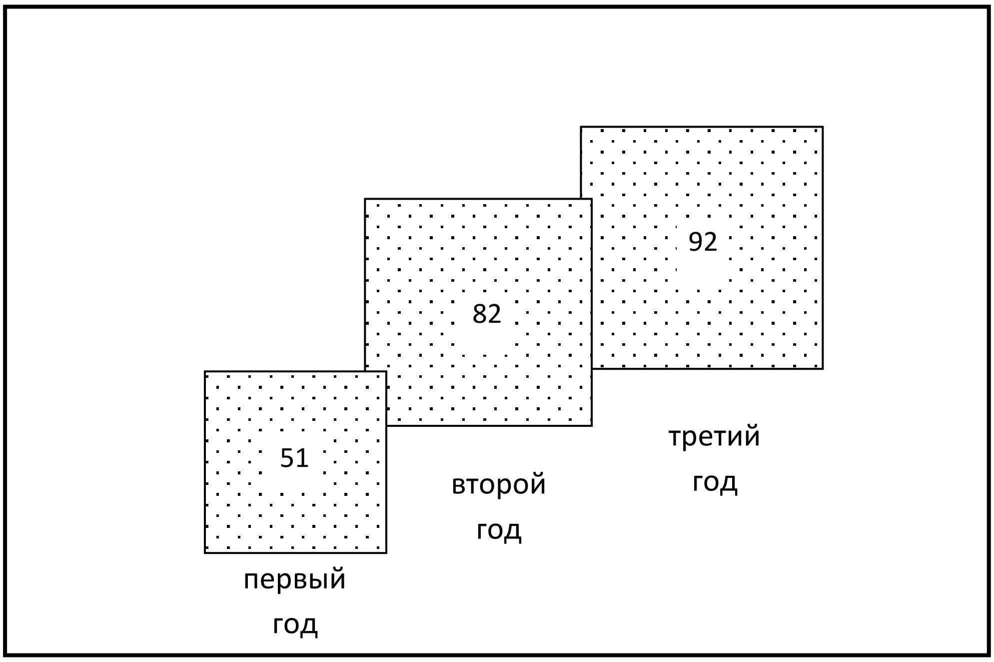


Рис. 4.7. Среднедушевое потребление овощей и бахчевых в РТ, кг

В круговых диаграммах величина признака отображается *площадью круга.* Для построения круговой диаграммы необходимо рассчитать *величину радиуса круга*, как *корень квадратный из величины признака* (рис.4.8).

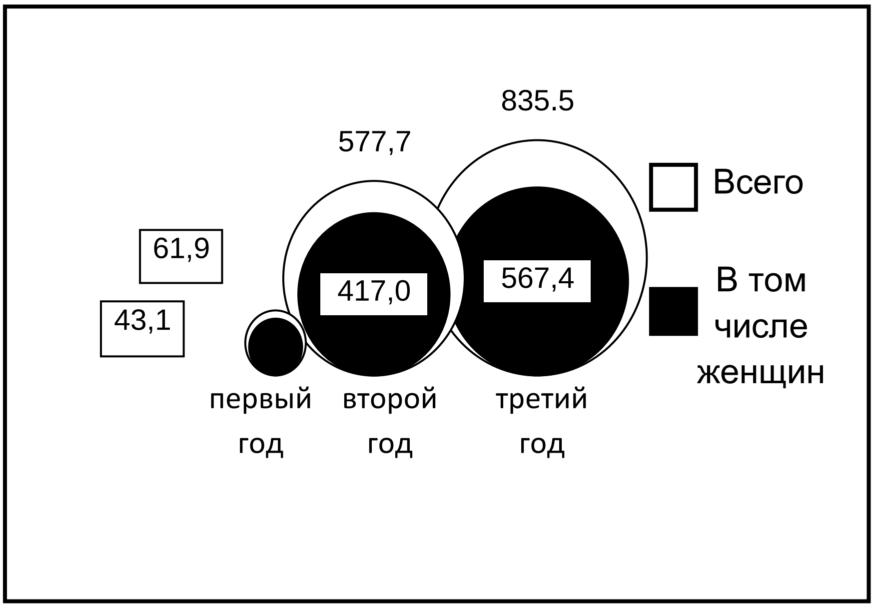


Рис. 4.8. Динамика численности безработных
в Российской Федерации на конец года, тыс. чел.

Для построения *круговой секторной диаграммы* в статистике чертят круг произвольного радиуса. Затем рассчитывают величину каждого сектора.

Если отображаемый признак выражен в процентах, то весь круг принимают за 100 %. Тогда каждый сектор рассчитывается исходя из соотношения $360° = 100\%$, а каждому проценту соответствует $3{,}6°$.

Если отображаемый признак выражен абсолютными величинами, то при расчете величины сектора $360°$ приравнивают всей величине признака в абсолютных единицах (рис. 4.9).

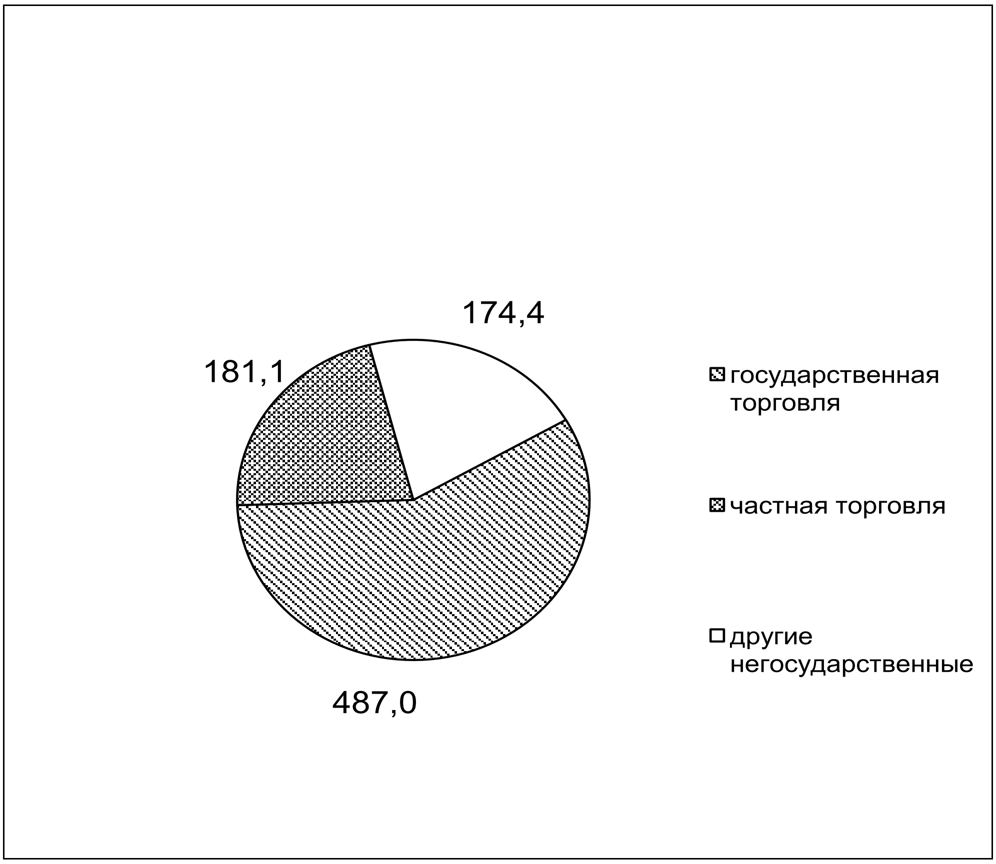


Рис. 4.9. Структура товарооборота РТ
по формам собственности за I полугодие, млрд. руб.

При построении секторной диаграммы структуры (в динамике рис. 4.10 или в сравнении) с отражением изменений только структуры явления круги чертят произвольным радиусом, а величину секторов рассчитывают как в графике на рис. 4.9.

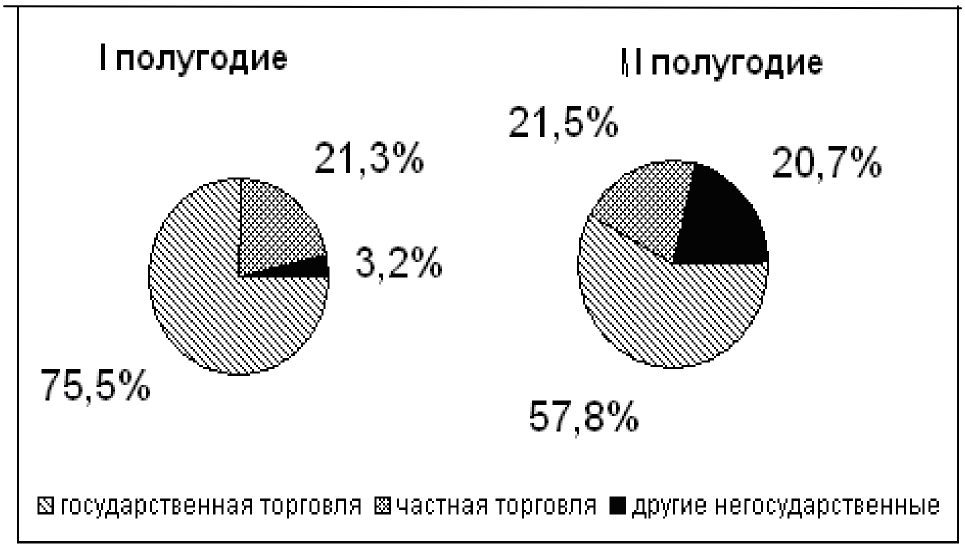


Рис.4.10. Динамика структуры товарооборота
по формам собственности (в %)

Рис. 4.11. Объем и структура розничного товарооборота, включая общественное питание, в РТ и Чувашской Республике по формам собственности за I полугодие (в %)

Знаки Варзара являются разновидностью плоскостных диаграмм. Они позволяют отразить на графике совокупность по признакам, произведение двух из которых имеет определенный экономический смысл. Например, произведение урожайности и посевной площади дает валовой сбор. Поэтому, если в прямоугольнике одну сторону брать пропорциональной урожайности, а другую – посевной площади, то площадь прямоугольника будет пропорциональна валовому сбору (рис.4.12)

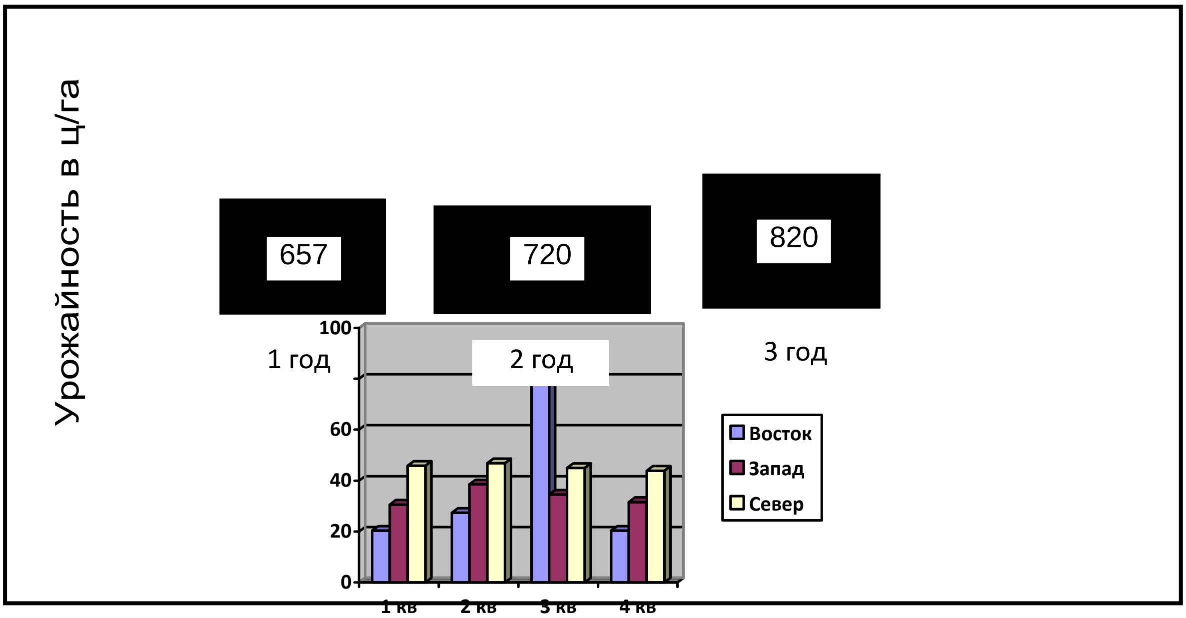
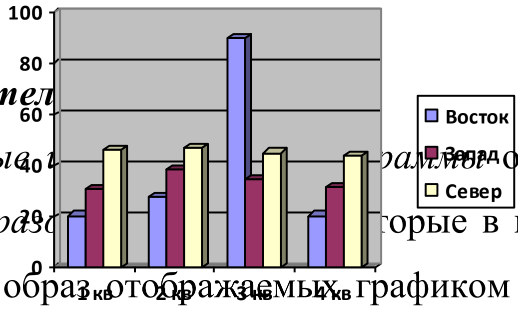


Рис. 4.12. Динамика валового сбора чая в Индии, тыс. тонн



4.5.3. Изобразител

Изобразительные ... *аммы* отображают явления с помощью *фигур - образ*... ...торые в какой-то степени воспроизводят внешний образ отображаемых графиком статистических совокупностей. Для этого используют *знаки-символы* (рис.4.13) и *фигуры – масштабные* знаки (4.14).

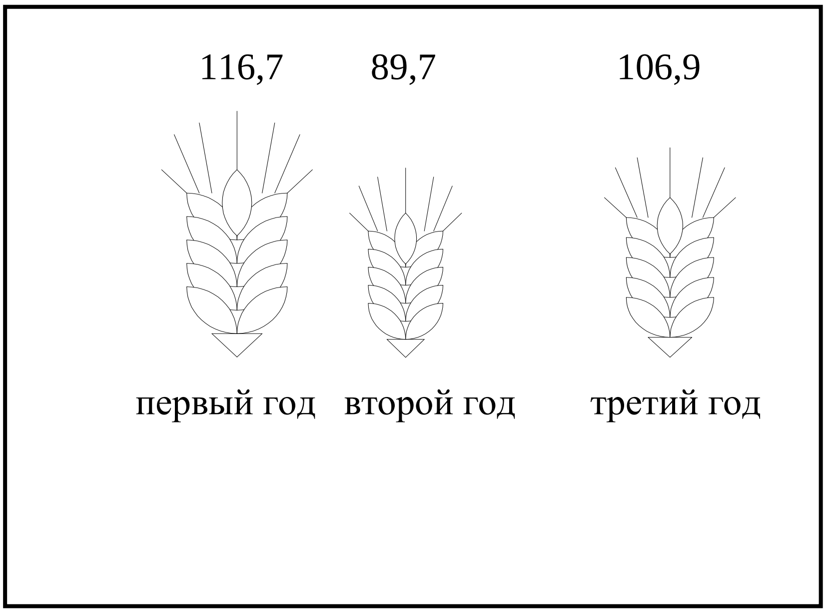


Рис 4.13. Валовой сбор зерна в России (в среднем за год), млн. тонн

Изобразительные диаграммы с использованием фигур – масштабных *знаков* более точно отражают величину признака. В этом случае фигуре-символу условно придается определенное численное значение, а самому знаку определенная стандартная ширина. Из этих знаков комплектуется полоса определенной длины, *т.е. полосовая диаграмма, состоящая из знаков-звеньев.*

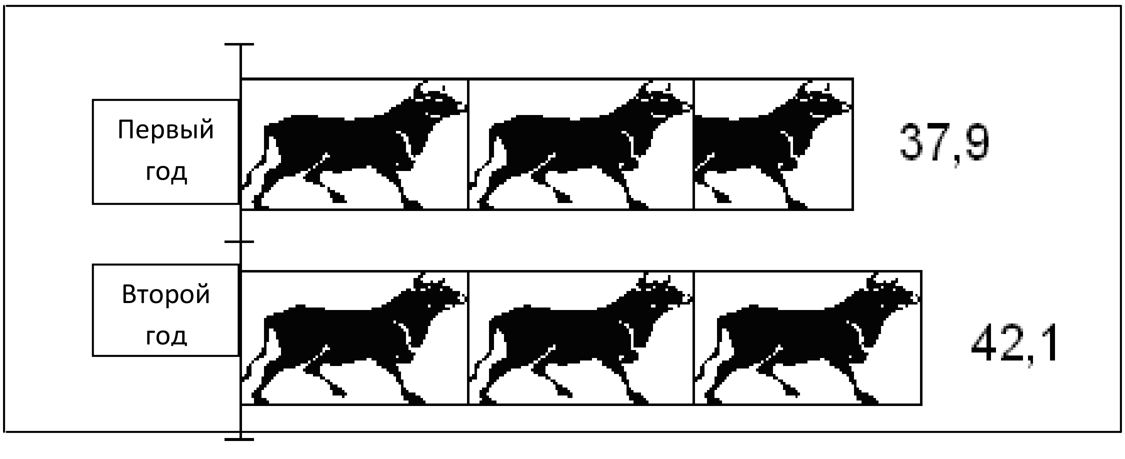


Рис. 4.14. Численность крупного рогатого скота
в общественных хозяйствах России на 1 июля, млн. голов

4.5.4. Линейные графики

Линейные графики изображают изменяющиеся во времени процессы *в виде ломаных линий.* Геометрическими знаками-символами в таких графиках являются точки и последовательно соединяющие их линии, которые складываются в ломаные кривые. Ориентировка точек и соединяющих их отрезков задается осями координат. Ось абсцисс является осью времени, а ось ординат – осью значений признака. Поскольку в линейном графике за аргумент принимается время, его можно считать геометрической формой динамического графика отображающего изменение экономических явлений как функцию времени.

Конфигурация каждой кривой в линейном графике определяется ее изломами, т.е. углами наклона ее отрезков в отношении горизонтальной оси и принятым масштабом.

Линейные графики имеют много разновидностей. Они отличаются по строению координатных сеток (прямоугольной, полулогарифмической и полярных координат).

Одним из важнейших условий построения линейных графиков в прямоугольной системе координат является выбор формата графической части, который определяется масштабами обеих шкал. Масштабы оси абсцисс - оси времени и оси ординат - оси показателей выбираются независимо и в разных единицах измерения, соотношение их между собой тоже не задано. Однако от их выбора зависит зрительное впечатление изображаемой кривой (рис.4.15).

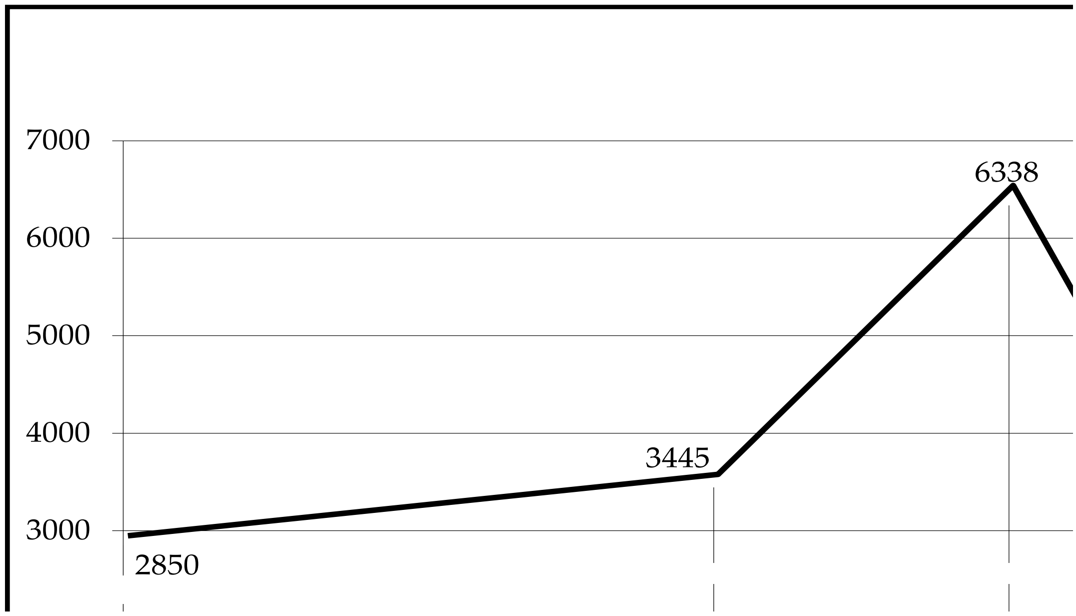


Рис. 4.15. Динамика численности населения РТ, тыс. чел.

Графики, отражающие двусторонние процессы (*балансовые)*: эмиграцию и иммиграцию, экспорт и импорт, обороты по платежному балансу, кассовые операции и т. д. *наносят по обе стороны от нулевой линии* (рис. 4.16).

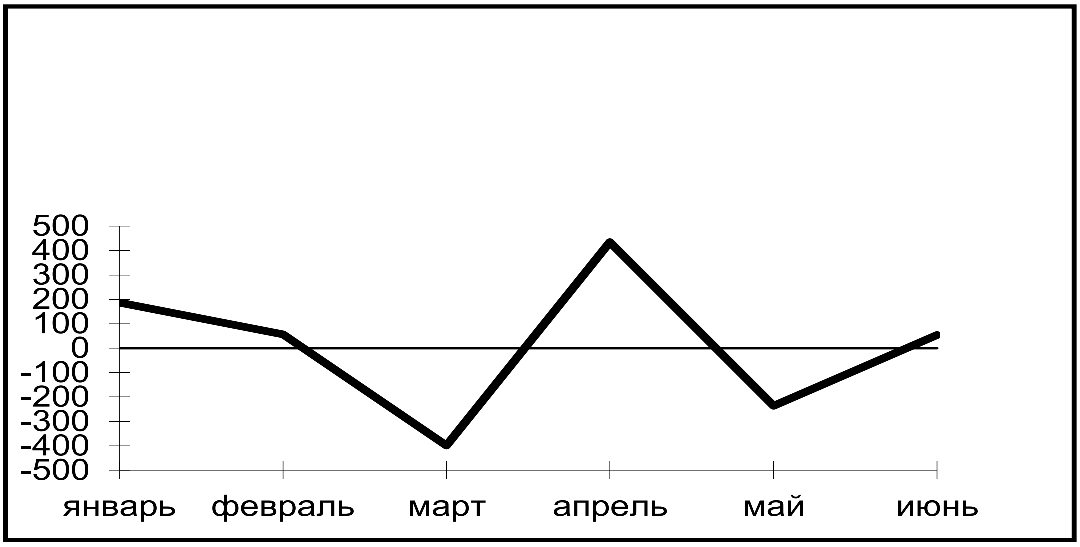


Рис. 4.16. Сальдо консолидированного бюджета России в I полугодии, млрд. руб.

Спиральные (радиальные) диаграммы предназначены для отражения явлений периодически повторяющихся во времени. Например, сезонных колебаний. Для построения спиральной диаграммы берется круг произвольного радиуса. Окружность разбивается на части по числу периодов. Например, 12 месяцев. Радиус круга за каждый период определяет величину явления (рис. 4.17).

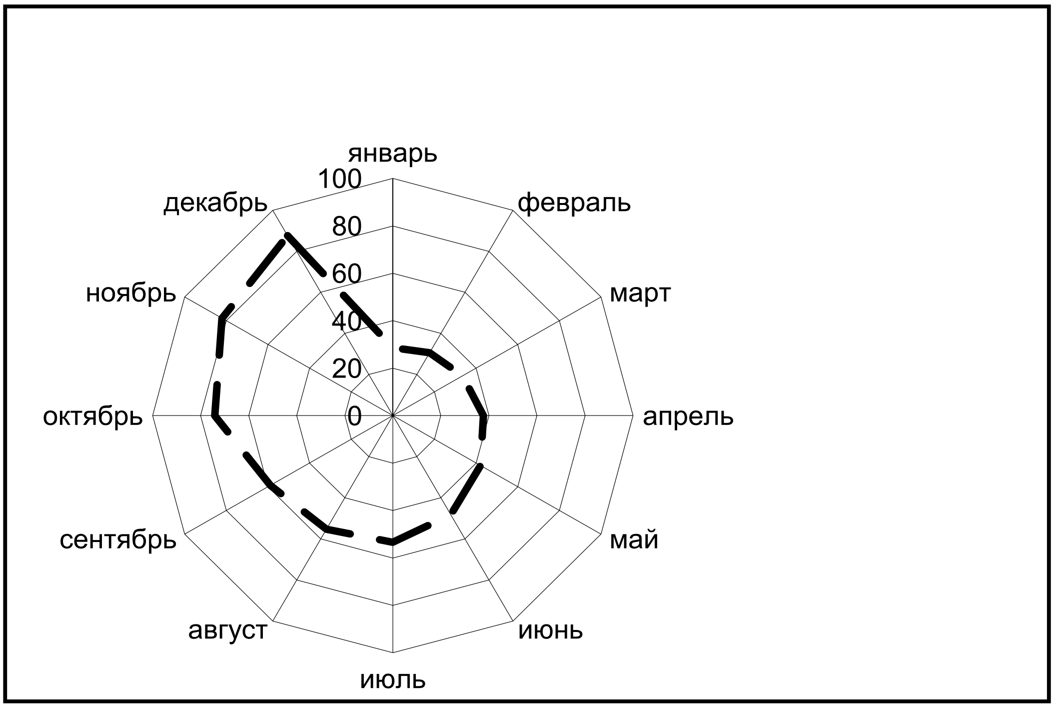


Рис. 4.17. Производство валового внутреннего продукта РФ (в текущих ценах)

Контрольно-плановые графики имеют несколько разновидностей. Простейшим из них является график, в котором используются две линии плановая и фактическая. При необходимости наглядного контроля выполнения плана одновременно по нескольким объектам строят контрольно-плановые графики на специальной сетке, имеющей форму таблицы (рис. 4.18).

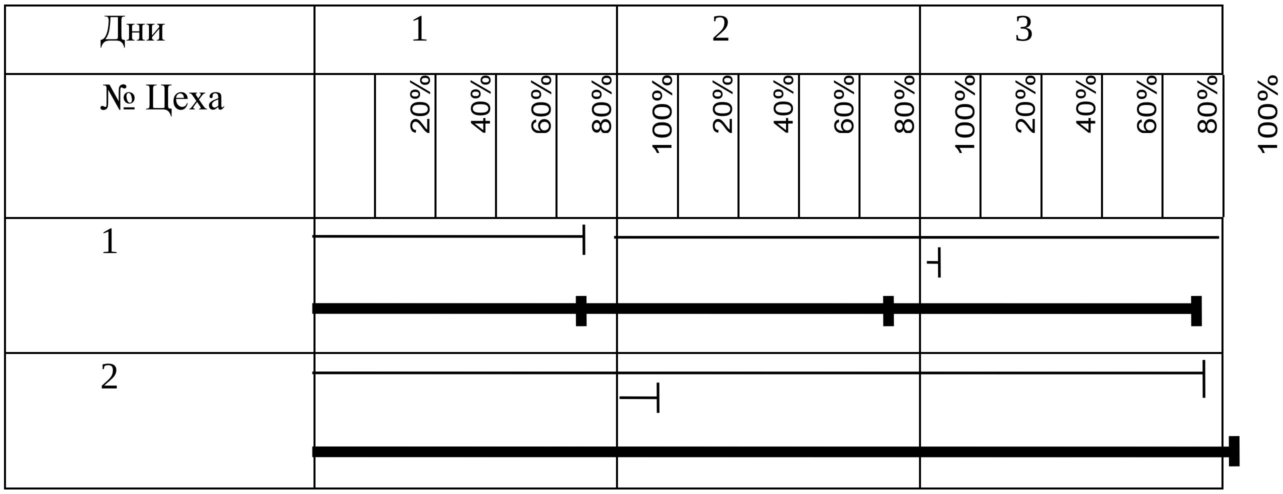


Рис. 4.18. Выполнение плана производства продукции
по подразделениям предприятия

4.5.5. Аналитические графики

Применяемые в математико-статистическом анализе *аналитические графики* представляют собой одну из форм математического моделирования. Из множества графиков этого типа рассмотрим *графики рядов распределения и графики для изучения взаимосвязей.*

Для графического изображения связи между двумя взаимосвязанными признаками в прямоугольной системе координат на оси абсцисс располагают *факторный (независимый) признак* «X», а на оси ординат – *результативный (зависимый) признак* «Y». Соединив точки пересечения прямыми, получаем ломаную линию регрессии (рис. 4.19).

С помощью графического метода в анализе взаимосвязей определяется не только направление связей, но и ее формы (прямолинейная или криволинейная), что способствует решению одной из трудных задач корреляционного анализа.

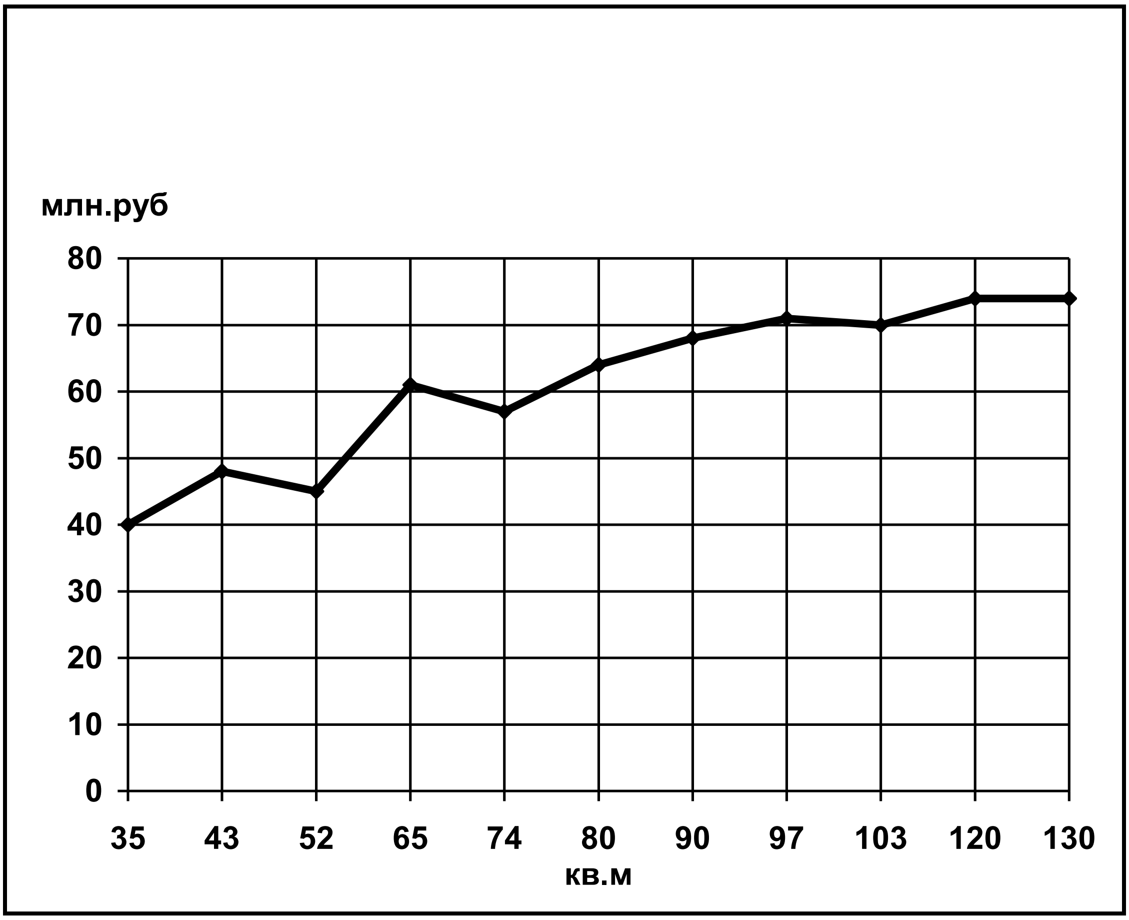


Рис. 4.19. Зависимость товарооборота на одного работника от площади торгового зала магазина

Полигоном чаще изображаются *дискретные ряды* распределения (рис. 4.20). В прямоугольной системе координат на оси абсцисс наносятся значения варьирующего признака, а на оси ординат – соответствующие им частоты (частости). Затем полученные точки соединить отрезками прямой. Полученный многоугольник распределения и будет полигоном распределения.

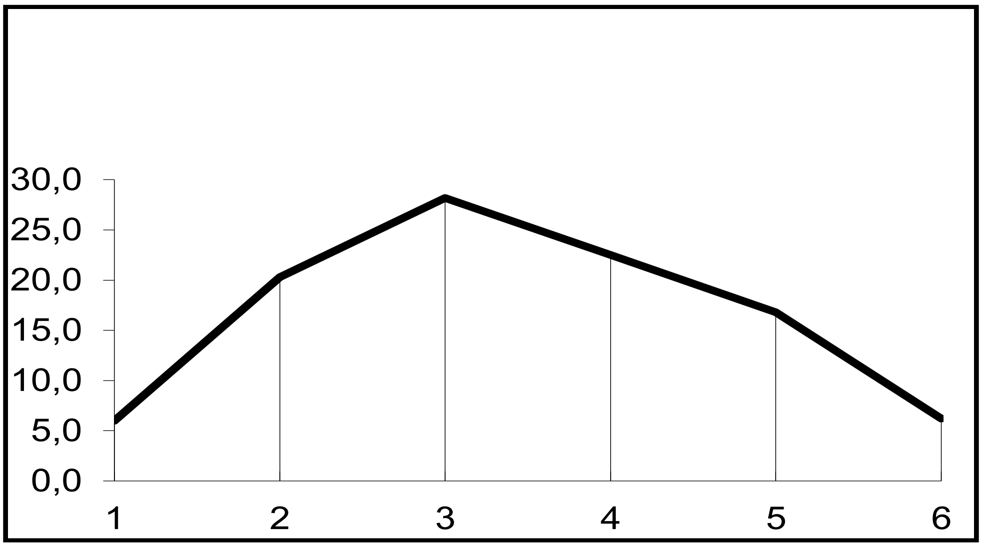


Рис. 4.20. Полигон распределения рабочих промышленности по тарифным разрядам

Гистограммой изображаются *интервальные ряды* распределения.

Строится гистограмма в прямоугольной системе координат. По оси абсцисс откладываются отрезки, изображающие интервалы значений варьирующего признака. На этих отрезках, как на основаниях, строят прямоугольники, высота которых при равных интервалах соответствует частотам, при неравных интервалах – плотностям распределения соответствующих интервалов. Получается ступенчатая фигура в виде сдвинутых друг к другу прямоугольников, площади которых пропорциональны частотам (частостям) (рис.4.21).

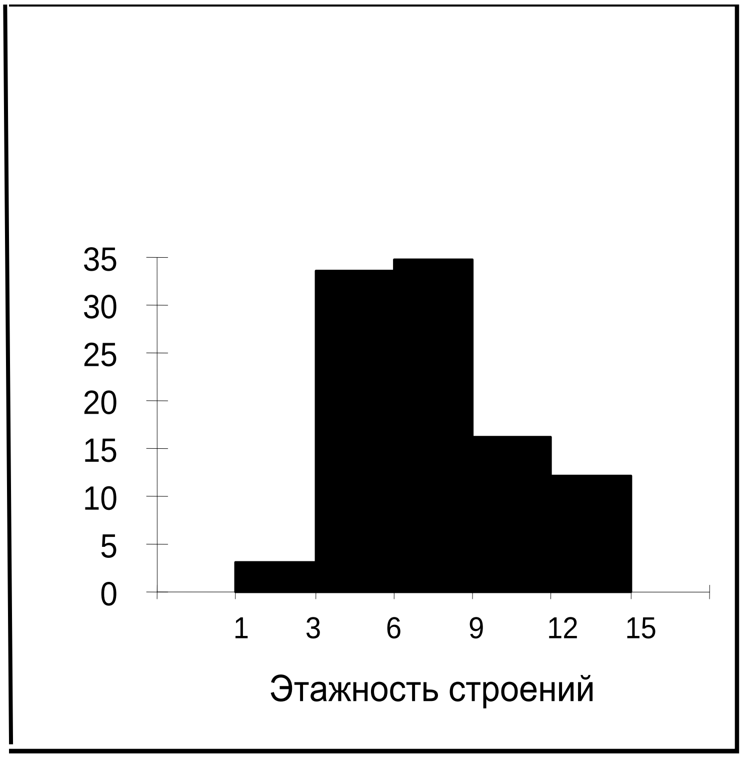


Рис. 4.21. Гистограмма распределения жилищного фонда Москвы по этажности строений на начало года, %

Кумулятой изображаются *кумулятивные ряды распределения*, где по оси абсцисс откладываются варианты ряда, а на оси ординат – накопленные частоты, показывающие, сколько единиц совокупности имеют значения признака, не превосходящие данное значение (рис. 4.22).

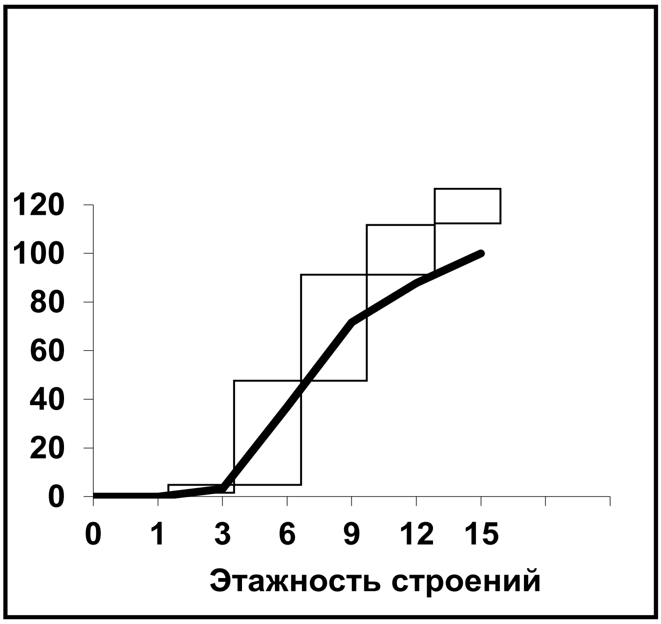


Рис. 4.22. Кумулята распределения жилищного фонда Москвы
по этажности строений на начало года, %

Такое изображение удобно при сравнении различных статистических рядов, а также в экономических исследованиях, в частности, для анализа концентрации производства.

4.5.6. Статистические карты

Картограмма показывает среднюю интенсивность развития, какого-либо явления по отдельным районам, странам и т.д.

Для наглядности каждую территориальную единицу раскрашивают или штрихуют в соответствии с интенсивностью явления.

Среди картограмм выделяют *фоновые* и *точечные* картограммы.

Фоновые картограммы отображают интенсивность явления штриховкой (рис. 4.23). Чем больше интенсивность явления, тем гуще штриховка и наоборот. Так, если по отношению к плотности населения России плотность населения в Белоруссии больше в 5,5 раз, на Украине в 9,5 раз, Молдавии в 14 раз. Следовательно, и плотность штриховки должна быть в таком же соотношении.

Рис. 4.23. Плотность населения

Точечные картограммы отображают явления числом точек. Чаще всего точечные картограммы используют для характеристики абсолютных показателей: численность населения, производство продукции и т.п. При построении точечной картограммы точки располагаются равномерно на территории (при географическом аспекте в действительном месторасположении объекта), определяется численное значение одной точки и величина её диаметра (рис. 4.24).

Рис.4.24. Численность населения

Картодиаграмма – это карта, показывающая с помощью диаграммной фигуры суммарную величину (иногда структуру и динамику) какого-либо статистического показателя в пределах территориального деления (рис. 4.25).

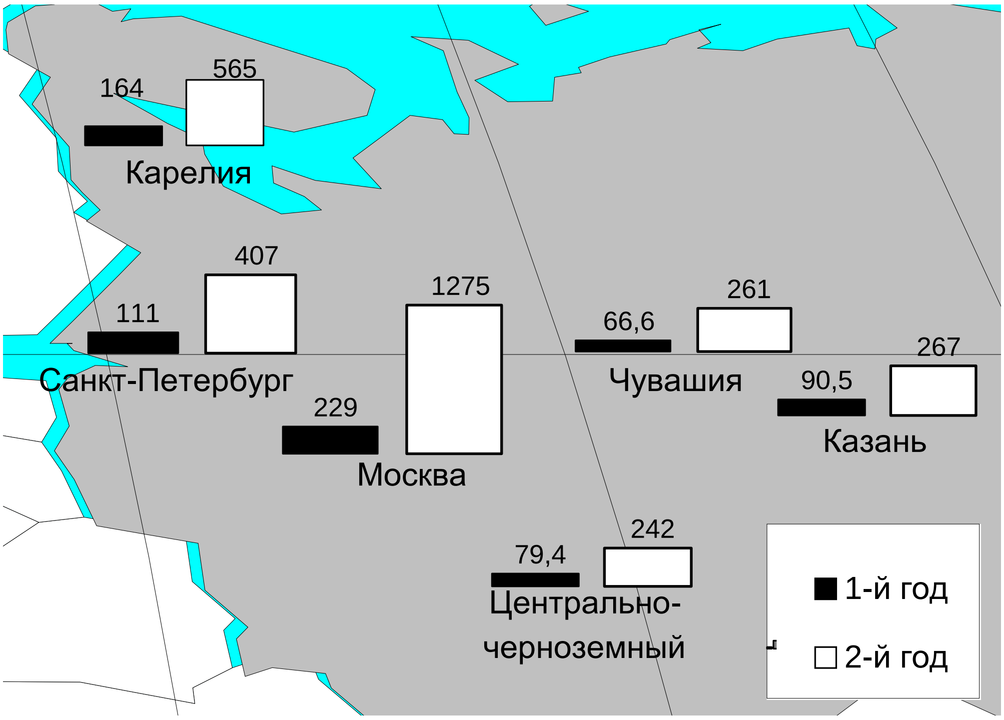


Рис. 4.25. Денежные доходны населения на душу населения, тыс. руб.

Центрограммы или историко-географические карты отображают долговременные динамические процессы в историко-географическом разрезе. Они бывают либо в виде *статистических таблиц-карт*, либо в виде *центрограмм перемещения* (рис. 4.26).

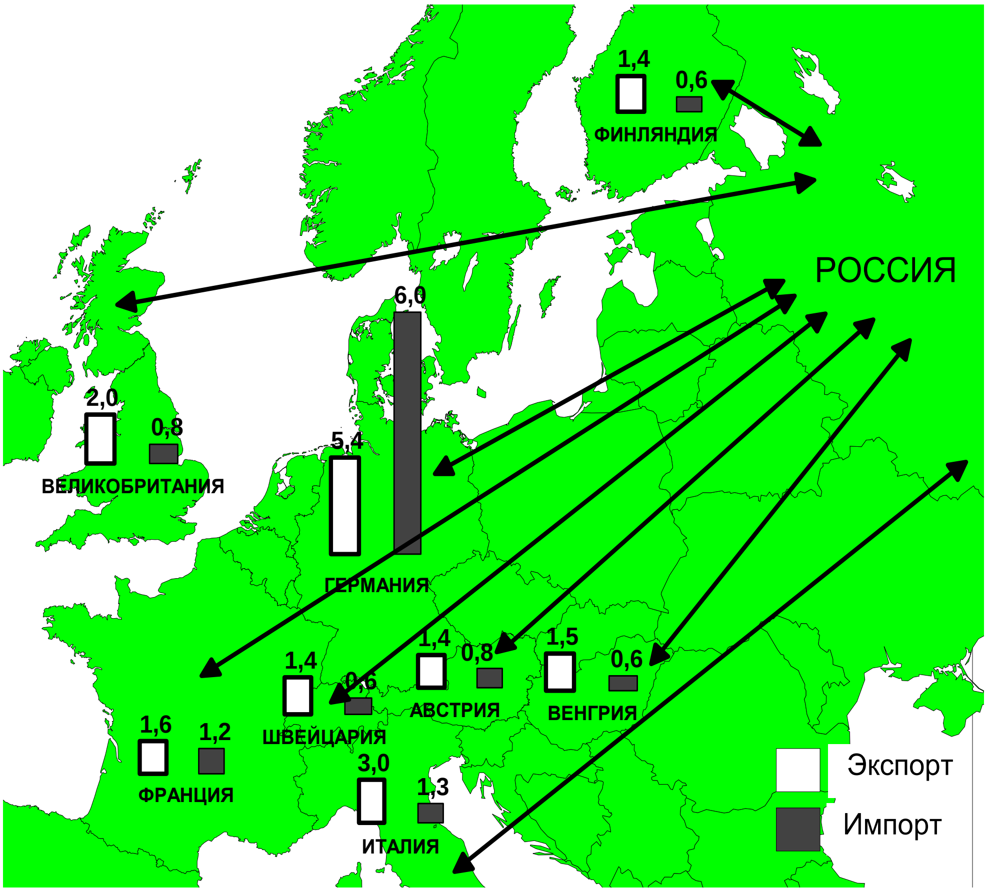


Рис. 4.26. Внешнеторговый оборот России, млрд. долл.

Для наглядного изображения статистических данных часто используются различные комбинации статистических графиков. Несколько разновидностей таких комбинаций графиков приведены в приложении 1.

ГЛАВА 5. АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

5.1. Абсолютные статистические величины

Исходной, первичной формой выражения статистических показателей являются абсолютные величины. *Абсолютные величины* характеризуют размер явлений в мерах массы, площади, объема, протяженности, времени и т.д. *Индивидуальные абсолютные показатели* получаются, как правило, непосредственно в процессе наблюдения в результате замера, взвешивания, подсчета, оценки. В некоторых случаях абсолютные индивидуальные показатели представляют собой разность. *Сводные, итоговые объемные абсолютные показатели* получают в результате сводки и группировки.

Абсолютные статистические показатели всегда являются числами именованными, т.е. *имеют единицы измерения*. Существует 3 типа единиц измерения абсолютных величин: натуральные, трудовые и стоимостные.

Натуральные единицы измерения – выражают величину явления в физических мерах, т.е. мерах веса, объема, протяженности, времени, счета. Например, в килограммах, кубических метрах, километрах, часах, штуках и т.д.

Разновидностью натуральных единиц *являются условно-натуральные единицы измерения,* которые используются для сведения воедино нескольких разновидностей одной и той же потребительной стоимости. Одну из них принимают за эталон, а другие пересчитываются с помощью специальных коэффициентов в единицы меры этого эталона. Так, например, мыло с разным содержанием жирных кислот пересчитывают на 40 % содержание жирных кислот. В отдельных случаях для характеристики какого-либо явления одной единицы измерения недостаточно, и используется произведение двух единиц измерения. Примером может служить грузооборот в тонно-километрах, производство электроэнергии в киловатт-часах и др.

В условиях рыночной экономики наибольшее значение имеют *стои-*

мостные (денежные) единицы измерения (рубль, доллар, марка и т.д.). Они позволяют получить денежную оценку любых социально-экономических явлений (объем продукции, товарооборота, национального дохода и т.п.). Однако следует помнить, что в условиях высоких темпов инфляции показатели в денежной оценке становятся несопоставимыми. Это следует учитывать при анализе стоимостных показателей в динамике. Для достижения сопоставимости показатели необходимо пересчитывать в сопоставимые цены.

Трудовые единицы измерения (человеко-часы, человеко-дни) используются для определения затрат труда на производство продукции, на выполнение какой-нибудь работы и т.п.

5.2. Относительные статистические величины, их сущность и формы выражения

Относительными величинами в статистике называются величины, выражающие количественное соотношение между явлениями общественной жизни. Они получаются в результате деления одной величины на другую. Величина, с которой производится сравнение (знаменатель), называется основанием, *базой сравнения*. Величина, которая сравнивается, (числитель) – называется, *сравниваемой, отчетной или текущей* величиной.

Относительная величина показывает, во сколько раз сравниваемая величина больше или меньше базисной, или какую долю, первая составляет от второй; а в отдельных случаях – сколько единиц одной величины приходится на единицу (или на 100, на 1000 и т.д.) другой (базисной) величины.

В результате сопоставления одноименных абсолютных величин получаются отвлеченные неименованные относительные величины, показывающие во сколько раз данная величина больше или меньше базисной. В этом случае базисная величина принимается за единицу (в результате получается *коэффициент)*. Кроме коэффициента широко распространенной формой выражения относительных величин являются *проценты «%»*. В этом случае базисная величина принимается за 100 единиц.

Относительные величины могут выражаться в *промилле (‰)*,в *продецимилле* ($^0/_{000}$). В этих случаях база сравнения принимается соответственно за 1000 и за 10000. В отдельных случаях база сравнения может быть принята и за 100000.

Относительные величины *могут быть числами именованными*. Ее наименование представляет собой сочетание наименований сравниваемого и базисного показателей. Например, плотность населения чел/кв.км (сколько человек приходится на 1 квадратный километр).

5.3. Виды относительных величин

Виды относительных величин подразделяются в зависимости от их содержания. Это относительные величины: планового задания, выполнения плана, динамики, структуры, координации, интенсивности, уровня экономического развития и сравнения.

Относительная величина *планового задания* представляет собой отношение величины показателя, устанавливаемой на планируемый период к величине его, достигнутой к планируемому периоду.

Относительной величиной *выполнения плана* называется величина, выражающая соотношение между фактическим и плановым уровнем показателя.

Относительная величина *динамики* представляет собой отношение уровня показателя за данный период к уровню этого же показателя в прошлом.

Три вышеперечисленные относительные величины связаны между собой, а именно: относительная величина динамики равна произведению относительных величин планового задания и выполнения плана.

Относительная величина *структуры* представляет собой отношение размеров части к целому. Она характеризует структуру, состав той или иной совокупности. Например, состав населения по полу.

$$Доля\ \ женщин = \frac{численность\ \ женщин}{все\ \ население}$$

$$Доля\ \ мужчин = \frac{численность\ \ мужчин}{все\ \ население}$$

Эти же величины в процентах называют *удельным весом.*

Относительной величиной *координации* называют соотношение частей целого между собой. В результате получают, во сколько раз данная часть больше базисной. Или сколько процентов от нее составляет или сколько единиц данной структурной части приходится на 1 единицу (100 или 1000 и т.д. единиц) базисной структурной части. Например, на 100 родившихся девочек приходится 105 родившихся мальчиков ($\frac{родившиеся\ \ девочки}{родившиеся\ \ мальчики} \cdot 100$).

Относительная величина *интенсивности* характеризует развитие изучаемого явления или процесса в другой среде. Это отношение двух взаимосвязанных явлений, но разных. Оно может быть выражено и в процентах, и в промилле, и продецимилле, и именованной. Например, число вакансий на 100 незанятых граждан = $\frac{число\ вакансий}{число\ незанятых} \cdot 100$ или коэффициент рождаемости в ‰ = $\frac{число\ родившихся\ \ за\ период}{средняя\ численность\ населения\ \ за\ период} \cdot 1000$, или плотность населения = $\frac{численность\ \ населения, чел.}{территория, кв.км.} (чел / кв.км.)$.

Разновидностью относительной величины интенсивности является показатель *уровня экономического развития*, характеризующий производство продукции на душу населения. Например, производство мяса на душу населения = $\frac{производство\ мяса\ за\ период, кг.}{средняя\ численность\ населения\ \ за\ период, чел}$.

Относительная величина *сравнения* представляет собой соотношение одноименных абсолютных показателей по разным объектам (предприятиям, районам, областям, странам и т.д.). Она может быть выражена как в коэффициентах, так и в процентах. Рассчитывая относительные величины сравнения, следует обращать внимание на сопоставимость сравниваемых показателей с позиций методологии их исчисления, поскольку по целому ряду показателей методы их исчисления в разных странах или в разные периоды времени неодинаковы. Поэтому прежде чем рассчитывать относительные показатели сравнения, приходится решать задачу пересчета сравниваемых показателей по единой методологии.

5.4. Сущность и виды средних величин

Статистика, как известно, изучает массовые социально-экономические явления. Каждое из этих явлений может иметь различное количественное выражение одного и того же признака. Например, заработная плата одной и той же профессии рабочих или цены на рынке на один и тот же товар и т.д.

Для изучения какой-либо совокупности по варьирующим (количественно изменяющимся) признакам статистика использует средние величины.

Средняя величина – это обобщающая количественная характеристика совокупности однотипных явлений *по одному* варьирующему признаку. Важнейшее свойство средней величины заключается в том, что она представляет значение определенного признака во всей совокупности одним числом, несмотря на количественные различия его у отдельных единиц совокупности, и выражает то общее, что присуще всем единицам изучаемой совокупности. Таким образом, через характеристику единицы совокупности она характеризует всю совокупность в целом.

Средние величины связаны с законом больших чисел. Суть этой связи заключается в том, что при осреднении случайные отклонения индивидуальных величин в силу действия закона больших чисел взаимопогашаются и в средней выявляется основная тенденция развития, необходимость, закономерность. Однако для этого среднюю необходимо вычислять на основе обобщения *массы* фактов.

Средние величины позволяют сравнивать показатели, относящиеся к совокупностям с различной численностью единиц.

Важнейшим условием научного использования средних величин в статистическом анализе общественных явлений является *однородность* совокупности, для которой исчисляется средняя. Одинаковая по форме и технике вычисления средняя в одних условиях (для неоднородной совокупности) фиктивная, а в других (для однородной совокупности) соответствует действительности. Качественная однородность совокупности определяется на основе всестороннего теоретического анализа сущности явления. Так, например, при исчислении средней урожайности требуется, чтобы исходные данные относились к одной и той же культуре (средняя урожайность пшеницы) или группе культур (средняя урожайность зерновых). Нельзя вычислять среднюю для разнородных культур.

Математические приемы, используемые в различных разделах статистики, непосредственно связаны с вычислением средних величин.

Средние в общественных явлениях обладают относительным постоянством, т.е. в течение какого-то определенного промежутка времени однотипные явления характеризуются примерно одинаковыми средними.

Средине величины очень тесно связаны с методом группировок, т.к. для характеристики явлений необходимо исчислять не только общие (для всего явления) средние, но и групповые (для типических групп этого явления по изучаемому признаку).

От того, в каком виде представлены исходные данные для расчета средней величины, зависит, по какой формуле она определяется. Рассмотрим наиболее часто применяемые в статистике виды средних величин:

- среднюю арифметическую;
- среднюю гармоническую;
- среднюю геометрическую;

- среднюю квадратическую.

Для этого введем следующие понятия и обозначения:

1. Признак, по которому находится средняя, называемый *осредняемым признаком,* обозначим буквой $"x"$.

2. Значения признака, которые встречаются у группы единиц или отдельных единиц совокупности (не повторяясь) называются *вариантами признака* и обозначаются через x_1, x_2, x_3 и т.д. *Средняя величина* этих значений обозначается через $\bar{x}$, численность вариантов признака – через n.

Средняя арифметическая величина может быть простой и взвешенной.

Средняя арифметическая простая рассчитывается по формуле:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{x_1 + x_2 + ... + x_n}{n},$$

т.е. как сумма вариантов признака, деленная на их число. Средняя арифметическая простая применяется в тех случаях, когда каждая варианта признака встречается в совокупности один или равное число раз.

Средняя арифметическая взвешенная вычисляется по формуле:

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i},$$

где f_i – частота повторения i*- ых* вариантов признака, называемая *весом.* Таким образом, средняя арифметическая величина взвешенная равна сумме взвешенных вариантов признака, деленная на сумму весов:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + ... + x_n f_n}{f_1 + f_2 + ... + f_n}$$

Она применяется в тех случаях, когда каждая варианта признака встречается несколько (неравное) число раз.

При расчете средней по *интервальному вариационному ряду* необходимо сначала найти *середину интервалов*. Это и будут значения x_i а количество единиц совокупности в каждой группе f_i .

Пример 5.1.

Необходимо определить средний возраст рабочих в цехе из 106 человек. Данные представлены в виде интервального ряда.

Таблица 5.1

Возраст рабочего, лет	Число рабочих, чел *(f_i)*	Середина возрастного интервала, лет (x_i)
А	1	2
20-30	7	25
30-40	13	35
40-50	48	45
50-60	32	55
60 и более	6	65
Итого	106	Х

Средний возраст рабочих цеха будет равен:

$$\bar{x} = \frac{25 \cdot 7 + 35 \cdot 13 + 45 \cdot 48 + 55 \cdot 32 + 65 \cdot 6}{106} = 47 \text{ лет.}$$

Средняя гармоническая величина является преобразованной средней арифметической величиной. Применяется она тогда, когда необходимые веса *(f_i)* в исходных данных не заданы непосредственно, а входят сомножителем в одни из имеющихся показателей. Она также может быть простой и взвешенной.

Средняя гармоническая простая рассчитывается по формуле:

$$\bar{x} = \frac{n}{\Sigma \frac{1}{x_i}},$$

т.е. это обратная величина средней арифметической простой из обратных значений признака.

Формула *средней гармонической взвешенной:*

$$\bar{x} = \frac{\Sigma M_i}{\Sigma \frac{M_i}{x_i}},$$

где $M_i = x_i \cdot f_i$ (по содержанию).

Пример 5.2.

По одному из районов имеются данные о валовом сборе и урожайности технических культур. Необходимо определить среднюю урожайность всех технических культур на основании следующих данных:

Таблица 5.2

Культуры	Валовой сбор, ц .(M_i)	Урожайность, ц/га (x_i)
А	1	2
Хлопчатник	97,2	30,4
Сахарная свекла	601,2	467,0
Подсолнечник	46,3	11,0
Льноволокно	2,6	2,9
Итого	743,3	Х

Здесь в исходной информации веса (площадь под культурами) не заданы, но входят сомножителем в валовой сбор, равный урожайности, умноженной на площадь:

$M_i = x_i \cdot f_i$, поэтому $f_i = \frac{M_i}{x_i}$, а средняя урожайность будет равна:

$$\bar{x} = \frac{\Sigma M_i}{\Sigma \frac{M_i}{x_i}} = \frac{747.3}{\frac{97.2}{30.4} + \frac{601.2}{167.0} + \frac{46.3}{11.0} + \frac{2.6}{2.9}} = \frac{747.3ц}{11,9га} = 62,9ц / га$$

Средняя геометрическая также может быть простой и взвешенной.

Применяется она главным образом при нахождении средних коэффициентов роста.

Средняя геометрическая простая находится по формуле:

$$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} = \sqrt[n]{\Pi x_i}$$

Средняя геометрическая взвешенная – по формуле:

$$\bar{x} = \sqrt[\Sigma m]{x_1^{m_1} \cdot x_2^{m_2} \cdot \ldots \cdot x_n^{m_n}} = \sqrt[\Sigma m]{\Pi x_i^{m_i}}$$

Эту среднюю используют в основном для нахождения средних коэффициентов роста.

Средняя квадратическая применяется в тех случаях, когда приходится осереднять величины, входящие в исходную информацию в виде квадратических функций.

Простая средняя квадратическая:

$$\bar{x} = \sqrt{\frac{\Sigma x_i^2}{n}}$$

Взвешенная:

$$\bar{x} = \sqrt{\frac{x_i^2 f_i}{\Sigma f_i}}\ .$$

Наиболее широко этот вид средней используется при расчете показателей вариации.

5.5. Структурные средние

Для характеристики структуры вариационных рядов применяются так называемые структурные средние. Наиболее часто используются в экономической практике мода и медиана.

Мода – это наиболее часто встречающаяся варианта признака в данной совокупности.

В *дискретных вариационных рядах* мода определяется по наибольшей частоте.

Пример 5.3.

Предположим товар «А» реализуют в городе девять фирм по цене в рублях: 44; 43; 44; 45; 43; 46; 42; 46;43.

Так как чаще всего встречается цена 43 рубля, то она и будет модальной.

В *интервальных вариационных рядах* моду определяют приближенно по формуле:

$$M_0 = x_0 + i_{Mo} \cdot \frac{f_{Mo} - f_{Mo-1}}{f_{Mo} - f_{Mo-1} + f_{Mo} - f_{Mo+1}}, \text{ где:}$$

x_0 - нижняя граница модального интервала;

i_{Mo} - величина модального интервала;

f_{Mo} - частота модального интервала;

f_{Mo-1} - частота интервала, предшествующая модальному;

f_{Mo+1} - частота интервала, следующая за модальным.

Мода применяется для решения некоторых практических задач. Так, например, при изучении товарооборота рынка берется модальная цена, для изучения спроса на обувь, одежду используют модальные размеры обуви и одежды и др.

Пример 5.4.

Необходимо определить модальное значение среднедушевого месячного дохода по данным таблицы 5.3.

Таблица 5.3

Распределение населения по уровню среднедушевого месячного дохода

Среднедушевой месячный доход, тыс.руб.	Удельный вес населения, % (f_i)	Накопленная частота, % (S_i)
А	1	2
5-7	37,9	37,9
7-9	30,0	67,9
9-11	15,7	83,6
11-13	7,7	91,3
13 и выше	8,7	100,0
Всего	100,0	Х

Расположение модального интервала определяют по наибольшей частоте.

Интервал 5000-7000 в данном распределении будет модальным, т.к. он имеет наибольшую частоту ($f_{Mo} = 37{,}9$). Тогда по вышеуказанной формуле мода будет равна:

$$M_0 = 5 + 2\frac{37{,}9 - 0}{37{,}9 - 0 + 37{,}9 - 30{,}0} = 6{,}655 \text{ тыс. руб.}$$

Медиана – это численное значение признака у той единицы совокупности, которая находится в середине ранжированного ряда (построенного в порядке возрастания, либо убывания значения изучаемого признака). Медиану иногда называют *серединной* вариантой, т.к. она делит совокупность на две равные части. В *дискретных вариационных рядах* с нечетным числом единиц совокупности – это конкретное численное значение в середине ряда. Так в группе студентов из 27 человек медианным будет рост у 14-го, если они выстроятся по росту. Если число единиц совокупности четное, то медианой будет средняя арифметическая из значений признака у двух средних членов ряда. Так, если в группе 26 человек, то медианным будет рост средний 13-го и 14-го студентов. В интервальных вариационных рядах медиана определяется по формуле:

$$Me = x_0 + i_{Me} \cdot \frac{\frac{1}{2}\Sigma f_i - S_{Me-1}}{f_{Me}} \quad ,$$

где x_0 – нижняя граница медианного интервала;

i_{Me} – величина медианного интервала;

S_{Me-1} - сумма накопленных частот до медианного интервала;

f_{Me} - частота медианного интервала.

Пример 5.5.

По данным таблицы 5.3. определим медианное значение среднедушевого дохода. Для этого необходимо определить, какой интервал будет медианным. Используя формулу накопленной частоты до медианы, т.е. середины:

$$S_{Me} = \frac{\Sigma f + 1}{2} = \frac{100+1}{2} = 50.5 \ (\%)$$

Дробное значение S_{Me} (всегда при четном числе членов) равное 50,5 % говорит о том, что середина ряда находится между 50 % и 51 %, т.е. в третьем интервале. Отсюда медиана по формуле будет определена:

$$Me = 7 + 2 \cdot \frac{0{,}5 \cdot 100 - 37{,}9}{30} = 7{,}81 \text{ тыс. руб.}$$

Мода и медиана, как правило, отличаются от значения средней, совпадая с ней только в случае симметричного распределения частот вариационного ряда. Соотношение моды, медианы и средней арифметической указывает на характер распределения признака в совокупности, позволяет оценить его асимметрию. Если $Mo < Me < \bar{x}$ имеет место правосторонняя асимметрия. Если же $\bar{x} < Me < Mo$ – левосторонняя асимметрия ряда. По приведенному примеру можно сделать заключение, что наиболее распространенным является доход порядка 6,655 тыс. руб. в месяц. В то же время более половины населения располагают доходом свыше 7,81 тыс. руб., при среднем уровне 8,386 тыс. руб.:

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{838{,}6}{100} = 8{,}386 \text{ тыс. руб.}$$

Из соотношения этих показателей следует сделать вывод о правосторонней асимметрии распределения населения по уровню среднедушевого денежного дохода.

Аналогично медиане вычисляются значения признака, делящие совокупность на четыре равные (по числу единиц) части – *квартили,* на десять частей – *децили*, на сто частей – *перцентили.* Так формула первого квартиля будет:

$$Q_1 = x_{Q_1} + i_{Q_1} \cdot \frac{\frac{1}{4}\sum f - S_{Q_1-1}}{f_{Q_1}}.$$

Второй квартиль равен медиане. Формула третьего квартиля будет:

$$Q_3 = x_{Q_3} + i_{Q_3} \cdot \frac{\frac{3}{4}\sum f - S_{Q_3-1}}{f_{Q_{31}}}.$$

Аналогичны формулы децилей. Пятый дециль равен медиане.

Среди множества варьирующих признаков существуют признаки, которыми одни единицы совокупности обладают, а другие не обладают. Такие признаки называются *альтернативными.* Примером таких признаков являются: наличие бракованной продукции, ученая степень у преподавателя, наличие академической задолженности у студента и др. Обозначим: 1 – наличие интересующего нас признака; 0 – его отсутствие; p – доля единиц, обладающих данным признаком; q – доля единиц, не обладающих данным признаком; тогда $p + q = 1$.

Среднее значение альтернативного признака:

$$\bar{\bar{x}} = \frac{\sum xf}{\sum f} = \frac{1 \cdot p + 0 \cdot q}{p + q} = p, \text{ так как } p + q = 1.$$

Следовательно, средняя арифметическая величина альтернативного признака равна доле единиц обладающих признаком.

Пример 5.6.

Предположим в группе 10 % студентов имеют академическую задолженность по результатам сессии. Следовательно 90 % не имеют этой задолженности. Отсюда $p = 0{,}1$, а $q = 0{,}9$. Среднее же значение альтернативного признака – академическая задолженность – будет равно:

$$\bar{x} = \frac{1 \cdot 0{,}1 + 0 \cdot 0{,}9}{0{,}1 + 0{,}9} = 0{,}1$$

5.6. Свойства средней арифметической величины

Средняя арифметическая величина обладает рядом математических свойств, которые позволяют упростить ее вычисление.

Свойство 1. Сумма отклонений индивидуальных значений признака от средней арифметической равна нулю:

$$\sum (x - \bar{x}) = 0$$

Свойство 2. Если все индивидуальные значения признака (т.е. все варианты) уменьшить или увеличить в i раз, то среднее значение нового признака соответственно уменьшится или увеличится в i раз:

$$\frac{\sum \frac{x}{i} f}{\sum f} = \frac{\bar{x}}{i}$$

Свойство 3. Если все варианты осредняемого признака уменьшить или увеличить на число A, то средняя арифметическая соответственно уменьшится или увеличится на это же число A:

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

Свойство 4. Если веса всех осередняемых вариантов уменьшить или увеличить в k раз, то средняя арифметическая не изменится:

$$\frac{\sum x\frac{f}{k}}{\sum \frac{f}{k}} = \bar{x}$$

5.7. Упрощенный метод расчета средней арифметической величины

Для упрощения расчетов средней арифметической величины идут по пути уменьшения значений вариантов и частот. Используют, так называемый, *метод моментов* или *способ отсчета от условного нуля.* Этот способ может использоваться только в интервальных рядах *с равным интервалом*. Основан он на использовании свойств средней арифметической величины и предполагает следующие действия:

1. Если возможно, сокращаются все веса в k раз и получают значения:

$f' = \dfrac{f}{k}$ (свойство 4).

2. Находятся средние значения в каждой группе интервального ряда по простой средней арифметической величине как сумма крайних значений интервала деленная на 2. Для определения отсутствующей нижней границы первой группы следует из верхнего значения вычесть величину интервала, а для определения верхней границы последнего интервала к нижнему значению прибавить величину интервала.

3. Уменьшаются все варианты на A единиц. В качестве A выбирается значение одного из средних групповых вариантов (находящегося посередине ряда, либо обладающего наибольшей частотой (свойство 3).

4. Уменьшенные на A единиц варианты уменьшают в i раз. В качестве i берется величина равновеликого интервала (свойство 2).

5. Находится средняя арифметическая взвешенная величина из сокращенных вариантов и весов, т.е. определяется, так называемый момент

первой степени m_1 по формуле:

$$m_1 = \frac{\sum \frac{x-A}{i} \cdot f'}{\sum f'}$$ (если веса не имеют общего множителя, в формуле вместо f' будет f.

6. Находим среднюю арифметическую величину для всей совокупности по формуле:

$\bar{x} = m_1 \cdot i + A$ – это и будет формула средней *по способу моментов*.

Применение способа моментов настолько облегчает расчеты, что позволяет их выполнять без использования вычислительной техники даже при больших и многозначных числах, характеризующих индивидуальные значения осредняемых показателей.

Пример 5.7.

Рассмотрим нахождение среднего душевого месячного дохода методом моментов используя данные таблицы 5.3.

Таблица 5.3.1

Распределение населения по уровню среднедушевого месячного дохода

Среднедушевой месячный доход, тыс. руб.	Удельный вес населения, % (f_i)	Середина интервала x	$x-A$ $A=8$	$\frac{x-A}{i}$ $i=2$	$\frac{x-A}{i} \cdot f$	$\left(\frac{x-A}{i}\right)^2 \cdot f$
A	1	2	3	4	5	6
5-7	37,9	37,9	-2	-1	-37,9	37,9
7-9	30,0	67,9	0	0	0	0
9-11	15,7	83,6	2	1	15,7	15,7
11-13	7,7	91,3	4	2	15,4	30,8
13 и выше	8,7	100,0	6	3	26,1	78,3
Всего	100,0	X	X	X	19,3	162,7

Величина A выбирается из значений x. Наиболее рационально, если это будет значение находящееся в середине ряда. В данном случае это 8 тыс. руб. Тогда момент первой степени будет:

$$m_1 = \frac{\sum \frac{x - A}{i} \cdot f}{\sum f} = \frac{19,3}{100} = 0,193,$$

а средний доход:

$$\bar{x} = m_1 \cdot i + A = 0,193 \cdot 2 + 8 = 8,286\,\textit{тыс.руб}.$$

ГЛАВА 6. ВАРИАЦИОННЫЙ АНАЛИЗ

6.1. Понятие и сущность вариации

Вариация – это различие в значениях какого-либо признака у разных единиц данной совокупности в один и тот же период или момент времени. Например, работники фирмы различаются по доходам, затратам времени на работу, росту, весу, любимому занятию в свободное время и т.д. Она возникает в результате того, что индивидуальные значения признака складываются под совокупным влиянием разнообразных факторов (условий), которые по-разному сочетаются в каждом отдельном случае. Таким образом, величина каждого варианта объективна.

Исследование вариации в статистике имеет большое значение, помогает познать сущность изучаемого явления. Особенно актуально оно в период формирования многоукладной экономики. Измерение вариации, выяснение ее причины, выявление влияния отдельных факторов дает важную информацию (например, о продолжительности жизни людей, доходах и расходах населения, финансовом положении предприятия и т.п.) для принятия научно-обоснованных управленческих решений.

Средняя величина дает обобщающую характеристику признака изучаемой совокупности, но она не раскрывает строения совокупности, которое весьма существенно для ее познания. Средняя не показывает, как располагаются около нее варианты осредняемого признака, сосредоточены ли они вблизи средней или значительно отклоняются от нее. Средняя величина признака в двух совокупностях может быть одинаковой, но в одном случае все индивидуальные значения отличаются от нее мало, а в другом – эти отличия велики, т.е. в одном случае вариация признака мала, а в другом – велика, это имеет весьма важное значение для характеристики *надежности* средней величины.

Чем больше варианты отдельных единиц совокупности различаются между собой, тем больше они отличаются от своей средней, и наоборот, –

чем меньше варианты отличаются друг от друга, тем меньше они отличаются от средней, которая в таком случае будет более реально представлять всю совокупность. Следовательно, необходим расчет показателей вариации.

6.2. Абсолютные и относительные показатели вариации

К абсолютным показателям вариации показателям вариации относятся: *размах вариации, среднее линейное отклонение, дисперсия, среднее квадратическое отклонение.*

Самым элементарным показателем вариации признака является *размах вариации* R, представляющий собой разность между максимальным и минимальным значениями признака: $R = x_{\max} - x_{\min}$. Он имеет те же единицы измерения, что и осередняемый признак. Недостатком этого показателя является то, что он оценивает только границы варьирования признака и не отражает его колеблемость внутри этих границ.

Среднее линейное отклонение ($\overline{d}$) представляет собой среднюю арифметическую величину из абсолютных отклонений отдельных вариантов от средней арифметической, при этом всегда предполагают, что среднюю вычитают из варианта $(x - \overline{x})$. Среднее линейное отклонение, как и всякая средняя, может быть простым и взвешенным. Простое среднее линейное отклонение определяется по формуле:

$$\overline{d} = \frac{\sum |x - \overline{x}|}{n},$$ где n — число членов ряда.

Взвешенное среднее линейное отклонение определяется по формуле:

$$\overline{d} = \frac{\sum |x - \overline{x}| \cdot f}{\sum f},$$

где $\sum f$ — сумма частот вариационного ряда.

Среднее линейное отклонение как меру вариации признака применяют в статистической практике редко (только в тех случаях, когда суммирование показателей без учета знаков имеет экономический смысл). С его помощью, например, анализируется состав работающих, ритмичность производства, оборот внешней торговли. Среднее линейное отклонение измеряется в тех же единицах изменения, что и осередняемый признак и не может быть отрицательной величиной.

Дисперсия признака представляет собой средний квадрат отклонений вариантов от их средней величины, она вычисляется также по формулам простой и взвешенной дисперсий (в зависимости от исходных данных):

Простая дисперсия:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}$$

Взвешенная дисперсия:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f}.$$

Дисперсия, как средний квадрат, не имеет единиц измерения.

Дисперсия альтернативного признака

$$\sigma_p^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1-p)^2 p + (0-p)^2 q}{p+q}.$$

Подставив в формулу дисперсии $q = 1 - p$, получим:

$$\sigma_p^2 = \frac{q^2 p + p^2 q}{p+q} = \frac{pq(q+p)}{p+q} = pq.$$

Таким образом, $\sigma_p^2 = pq$ – дисперсия альтернативного признака равна произведению доли единиц, обладающих признаком, на долю единиц, не обладающих данным признаком.

Среднее квадратическое отклонение(σ *)*равно корню квадратному из дисперсии:

простое среднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}},$$

взвешенноесреднее квадратическое отклонение:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}},$$

Среднее квадратическое отклонение – это обобщающая характеристика размеров вариации признака в совокупности; оно показывает, на сколько в среднем отклоняются конкретные варианты от их среднего значения; является абсолютной мерой колеблемости признака и выражается в тех же единицах, что и варианты, поэтому экономически хорошо интерпретируется.

Пример 6.1. Предположим имеются данные о стоимости коттеджей, предлагаемых к продаже в Подмосковье.

Таблица 6.1

Цена 1 кв.м., долл. США	Общая площадь, тыс.кв.м. f	Середина интервала x	xf	$x - \bar{x}$	$(x-\bar{x})^2 \cdot f$
300-400	29,4	350	10290	-106	330338
400-500	20,5	450	9225	-6	738
500-600	7,3	550	4015	94	64503
600-700	7,0	650	4550	194	263452
700-800	4,0	750	3000	294	345744
Итого	68,2	X	31080	X	1004775

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{31080}{68{,}2} = 456 \text{ } \textit{долл.}$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{1004775}{68{,}2} = 14733$$

$\sigma = \sqrt{14733} = 121$ *долл.*.

Рассматриваемая величина показывает, что цены на коттеджи на фондовом рынке отклоняются от средней взвешенной цены на 121 доллар.

Среднее квадратическое отклонение альтернативного признака:

$$\sigma_p = \sqrt{pq} = \sqrt{p(1-p)}.$$

В статистической практике часто возникает необходимость сравнения вариаций различных признаков. Например, большой интерес представляет сравнение вариаций возраста рабочих и их квалификации, стажа работы и размера заработной платы, себестоимости и прибыли, стажа работы и производительности труда и т.д. Для подобных сопоставлений показатели абсолютной колеблемости признаков не пригодны: нельзя сравнивать колеблемость стажа работы, выраженного в годах, с вариацией заработной платы, выраженной в рублях.

Для осуществления такого рода сравнений, а также сравнений колеблемости одного и того же признака в нескольких совокупностях с различной средней арифметической используют относительные показатели вариации

Относительные показатели вариации определяются как отношение абсолютных показателей вариации к средней арифметической.

Это *коэффициент осцилляции,* определяемый как отношение размаха вариации к средней арифметической величине в процентах $V_R = \frac{R}{\bar{x}} \cdot 100\%$.

Линейный коэффициент вариации определяется аналогично, но по среднему линейному отклонению $V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100\%$.

Наиболее распространенными из них являются коэффициент вариации.

Коэффициент вариации представляет собой выраженное в процентах отношение среднего квадратического отклонения к средней арифметической:

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\ \% \ .$$

Относительные показатели вариации характеризуют степень колеблемости признака внутри средней величины. По величине, например, коэффициента вариации можно определить степень однородности изучаемой совокупности. Совокупность считается достаточно однородной, если коэффициент вариации не превышает 33 %. Для оценки качества, устойчивости средней величины установлены пределы. Самыми лучшими значениями коэффициента вариации являются $V_{\sigma} \leq 10\ \%$; допустимыми считаются значения до 50 %.

В рассматриваемом примере 6.1. коэффициент вариации равен: $V_{\sigma} = \frac{\sigma}{\bar{x}} \cdot 100\% = \frac{121}{456} \cdot 100\% = 26{,}5\%$. Следовательно, колеблемость цен небольшая, совокупность однородная и средняя цена может характеризовать совокупность.

6.3. Свойства дисперсии и упрощенные методы ее расчета

Техника вычисления дисперсии по формулам достаточно сложна, а при больших значениях вариантов и частот может быть громоздкой. Расчет можно упростить, используя свойства дисперсии (доказываемые в математической статистике):

Первое свойство – если все значения признака уменьшить на одну и ту же постоянную величину A, то дисперсия от этого не изменится:

$$\sigma^2_{(x-A)} = \sigma^2_x$$

Второе свойство – если все значения признака уменьшить в одно и то же число i раз, то дисперсия соответственно уменьшится в i^2 раз:

$$\sigma^2_{(\frac{x}{i})} = \frac{\sigma^2_x}{i^2}$$

Третье свойство (свойство минимальности) – средний квадрат отклонений от любой величины A (отличной от средней арифметической) большедисперсии признака на квадрат разности между средней арифметической и величиной A:

$$\sigma^2_A = \sigma^2_x + (x - A)^2$$

Используя свойства дисперсии, получим следующую *упрощенную формулу* вычисления дисперсии в вариационных рядах с равными интервалами *по способу моментов:*

$$\sigma^2 = i^2 \cdot (m_2 - m_1^2)$$

$$m_2 = \frac{\sum \left(\frac{x - A}{i} \right)^2 f^{/}}{\sum f^{/}}$$ – момент второго порядка;

$$m^2{}_1 = \left(\frac{\sum \frac{x - A}{i} \cdot f^{/}}{\sum f^{/}} \right)^2$$ – квадрат момента первого порядка.

Так на основании данных таблицы 5.3.1. при использовании метода моментов расчет дисперсии среднедушевого дохода будет включать расчет:

$$m_2 = \frac{\sum \left(\frac{x - A}{i} \right)^2 f}{\sum f} = \frac{162{,}7}{100} = 1{,}162$$

$$\sigma^2 = i^2 \cdot (m_2 - m_1^2) = 2^2 \cdot (1{,}162 - 0{,}193^2) = 4 \cdot 1{,}125 = 4{,}5$$

На основании последнего свойства дисперсии упрщенная формула дисперсии для любого ряда (дискретного, интервального с равным и неравным интервалами) примет вид:

$$\sigma^2 = \overline{x}^2 - (\overline{x})^2 = \frac{\sum x^2 \cdot f}{\sum f} - (\frac{\sum x \cdot f}{\sum f})^2$$

Пример 6.2. Дисперсия цен на основании данных таблицы 6.1 может быть определена следующим образом:

Таблица 6.1.1

Цена 1 кв.м., долл. США	Общая площадь, тыс.кв.м. f	Середина интервала x	xf	$x^2 f$
1	2	3	4	5
300-400	29,4	350	10290	3601500
400-500	20,5	450	9225	4151250
500-600	7,3	550	4015	2208250
600-700	7,0	650	4550	2957500
700-800	4,0	750	3000	2250000
Итого	68,2	Х	31080	15168500

$$\sigma^2 = \frac{\sum x^2 \cdot f}{\sum f} - \left(\frac{\sum x \cdot f}{\sum f} \right)^2 = \frac{15168500}{68,2} - \left(\frac{31080}{68,2} \right)^2 = 222412 - 207680 =$$

$$= 14733$$

6.4. Виды дисперсий

Вариация признака обусловлена различными факторами, некоторые из этих факторов можно выделить, если статистическую совокупность разбить на группы по какому-либо признаку. Тогда, наряду с изучением вариации признака по всей совокупности в целом, становится возможным изучить вариацию для каждой из составляющих ее группы, а также и между этими группами. В простейшем случае, когда совокупность расчленена на группы по одному фактору, изучение вариации достигается посредством исчисления и анализа трех видов дисперсий: *общей, межгрупповой и внутригрупповой.*

Общая дисперсия σ^2 измеряет вариацию признака по всей совокупности под влиянием всех факторов, обусловивших эту вариацию. Она равна среднему квадрату отклонений отдельных значений признака xот общей средней $\bar{x}$ и может быть вычислена как *простая дисперсия* или *взвешенная дисперсия*.

Межгрупповая дисперсия δ^2 характеризует систематическую вариацию результативного признака, обусловленную влиянием признака-фактора, положенного в основание группировки. Она равна среднему квадрату отклонений групповых (частных) средних:

$$\overline{\overline{x_i}} = \frac{\sum x_i}{n_i} \quad \text{или} \quad \bar{x} = \frac{\sum x_i \cdot f_i}{\sum f_i},$$

от общей средней $\bar{x} = \frac{\sum x}{n}$ *или* $\bar{x} = \frac{\sum x \cdot f}{\sum f}$

и может быть исчислена как *простая дисперсия* или как *взвешенная дисперсия* по формулам, соответственно:

$$\delta^2 = \frac{\sum (\overline{x_i} - \bar{x})^2}{n_i};$$

$$\delta^2 = \frac{\sum (\overline{x_i} - \bar{x})^2 f_i}{\sum f_i}$$

Внутригрупповая (частная) дисперсия (в каждой группе) σ_i^2, отражает случайную вариацию, т.е. часть вариации, обусловленную влиянием неучтенных факторов и не зависящую от признака-фактора, положенного в основание группировки. Она равна среднему квадрату отклонений отдельных значений признака внутри группы x от средней арифметической этой группы $\overline{\overline{x_i}}$, (групповой средней) и может быть исчислена как *простая дисперсия* или как *взвешенная дисперсия* по формулам, соответственно:

$$\sigma_i^{\ 2} = \frac{\sum (x - \bar{x}_i)^2}{n_i};$$

$$\sigma_i^2 = \frac{\sum (x - \bar{x}_i)^2 f_i}{\sum f_i}$$

На основании внутригрупповых дисперсий по каждой группе, т.е. на основании σ_i^2 можно определить *среднюю из внутригрупповых дисперсий:*

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i^2}{n} \text{ или } \overline{\sigma_i^2} = \frac{\sum \sigma_i^2 f_i}{\sum f_i}.$$

Согласно *правилу сложения дисперсий* общая дисперсия равна сумме средней из внутригрупповых и межгрупповой дисперсий:

$$\sigma^2 = \overline{\sigma_i^2} + \delta^2.$$

Пользуясь правилом сложения дисперсий, можно всегда по двум известным дисперсиям определить третью - неизвестную, а также судить о силе влияния группировочного признака.

Долю вариации группировочного признака в совокупности характеризует эмпирический коэффициент детерминации (η^2).

$$\eta^2 = \frac{\delta^2}{\sigma^2}$$

Пример 6.3. Предположим имеются данные о товарообороте магазинов в двух районах:

Таблица 6.2

Группы магазинов по величине товарооборота, тыс. руб.	Число магазинов		
	Район A (f_1)	Район $Б$ (f_2)	Итого (f)
А	1	2	3
400-600	6	3	9
600-800	17	20	37
800-1000	35	22	57
1000-1200	33	40	73
1200-1400	9	25	34
Итого	100	110	210

Необходимо дать характеристику вариации товарооборота магазинов в целом и в том числе по районам. Для расчета показателей вариации используем упрощенный метод моментов, т.к. интервал группировки равный. Определяем середину интервалов и выбираем значение «A»$=900$

Таблица 6.2.1

Середина интервала (x)	$x-A$ $A=900$	$\frac{x-A}{i}$ $i=200$	$\left(\frac{x-A}{i}\right)f_1$	$\left(\frac{x-A}{i}\right)f_2$	$\left(\frac{x-A}{i}\right)f$	$\left(\frac{x-A}{i}\right)^2 f_1$	$\left(\frac{x-A}{i}\right)^2 f_2$	$\left(\frac{x-A}{i}\right)^2 f$
4	5	6	7	8	9	10	11	12
500	- 400	-2	-12	-6	-18	24	12	36
700	- 200	-1	-17	-20	-37	17	20	37
900	-	-	-	-	-	-	-	-
1100	200	1	33	40	73	33	40	73
1300	400	2	18	50	68	36	100	136
x	x	x	22	64	86	110	172	282

где $m_1 = \dfrac{\sum\left(\frac{x-A}{i}\right)f}{\sum f}$ (по методу моментов)

$\bar{x}_1 = \dfrac{22}{100}200+900 = 0{,}22\cdot 200+900 = 944$ тыс. руб.- средний товарооборот одного магазина в районе А;

$\bar{x}_2 = \dfrac{64}{110}200+900 = 0{,}58\cdot 200+900 = 1016$ тыс. руб.- средний товарооборот одного магазина в районе Б;

$x = \dfrac{86}{210}200+900 = 0{,}41\cdot 200+900 = 982$ тыс. руб.- средний товарооборот одного магазина по всей совокупности магазинов.

$\sigma^2 = i^2 \cdot (m_2 - m_1{}^2)$, где $m_2 = \frac{\sum \left(\frac{x-A}{i}\right)^2 f}{\sum f}$ (по методу моментов) Тогда:

$$\sigma_1^2 = 200^2 \left(\frac{110}{100} - 0{,}22^2\right) = 42064$$ - дисперсия по району А;

$$\sigma_2^2 = 200^2 \left(\frac{172}{110} - 0{,}58^2\right) = 49088$$ - дисперсия по району Б.

$$\sigma^2 = 200^2 \left(\frac{282}{210} - 0{,}41^2\right) = 46992$$ - дисперсия общая.

Средняя из внутригрупповых дисперсий:

$$\overline{\sigma}_i^2 = \frac{\sum \sigma_i^2 f_i}{\sum f_i} = \frac{\sigma_1^2 f_1 + \sigma_2^2 f_2}{\sum f_1 + \sum f_2} = \frac{42064 \cdot 100 + 49088 \cdot 110}{210} = \frac{9606080}{210} = 45743$$

Межгрупповая дисперсия:

$$\delta^2 = \frac{\sum (\overline{x}_i - \overline{x})^2 f_i}{\sum f_i} = \frac{(\overline{x}_1 - \overline{x})^2 \sum f_1 + (\overline{x}_2 - \overline{x})^2 \sum f_2}{\sum f_1 + \sum f_2} =$$

$$= \frac{(944-982)^2 100 + (1016-982)^2 110}{210} = \frac{271560}{210} = 1293$$

Проверка правилом сложения дисперсий:

$$\sigma^2 = \overline{\sigma}_i^2 + \delta^2 = 45743 + 193 = 47036$$

т.е. σ^2,полученная по методу моментов $46992 \approx 47000$, равна σ^2,полученной как сумма средней из внутригрупповых дисперсий и межгрупповой дисперсии.

Эмпирический коэффициент детерминации:

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{1293}{47036} = 0{,}027$$

Следовательно, на группировочный территориальный признак (деление на районы) приходится лишь 2,7 % вариации.

ГЛАВА 7. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

7.1. Выборочное наблюдение как источник статистической информации

Под выборочным наблюдением понимается такое несплошное наблюдение, при котором статистическому обследованию подвергаются единицы совокупности, отобранные в случайном порядке.

Переход статистики РФ на международные стандарты требует более широкого применения выборки для получения и анализа показателей во многих секторах экономики. К выборочному наблюдению статистика прибегает по различным причинам. Существование множества субъектов хозяйственной деятельности, которые характерны для рыночной экономики, не позволяет использовать сплошное обследование из-за огромных материальных, финансовых и трудовых затрат. Выборочное наблюдение экономит ресурсы, позволяет расширить программу наблюдения и использовать более квалифицированные кадры для проведения наблюдения. Выборочное наблюдение используют и для решения таких задач, где сплошное наблюдение применять невозможно (изучение качества продукции) или нецелесообразно, а также для уточнения и проверки результатов сплошного наблюдения. В отличие от других видов несплошного наблюдения выборочное наблюдение позволяет получить необходимые сведения приемлемой точности.

Совокупность отобранных для обследования единиц в статистике называют *выборочной,* а совокупность единиц, из которых производится отбор – *генеральной.*

Результаты выборочного статистического исследования во многом зависят от *уровня подготовки процесса наблюдения*. В данном случае подразумевается соблюдение определенных правил и принципов проектирования выборочного обследования. Особенно важным является составление организационного плана выборочного наблюдения. В организационный план включаются следующие вопросы:

1. Постановка цели и задачи наблюдения.

2. Определение границ объекта исследования.

3. Отработка программы наблюдения и разработки ее материалов.

4. Определение процедуры отбора, способа отбора и объема выборки.

5. Подготовка кадров для проведения наблюдения, тиражирование формуляров, инструктивных материалов.

6. Расчет выборочных характеристик и определение ошибок выборки.

7. Распространение выборочных данных на всю генеральную совокупность.

Основные характеристики параметров генеральной и выборочной совокупности обозначаются определенными символами (таблица 7.1.).

Таблица 7.1.

Символы основных характеристик параметров генеральной и выборочной совокупностей.

Характеристика	Генеральная совокупность	Выборочная совокупность
Объем совокупности (численность единиц)	N	n
Численность единиц, обладающих обследуемым признаком	M	m
Доля единиц, обладающих обследуемым признаком	$P = \frac{M}{N}$	$W = \frac{m}{n}$
Средний размер признака	$\bar{x} = \frac{\sum x_i}{N}$	$\tilde{x} = \frac{\sum x_i}{n}$
Дисперсия количественного признака	$\sigma_{\bar{x}}^2 = \frac{\sum (x_i - \bar{x})^2}{N}$	$\sigma_{\tilde{x}}^2 = \frac{\sum (x_i - \tilde{x})^2}{n}$
Дисперсия доли	$\sigma_p^2 = pq$	$\sigma_w^2 = W \cdot (1 - W)$

7.2. Основные способы формирования выборочной совокупности

Достоверность рассчитанных по выборочным данным характеристик зависит от способа отбора единиц из генеральной совокупности. В каждом конкретном случае в зависимости от ряда условий выбирают наиболее предпочтительную систему организации отбора, которая определяется видом, методом и способом отбора.

По виду различают индивидуальный, групповой и комбинированный отбор. При *индивидуальном* отборе в выборочную совокупность отбираются отдельные единицы генеральной совокупности. При *групповом* отборе отбираются группы единиц. *Комбинированный* отбор предполагает сочетание индивидуального и группового отбора.

Метод отбора определяет возможность продолжения участия отобранной единицы в процедуре отбора.

Бесповторным называется отбор, при котором попавшая в выборку единица не возвращается в совокупность, из которой осуществляется дальнейший отбор.

При *повторном* отборе попавшая в выборку единица после регистрации наблюдаемых признаков возвращается в исходную (генеральную) совокупность для участия в дальнейшем отборе. Повторный метод отбора применяется в тех случаях, когда характер исследования предполагает возможность повторной регистрации единиц. Например, в выборочных обследованиях населения в качестве покупателей, избирателей, абитуриентов и т. д.

Способ отбора определяет конкретный механизм выборки единиц из генеральной совокупности. В практике обследований получили распространение следующие виды выборки:

- собственно - случайная;
- механическая;
- типическая;
- серийная;
- комбинированная.

Собственно-случайная выборка заключается в отборе единиц из генеральной совокупности наугад, без какой либо системности. Технически этот отбор проводят методом жеребьевки (использование фишек, шаров, карточек и т.д. в количестве генеральной совокупности) или по таблице случайных чисел (произвольные столбцы цифр).

Собственно-случайный отбор может быть повторным и бесповторным.

После проведения отбора для определения возможных границ генеральных характеристик рассчитывается средняя и предельная ошибки выборки.

Величина *средней ошибки выборки* рассчитывается дифференцированно в зависимости от способа отбора по формулам:

при повторном отборе $\mu = \dfrac{\sigma}{\sqrt{n}}$,

при бесповторном отборе $\mu = \sqrt{\dfrac{\sigma^2}{n}\left(1 - \dfrac{n}{N}\right)}$, где

σ^2 - выборочная (или генеральная) дисперсия;

σ - выборочное (или генеральное) среднее квадратическое отклонение;

n - объем выборочной совокупности;

N -объем генеральной совокупности

Предельная ошибка выборки связана со средней ошибкой выборки соотношением: $\Delta = t \cdot \mu$, где

Δ - предельная ошибка выборки;

μ - средняя ошибка выборки;

t - коэффициент доверия, определяемый в зависимости от уровня вероятности *P*.

Таблица 7.2

Значения t в зависимости от уровня вероятности

Вероятность, P_i	0,683	0,866	0,954	0,988	0,997	0,999
Значение t	1,0	1,5	2,0	2,5	3,0	3,5

Расчет средней и предельной ошибок выборки позволяет определить возможные пределы, в которых будут находиться характеристики генеральной совокупности. Так для средней в генеральной совокупности эти пределы будут $\bar{x} = \tilde{x} \pm \Delta_{\tilde{x}}$, где:

$\bar{x}$ -генеральная средняя;

$\tilde{x}$ - выборочная средняя;

$\Delta_{\tilde{x}}$ -предельная ошибка для выборочной средней.

$\Delta_{\tilde{x}} = t \cdot \dfrac{\sigma}{\sqrt{n}}$ *или* $\Delta_{\tilde{x}} = t \cdot \sqrt{\dfrac{\sigma^2}{n}}$, *а при бесповторном отборе* $\Delta_{\tilde{x}} = t \cdot \sqrt{\dfrac{\sigma^2}{n}\left(1 - \dfrac{n}{N}\right)}$.

Пример 7.1.Предположим в городе, где проживает 250 тыс.семей. Для определения числа детей в семье была организована 2%-я случайная бесповторная выборка семей. По ее результатам были получены следующие данные о распределении семей по числу детей.

Таблица 7.3

Число детей в семье	0	1	2	3	4	5
Количество семей	1000	2000	1200	400	200	200

С вероятностью 0,954 необходимо определить пределы, в которых будет находиться число детей в семье в генеральной совокупности. Для этого сначала находим среднее число детей в семье и дисперсию в выборочной совокупности.

Таблица 7.3.1

Число детей в семье X	Количество семей f	xf	$x^2 f$
1	2	3	4
0	1000	0	0
1	2000	2000	2000
2	1200	2400	4800
3	400	1200	3600
4	200	800	3200
5	200	1000	5000
Итого	5000	7400	18600

$$\tilde{x} = \frac{7400}{5000} = 1{,}5 \textit{чел.}$$

$$\sigma^2_{\tilde{x}} = \frac{\sum x^2 \cdot f}{\sum f} - \left(\frac{\sum x \cdot f}{\sum f}\right)^2 = \frac{18600}{5000} - (1{,}5)^2 = 3{,}72 - 2{,}25 = 1{,}47$$

Тогда предельная ошибка выборки будет равна $\Delta_{\tilde{x}} = t \cdot \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)} = 2 \cdot \sqrt{\frac{1{,}47}{5000} \cdot \left(1 - \frac{5000}{25000}\right)} = 0{,}015$, *т.к.* при вероятности $P = 0{,}954$ $t = 2$ и при 2% отборе $N = \frac{5000}{0{,}02} = 25000$ семей.

Следовательно, пределы генеральной средней будут:

$$\bar{x} = \tilde{x} \pm \Delta_{\tilde{x}} = 1{,}5 \pm 0{,}015$$

Таким образом, с вероятностью 0,954 можно утверждать, что среднее число детей в семьях города практически не отличается от 1,5, т.е. в среднем на каждые две семьи приходится 3 ребенка.

Эти же показатели могут быть определены и для доли признака. В этом случае особенности расчета связаны с определением дисперсии доли, которая определяется по формуле: $\sigma^2_w = w \cdot (1 - w)$, где $w = \frac{m}{n}$ – доля единиц, обладающих признаком в выборочной совокупности.

Тогда, например, при собственно-случайном отборе для определения предельной ошибки выборки при *повторном отборе* используется формула:

$$\Delta_w = t\sqrt{\frac{\sigma_w^2}{n}} = t \cdot \sqrt{\frac{w \cdot (1-w)}{n}}$$, а при *бесповторном отборе*

$$\Delta_w = t \cdot \sqrt{\frac{\sigma_w^2}{n} \cdot \left(1 - \frac{n}{N}\right)} = t \cdot \sqrt{\frac{w \cdot (1-w)}{n}\left(1 - \frac{n}{N}\right)}.$$

Пределы доли признака в генеральной совокупности будут $p = w \pm \Delta_w$.

По данным таблицы 7.3 определим долю семей с двумя детьми в генеральной совокупности с той же вероятностью 0,954.

Выборочная доля (доля семей с двумя детьми в выборочной совокупности) оказалась равной $w = \frac{1200}{5000} = 0{,}24$ или 24 %. Предельная ошибка для доли:

$$\Delta_w = t \cdot \sqrt{\frac{w \cdot (1-w)}{n}\left(1 - \frac{n}{N}\right)} = 2 \cdot \sqrt{\frac{0{,}24 \cdot (1 - 0{,}24)}{5000} \cdot \left(1 - \frac{5000}{25000}\right)} =$$

$$= 0{,}011 \ \textit{или} \ 1{,}1\%$$

Пределы доли признака в генеральной совокупности $24\% - 1{,}1\% \le P \le 24\% + 1{,}1\%$ или $22{,}9\% \le P \le 25{,}1\%$. Таким образом, с вероятностью 0,954 можно утверждать, что доля семей с двумя детьми в семьях города будет находиться в пределах от 22,9% до 25,1%.

Механическая выборка применяется в случаях, когда генеральная совокупность упорядочена, т.е. имеется определенная последовательность в расположении единиц (табельные номера работников, списки избирателей, телефонные номера, номера домов и т.д.).

Для механической выборки устанавливается пропорция отбора, которая определяется соотнесением объемов выборочной и генеральной совокупностей. Так, если генеральная совокупность 500 000 единиц и предполагается получить 2% выборку, т.е. отобрать 10 000 единиц, то пропорция отбора составит: $\frac{1}{50}=\left(\frac{1}{\frac{500000}{10000}}\right)$.

Отбор осуществляется в соответствии с установленной пропорцией через равные интервалы. Например, при пропорции $\frac{1}{50}$ (2% выборка) отбирается каждая 50-я единица.

Для определения средней ошибки механической выборки используется формула средней ошибки при собственно-случайном бесповторном отборе.

Типический отбор используется в тех случаях, когда все единицы генеральной совокупности можно разбить на несколько типических групп. Например, при обследовании населения это могут быть районы, возрастные или образовательные группы и т.д.

Средняя ошибка такой выборки находится по формулам:

при повторном отборе $\mu=\sqrt{\frac{\overline{\sigma_i^2}}{n}}$,

при бесповторном отборе $\mu=\sqrt{\frac{\overline{\sigma_i^2}}{n}\left(1-\frac{n}{N}\right)}$, где

$\overline{\sigma_i^2}=\frac{\sum\sigma_i^2\cdot f_i}{\sum f_i}$ - средняя из внутригрупповых дисперсий.

Определим пример7.2. типической выборки на условном примере. Предположим, 10% бесповторный типический отбор рабочих предприятия, пропорциональный размерам цехов, проведенный с целью оценки потерь из-за временной нетрудоспособности, привел к следующим результатам:

Таблица 7.4

Цех	Всего рабочих, чел.	Обследовано, чел.	Число дней временной нетрудоспособности за год	
			Средняя	дисперсия
1	1000	100	18	49
2	1400	140	25	25
3	800	80	15	16
Итого	3200	320	*x*	*x*

Определим среднюю из внутригрупповых дисперсий:

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i^2 \cdot f_i}{\sum f_i} = \frac{49 \cdot 100 + 25 \cdot 140 + 16 \cdot 80}{100 + 140 + 80} = 30{,}25$$

Определяем среднюю и предельную ошибки выборки с вероятностью 0,954:

$$\mu = \sqrt{\frac{\overline{\sigma_i^2}}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{30{,}25}{320} \cdot \left(1 - \frac{320}{3200}\right)} = 0{,}29;$$

$$\Delta_{\tilde{x}} = 2 \cdot 0{,}29 = 0{,}58$$

Определим выборочную среднюю:

$$\tilde{x} = \frac{\sum x_i \cdot f_i}{\sum f_i} = \frac{18 \cdot 100 + 25 \cdot 140 + 15 \cdot 80}{320} = 14{,}6 \textit{ дня.}$$

С вероятностью 0,954 можно сделать вывод, что среднее число дней нетрудоспособности одного рабочего в целом по предприятию находится в пределах $14{,}6 - 0{,}58 \le \bar{x} \le 14{,}6 + 0{,}58$, т.е. от 14 до 15 дней.

Серийный отбор удобен в тех случаях, когда единицы совокупности объединены в небольшие группы или серии. Поскольку внутри групп (серий) обследуются все без исключения единицы, средняя ошибка серийной выборки (при отборе равновеликих серий) зависит от величины только межгрупповой (межсерийной) дисперсии и определяется по следующим формулам:

при повторном отборе $\mu = \sqrt{\frac{\delta^2}{r}}$,

при бесповторном отборе $\mu = \sqrt{\frac{\delta^2}{r}\left(1 - \frac{r}{R}\right)}$, где

r -число отобранных серий (групп);

R - общее число серий (групп).

Межгрупповую дисперсию вычисляют по формуле: $\delta^2 = \frac{\sum(\tilde{x}_i - \tilde{x})^2}{r}$, где

$\tilde{x}_i$ - средняя i - ой серии (группы);

$\tilde{x}$ - общая средняя по всей выборочной совокупности.

Рассмотрим условный пример:

В области, состоящей из 20 районов, проводилось выборочное обследование урожайности на основе отбора серий (районов). Выборочные средние по районам составили($\tilde{x}_i$) соответственно 14,5 ц/га; 16,0 ц/га; 15,5 ц/га; 15,0 ц/га; 14,0ц/га. С вероятностью 0,954 определим пределы урожайности во всей области.

Общая средняя во всей выборочной совокупности:

$$\tilde{x} = \frac{14{,}5 + 16{,}0 + 15{,}5 + 15{,}0 + 14{,}0}{5} = 15{,}0\ ц/га$$

Межгрупповая (межсерийная) дисперсия равна:

$$\delta^2 = \frac{\sum(\tilde{x}_i - \tilde{x})^2}{r} =$$

$$= \frac{(14{,}5 - 15{,}0)^2 + (16{,}0 - 15{,}0)^2 + (15{,}5 - 15{,}0)^2 + (14{,}0 - 15{,}0)^2}{5} = 0{,}5$$

Тогда предельная ошибка серийной бесповторной выборки ($t = 2$ при P=0,954): $\Delta = t \cdot \sqrt{\frac{\delta^2}{r}\left(1 - \frac{r}{R}\right)} = 2 \cdot \sqrt{\frac{0{,}5}{5}\left(1 - \frac{5}{20}\right)} \approx 1{,}7$

Следовательно, урожайность в области будет находиться с вероятностью 0,954 в пределах $15{,}0 - 1{,}7 \le \overline{X} \le 15{,}0 + 1{,}7$ или $13{,}3\ ц/га \le \overline{X} \le 16{,}7\ ц/га$.

7.3. Определение необходимого объема выборки

При проектировании выборочного наблюдения вопрос о необходимой численности выборки. Эта численность может быть определена на базе допустимой ошибки при выборочном наблюдении, исходя из вероятности, на основе которой можно гарантировать величину установленной ошибки, а также на базе способа отбора.

Наиболее часто применяемые на практике *формулы объема выборки* для *собственно-случайной и механической выборки*:

$$n = \frac{t^2\sigma^2}{\Delta^2}\ (\textit{повторный отбор});$$

$$n = \frac{t^2\sigma^2 N}{\Delta^2 N + t^2\sigma^2}\ (\textit{бесповторный отбор});$$

Для *типической выборки:*

$$n = \frac{t^2\overline{\sigma_i}^2}{\Delta^2}\ (\textit{повторный отбор});$$

$$n = \frac{t^2\overline{\sigma_i}^2 N}{\Delta^2 N + t^2\overline{\sigma_i}^2}\ (\textit{бесповторный отбор});$$

Для *серийной выборки:*

$$n = \frac{t^2\delta^2}{\Delta^2}\ (\textit{повторный отбор});$$

$$n = \frac{t^2\delta^2 R}{\Delta^2 R + t^2\delta^2}\ (\textit{бесповторный отбор});$$

В зависимости от целей исследования дисперсии и ошибки выборки могут быть рассчитаны для средней величины и для доли признака.

Рассмотрим пример 7.3.

В 100 туристических агентствах города предполагается провести обследование количества реализованных путевок методом механического отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,683 ошибка не превышала 3 путевок, если по данным пробного обследования дисперсия составляет 225.

Необходимый объем выборки будет равен:

$$n = \frac{t^2 \overline{\sigma_i}^2 N}{\Delta^2 N + t^2 \overline{\sigma_i}^2} = \frac{1^2 \cdot 225 \cdot 100}{3^2 \cdot 100 + 1^2 \cdot 225} = \frac{22500}{1125} = 20 \text{ агентств}$$

7.4. Малая выборка

В практике статистического исследования в условиях рыночной экономики все чаще приходится сталкиваться с небольшими по объему так называемыми малыми выборками. Под *малой выборкой* понимается такое выборочное наблюдение, численность единиц которого не превышает 30.

В настоящее время малая выборка используется более широко, чем раньше за счет статистического изучения деятельности малых и средних предприятий, коммерческих банков, фермерских хозяйств и т. Д. Их количество, особенно при региональных исследованиях, а также величина характеризующих их показателей, часто незначительны. Поэтому хотя общий принцип выборочного обследования (с увеличением объема выборки точность выборочных данных повышается) остается, иногда приходится ограничиваться малым числом наблюдений. Необходимость в малой выборке возникает также в научно-исследовательской работе.

При оценке результатов малой выборки величина генеральной дисперсии в расчетах не используется. Для определения возможных пределов ошибки пользуются критерием Стьюдента, определяемым по формуле:

$t = \frac{\tilde{x} - \bar{x}}{\mu_{мв}}$, где $\mu_{мв} = \frac{\sigma}{\sqrt{n-1}}$. Распределение вероятности в олых выборках приведены в приложении 3.

Пример 7.4. Предположим, что выборочное обследование 10 рабочих малого предприятия показало, что на выполнение одной из производственных операций рабочие затрачивали времени (мин.): 3,4; 4,7;1,8; 3,9; 4,2; 3,9; 3,7; 3,2; 2,2; 3,9.

Определяем среднее время на операцию (выборочная средняя).

$$\tilde{x} = \frac{3{,}4 + 4{,}7 + 1{,}8 + 3{,}9 + 4{,}2 + 3{,}9 + 3{,}7 + 3{,}2 + 2{,}2 + 3{,}9}{10} = 3{,}49 \text{ мин.}$$

Выборочная дисперсия:

$$\sigma^2 = \frac{(3{,}4 - 3{,}49)^2 + (4{,}7 - 3{,}49)^2 + \ldots + (3{,}9 - 3{,}49)^2}{10} = 0{,}713$$

Отсюда средняя ошибка малой выборки равна:

$$\mu_{мв} = \frac{\sigma}{\sqrt{n-1}} = \sqrt{\frac{\sigma^2}{n-1}} = \sqrt{\frac{0{,}713}{10-1}} = 0{,}28 \text{ мин.}$$

По таблице 4 в приложении 3 находим, что для коэффициента доверия $t = 2$ и объема малой выборки $n = 10$ вероятность равна 0,924. таким образом, с вероятностью 0,924 можно утверждать, что расхождение между выборочной и генеральной средними лежит в пределах от -2μ до $+2\mu$.

Т.е. разность $\tilde{x} - \bar{x}$ не превысит по абсолютной величине $2 \cdot 0{,}28 = 0{,}56$. Следовательно, средние затраты времени во всей совокупности будут находиться в пределах от 2,93 до 4,05 мин. Вероятность того, что это предположение в действительности неверно и ошибка по случайным причинам будет по абсолютной величине больше, чем 0,56 равна $1 - 0{,}924 = 0{,}076$.

ГЛАВА 8. СТАТИСТИЧЕСКИЕ МЕТОДЫ ИЗУЧЕНИЯ СВЯЗЕЙ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

8.1. Виды взаимосвязей

Все явления в природе и обществе находятся во взаимной связи и взаимной обусловленности. Статистика изучает закономерности изменения одних явлений в связи с изменением других.

Экономика страны и отдельные предприятия характеризуются системой показателей, образующих диалектическое единство. Эти показатели связаны между собой и порождают друг друга.

Связь явлений имеет разнообразные проявления. Существуют различные формы и виды связей, которые отличаются по существу, характеру проявления, направлению, тесноте, аналитическому выражению и т.д.

По степени зависимости одного явления от другого различают в общем виде два типа связи: связь функциональную (полную) и связь стохастическую (неполную).

Функциональная связь – это связь, где каждому значению одной переменной (аргументу) соответствует одно вполне определенное значение другой переменной (функции). Такие связи широко распространены в технике, биологии, математике, Например, площадь круга определяется однозначно величиной радиуса $S = \pi R^2$.

При *стохастической форе связи* каждому значению одного признака (факторного) соответствует целый ряд значений другого признака (результативного). Следовательно, стохастическая связь проявляется не в каждом отдельном случае, а лишь в среднем для совокупности явлений данного вида.

Социально – экономические процессы и явления – это результат действия многочисленных факторов. Одни из них поддаются точному измерению, а другие – нет, т.е. их можно измерить только приближенно. Для социально- экономических явлений характерен тот факт, что наряду с факторами, определяющими исследуемую зависимость, действуют многочисленные случайные факторы. Поэтому зависимость проявляется не в каждом отдельном случае, а лишь, в общем, в среднем.

Статистика призвана определять наличие связи между явлениями, ее направление и форму выражения, измерять тесноту этой связи.

8.2. Методы изучения взаимосвязей

Для изучения связи между явлениями статистика использует ряд методов и приемов, важнейшие из которых: метод приведения параллельных рядов, метод группировок, индексный метод, балансовый метод и группа корреляционных методов.

Метод приведения параллельных рядов заключается в установлении связи между явлениями посредством сопоставления двух или нескольких рядов показателей. Такое сопоставление производится после того, как теоретически доказана возможность связи между изучаемыми показателями. Сопоставление параллельных рядов позволяет установить *наличие связи* и получить представление о ее *характере.* Сущность метода параллельных рядов заключается в следующем: факторный признак располагается в возрастающем (или убывающем) порядке и параллельно располагаются соответствующие значения одного или нескольких результативных признаков. Сравнивая, расположенные таким образом ряды показателей, выявляется существование связи и ее направление.

Метод параллельных рядов прост и достаточно эффективен на первых стадиях исследования.

Метод аналитических группировок позволяет не только констатировать наличие связи между изучаемыми признаками, но и выявлять причины этой связи. Чтобы анализировать сложные взаимные связи между несколькими признаками применяются комбинационные группировки. В основе группировки всегда *факторный* признак. Затем для каждой выделенной группы рассчитываются обобщающие показатели. В итоге рассматривают, какое влияние оказывает факторный признак на результативный. С помощью метода группировок можно рассматривать одновременное действие *нескольких признаков – факторов*, а также характеризовать структуру совокупности.

Балансовый метод заключается в построении различных балансовых равенств в виде *соотношений между наличием и распределением тех или иных ресурсов,* ввозом и вывозом и т. Д. Простейшим балансом такого рода является баланс материальных ресурсов на предприятии, Здесь балансовое равенство можно записать так:

Остаток на начало периода + поступление = расход + остаток на конец периода. Балансы позволяют выявить взаимосвязи в образовании и распределении ресурсов между предприятиями, районами и т. Д., позволяют анализировать сложившиеся пропорции и зависимости. Такого рода балансы распространены в торговле, балансовым методом изучают движение рабочей силы, финансов, основных фондов и т. Д. На основе балансов выявляют важные для анализа развития народного хозяйства показатели.

Индексный метод служит для определения *роли отдельных факторов* в изменении изучаемого явления с целью воздействия на положительно влияющие факторы. Исследование удельного веса факторов опирается на *взаимосвязи связанных явлений.* Факторный индексный анализ позволяет численно точно определить степень влияния каждого фактора в совместном влиянии факторов.

Корреляционные методы выявления взаимосвязей в отличие от вышеизложенных методов изучения взаимосвязей не только позволяют установить связь и выявить ее причины, но и позволяют измерить *степень тесноты связи.* Они дают возможность выразить эту связь *аналитически* в виде определенного математического уравнения. Корреляционные методы анализа являются основными в изучении связей между социально - экономическими явлениями. Корреляционная зависимость исследуется с помощью корреляционного и регрессионного анализов. Корреляционный анализ позволяет оценить *тесноту связи* с помощью парных, частных и множественных коэффициентов корреляции. Целью регрессионного анализа является оценка функциональной зависимости среднего значения результативного признака *(* y *)* от факторного *(* X *)* или факторных *(* $X_1, X_2, X_3, \dots X_n$).

Корреляционные методы изучения взаимосвязей можно разделить на две группы: непараметрические методы и методы собственно корреляции.

8.3. Непараметрические корреляционные методы изучения взаимосвязей

Непараметрические корреляционные методы исследования связей включают расчеты различных коэффициентов, с помощью которых орреляется теснота связи между явлениями, где обычные методы корреляции *недостаточны* или *невозможны*. Например, при определении тесноты связи между *качественными* признаками. Непараметрические методы не требуют никаких предположений о законе распределения исходных данных, т.к. при их использовании оперируют не значениями признаков, а их частотами, знаками, рангами и т.д. Это ранговый коэффициент Спирмена, коэффициент Фехнера, коэффициенты ассоциации и контингенции, коэффициенты взаимной сопряженности Пирсона и Чупрова, коэффициент корреляции рангов Кендалла.

Ранговый коэффициент Спирмена измеряет взаимосвязь между отдельными признаками с помощью условной оценки по рангам. Ранг ® – это порядковый номер значений признака, расположенных в порядке возрастания или убывания их величины на основе предпочтения (лучший – на первом месте, худший – на последнем). Рассчитывается он по формуле:

$$P = 1 - \frac{6\sum d_i^2}{n \cdot (n^2 - 1)}, \quad \text{где } d_i^2 = (R_x - R_y)^2, \; n \text{ – число наблюдений}$$

Коэффициент Спирмена изменяется от – 1 до + 1 и равен нулю при отсутствии связи. Эта формула используется, когда нет связанных (одинаковых в ряду) рангов. Если значения признака совпадают (появляются одинаковые в ряду ранги-связные), то определяется средний ранг путем деления суммы рангов на число значений. Коэффициент Спирмена в этом случае определяется по формуле

$$P = 1 - \frac{s \cdot \sum d_i^2}{n \cdot (n^2 - 1)}, \text{где } S - \text{средний ранг по совокупности}$$

Критерий тесноты связи для коэффициента Спирмена 0,5, т.е. $P \geq 0{,}5$.

Значимость коэффициента Спирмена проверяется на основе критерия Стьюдента. Расчетное значение критерия Стьюдента определяется по формуле $t_P = P \cdot \sqrt{\frac{n-2}{1-P^2}}$, которое сравнивается с теоретическим значением (t_T) при заданном уровне значимости и числе степеней свободы $(n - m)$. Значение коэффициента корреляции рангов Спирмена считается существенным, если $t_p \rangle t_T$.

Рассмотрим пример 8.1.:

По данным группы предприятий, выставившим акции на чековые аукционы необходимо определить с помощью коэффициента Спирмена зависимость между величиной уставного капитала и количеством выставленных акций.

Таблица 8.1

Расчет коэффициента Спирмена

№ пред-прия-тия	Уставный капитал (x)		Число корреляции акций (y)		Квадрат разности рангов $d^2 = (Rx - Ry)^2$
	Млн. руб.	Ранг Rx	Шт.	Ранг Ry	
1	2	3	4	5	6
1	2954	2	856	4	4
2	1605	10	930	2	64
3	4102	1	1563	1	0
4	2350	5	682	6	1
5	2625	4	616	8	16
6	1795	7	495	9	4
1	2	3	4	5	6
7	2813	3	815	5	4
8	1751	8	858	3	25
9	1700	9	467	10	1
10	2264	6	661	7	1
Итого	23959		7943		120

$$P = 1 - \frac{6\sum d_i^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \cdot 120}{10 \cdot (100 - 1)} = 1 - \frac{720}{990} = 1 - 0{,}727 = 0{,}273 \approx 0{,}3$$

Следовательно, связь прямая и умеренная. Значимость коэффициента Спирмена определяем по критерию Стьюдента

$$t_P = P \cdot \sqrt{\frac{n-2}{1-P^2}} = 0{,}3 \cdot \sqrt{\frac{10-2}{1-0{,}3^2}} = 0{,}889$$

Сопоставляем его с теоретическим значением (приложение 3). С вероятностью 0,9 и числом степеней свободы $9 = 10 - 1$ $t_T = 1{,}833$. Следовательно, т. К. $t_p \leq t_T$ найденная связь не значима.

Коэффициент Фехнера ($K_ф$) или коэффициент совпадения знаков основан на применении первых степеней отклонений от средних значений признаков двух связанных рядов показателей.

$$K_\phi = \frac{a-b}{a+b} \quad ,$$

где a – количество совпадений знаков отклонений;

b – количество несовпадений знаков отклонений.

Коэффициент Фехнера также изменяется от -1 до +1 и равен нулю при отсутствии связи.

По данным примера 8.1 рассчитаем коэффициент Фехнера.

Пример 8.2.

Таблица 8.2

Расчет коэффициента Фехнера

№ предприятия	Уставный капитал,млн.руб.. x	Число выставленных акций,шт. y	Знаки отклонений	
			$x-\bar{x}$	$y-\bar{y}$
1	2	3	4	5
1	2954	856	+	+
2	1605	930	-	+
3	4102	1563	+	+
4	2350	682	-	-
5	2625	616	+	-
6	1795	495	-	-
7	2813	815	+	+
8	1751	858	-	+
9	1700	467	-	-
10	2264	661	-	-
Итого	23959	7943		

$$\bar{x} = \frac{23959}{10} = 2395{,}9 \textit{ млн.руб.} \quad \bar{y} = \frac{7943}{10} = 794 \textit{ шт.}$$

$$K_\phi = \frac{a-b}{a+b} = \frac{7-3}{7+3} = 0{,}4$$

Связь прямая и умеренная.

Коэффициент корреляции рангов Кендалла также используется для измерения тесноты связи между качественными признаками, ранжированными по одному принципу. Расчет осуществляется по формуле:

$$\tau = \frac{2S}{n \cdot (n-1)}$$, где n – число наблюдений; $S = P + Q$.

Для нахождения P и Q надо произвести ранжирование по факторному признаку *(* x *)* в порядке предпочтительности и ранжирование по результативному признаку *(* y *)* соответственно предпочтительности факторного признака. Тогда P – это количество чисел, находящихся после каждого из элементов последовательности рангов переменной (y *)* и имеющих величину ранга больше ранга рассматриваемого элемента, а Q - это количество чисел находящихся после каждого из элементов последовательности рангов переменной (y *)*, имеющих величину ранга меньше ранга рассматриваемого элемента и взятых со знаком минус. Коэффициент Кендалла изменяется от -1 до + 1 и равен нулю при отсутствии связи. Значимость коэффициента Кендалла также определяется по *t*критерию Стьюдента.

Пример 8.3. Необходимо определить степень тесноты связи между уровнем механизации труда *(* x *)* и трудоемкостью единицы продукции *(* y *)*по данным 10 заводов:

Таблица 8.3

Номер завода	1	2	3	4	5	6	7	8	9	10
Уровень механизации труда, %	65	66	67	61	60	62	63	64	68	69
Трудоемкость единицы продукции, мин.	12	10	14	11	15	13	17	16	8	7
Ранг по x	5	4	3	9	10	8	7	6	2	1
Ранг по y	6	8	4	7	3	5	1	2	9	10

Ранг по x проставляется от большего к меньшему, т.к. лучшее значение большее. Ранг по y проставляется в соответствии с ранжированием x, т.е. тоже от большего к меньшему. Располагаем ранги по x в порядке возрастания, а по y в соответствии с x.

Таблица 8.4

Номер завода	5	4	6	7	8	1	2	3	9	10
Уровень механизации труда, %	60	61	62	63	64	65	66	67	68	69
Трудоемкость единицы продукции, мин.	15	11	13	17	16	12	10	14	8	7
Ранг по x	1	2	3	4	5	6	7	8	9	10
Ранг по y	10	9	4	8	6	2	1	5	7	3

Определяем $P = 0+0+4+0+1+3+3+1+0 = 12$, т.к. после 10 ранга по y нет чисел больше 10 (0), после 9 нет чисел больше 9 (0), после 4 четыре числа больше 4 (8; 6; 5; 7), после 8 нет чисел больше 8 (0), после 6 одно число больше 6 (7), после 2 три числа больше 2 (5; 7; 3), после 1 три числа больше 1 (5; 7; 3), после 5 одно число больше 5 (7), после 7 нет чисел больше 7 (0).

Определяем $Q = -9-8-3-6-4-1-0-1-1 = -33$, т.к. после 10 девять чисел меньше 10, после 9 восемь чисел меньше 9. после 4 три числа меньше 4 и т.д. Следовательно, $\tau = \dfrac{2S}{n\cdot(n-1)} = \dfrac{2\cdot(12-33)}{10\cdot(10-1)} = \dfrac{-42}{90} = -0{,}467$

связь умеренная и обратная. Проверяем значимость коэффициента Кендалла по t критерию Стьюдента

$$t_P = \tau\cdot\sqrt{\frac{n-2}{1-\tau^2}} = -0{,}467\cdot\sqrt{\frac{10-2}{1-0{,}467^2}} = -1{,}494.$$

Следовательно, с вероятностью 0,9 можно утверждать, что найденная связь yt значима, т.к. расчетное значение меньше теоретического ($t_T = 1{,}833$).

Если в изучаемой совокупности есть связанные ранги, то расчеты коэффициента Кендалла необходимо произвести по следующей формуле:

$$\tau = \frac{S}{\sqrt{\left[0{,}5 \cdot n(n-1) - v_x\right] \cdot \left[0.5 \cdot n(n-1) - v_y\right]}}, \text{ где}$$

$$v_x = 0{,}5 \cdot \sum t_j (t_j - 1); \quad v_y = 0{,}5 \cdot \sum t_j \cdot (t_j - 1);$$

t_j – *количество одинаковых рангов j – го значения соответственно призаков x или y*

В практике статистических исследований приходится иногда анализировать связь между *альтернативными признаками,* представленными только группами *с противоположными (взаимоисключающими) характеристиками.* Тесноту связи в этом случае можно оценить с помощью коэффициентов ассоциации и контингенции.

Коэффициент ассоциации определяется по формуле:

$K_a = \dfrac{ad - bc}{ad + bc}$ Связь считается подтвержденной, если $K_a \geq 0{,}5$.

Коэффициент контингенции определяется по формуле:

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

Связь считается подтвержденной, если $K_K \geq 0{,}3$.

Для расчета коэффициентов ассоциации и контингенции строится таблица сопряженности признаков, которая показывает связь между *двумя* явлениями, каждое из которых должно быть альтернативным.

Таблица 8.5

Группы по признаку y / Группы по признаку x	1	2	Итого
1	a	b	$a+b$
2	c	d	$c+d$
Итого	$a+c$	$b+d$	$a+b+c+d$

или

Таблица 8.6

Группы по признаку (y) / Группы по признаку (x)	1	2	Итого
1	a	c	$a+c$
2	b	d	$b+d$
Итого	$a+b$	$c+d$	$a+c+b+d$

Коэффициент контингенции всегда меньше коэффициента ассоциации, но оба изменяются от -1 до +1. При $ad\rangle bc$ связь прямая, при $ad\langle bc$ связь обратная, при $ad = bc$ связь отсутствует.

Пример 8.3.

Исследуем связь между участием в забастовках рабочих и уровнем их образования. Результаты обследования характеризуются следующими данными:

Таблица 8.7

Группы рабочих	Число рабочих		Итого
	Участвующих в забастовках	Не участвующих в забастовках	
Имеющие среднее образование	78 (a)	22 (y)	100 ($a+b$)
Не имеющие среднего образования	32 (c)	68 (d)	100 ($c+d$)
Итого	110 ($a+c$)	90 ($b+d$)	200 ($a+b+c+d$)

Коэффициент ассоциации:

$$K_a = \frac{ad-bc}{ad+bc} = \frac{78\cdot 68-32\cdot 22}{78\cdot 68+32\cdot 22} = \frac{4600}{6608} = 0{,}46 \approx 0{,}5$$

Связь подтверждается ($K_a = 0{,}5$). Она прямая и умеренная.

Коэффициент контингенции:

$$K_к = \frac{ad - bc}{\sqrt{(a+b)\cdot(b+d)\cdot(a+c)\cdot(c+d)}} = \frac{78\cdot 68 - 32\cdot 22}{\sqrt{100\cdot 90\cdot 110\cdot 100}} =$$

$$= \frac{4600}{9950} = 0{,}46 \approx 0{,}5$$

Связь подтверждается ($K_K \rangle 0{,}3$). Она прямая и умеренная.

Если по каждому из взаимосвязанных признаков число групп больше двух, то теснота связи между качественными признаками измеряется с помощью показателей взаимной сопряженности Пирсона и Чупрова.

Коэффициент взаимной сопряженности Пирсона определяется по формуле:

$$K_П = \sqrt{\frac{\varphi^2}{1+\varphi^2}}\text{, где } 1+\varphi^2 = \sum \frac{\sum \frac{n_{xy}^2}{n_y}}{n_x} \text{ или } 1+\varphi^2 = \sum \frac{\sum \frac{n_{xy}^2}{n_x}}{n_y}$$

Связь считается подтвержденной, если $K_П \geq 0{,}3$.

Коэффициент взаимной сопряженности Чупрова определяется по формуле:

$$K_Ч = \sqrt{\frac{\varphi^2}{\sqrt{(k_1-1)\cdot(k_2-1)}}},$$

где k_1 - количество групп по признаку x; k_2 -количество групп по признаку y.

Критерий тесноты связи $K_ч \geq 0{,}3$.

Для расчета коэффициентов Пирсона и Чупрова используется таблица сопряженности признаков, в которой количество групп по каждому признаку может быть более двух.

Таблица 8.8

Группы по признаку (y) / Группы по признаку (x)	1	2 и т.д.	Итого
1	n_{xy}	n_{xy}	n_x
2	n_{xy}	n_{xy}	n_x
3 и т.д.	n_{xy}	n_{xy}	n_x
Итого	n_y	n_y	n

Проверка значимости коэффициентов Пирсона и Чупрова осуществляется по критерию $\chi_p{}^2 = n\left[\sum\left(\frac{n_{xy}^2}{n_x \cdot n_y}\right) - 1\right]$, где n_{xy} – частота совместного появления признаков; n_x , n_y -суммы частот по строкам и столбцам соответственно; n -численность совокупности. Расчетное значение χ^2 должно быть больше табличного ($\chi_p^2 \rangle \chi_T{}^2$) при выбранном уровне вероятности. Формулы коэффициентов Пирсона и Чупрова через χ^2 будут соответственно:

Коэффициент Пирсона: $K_П = \sqrt{\frac{\chi^2}{\chi^2 + n}}$;

Коэффициент Чупрова $K_ч = \sqrt{\frac{\chi^2}{n \cdot \sqrt{(k_1 - 1) \cdot (k_2 - 1)}}}$.

Пример 8.4.

С помощью коэффициентов Пирсона и Чупрова исследуем связь между себестоимостью продукции и накладными расходами на реализацию.

Таблица 8.9

Накладные расходы	Себестоимость			Итого
	Низкая	Средняя	Высокая	
Низкие	19 n_{xy}	12 n_{xy}	9 n_{xy}	40 n_x
Средние	7 n_{xy}	18 n_{xy}	15 n_{xy}	40 n_x
Высокие	4 n_{xy}	10 n_{xy}	26 n_{xy}	40 n_x
Итого	30 n_y	40 n_y	50 n_y	120 n

Для определения коэффициентов Пирсона и Чупрова находим значение:

$$1+\varphi^2=\sum\frac{\sum\frac{n_{xy}^2}{n_x}}{n_y}=\frac{\frac{19^2}{30}+\frac{12^2}{40}+\frac{9^2}{50}}{40}+\frac{\frac{7^2}{30}+\frac{18^2}{40}+\frac{15^2}{50}}{40}+\frac{\frac{4^2}{30}+\frac{10^2}{40}+\frac{26^2}{50}}{40}=$$

$$=0{,}431+0{,}356+0{,}414=1{,}183 \qquad 1+\varphi^2=1{,}183 \qquad \varphi^2=0{,}183$$

Коэффициент взаимной сопряженности Пирсона будет равен:

$$K_П=\sqrt{\frac{\varphi^2}{1+\varphi^2}}=\sqrt{\frac{0{,}183}{1{,}183}}=\sqrt{0{,}155}=0{,}39\approx 0{,}4$$

Связь подтверждается $K_П>0{,}3$.

Коэффициент взаимной сопряженности Чупрова будет равен:

$$K_Ч=\sqrt{\frac{\varphi^2}{\sqrt{(k_1-1)\cdot(k_2-1)}}}=\sqrt{\frac{0{,}184}{\sqrt{(3-1)\cdot(3-1)}}}=\sqrt{\frac{0{,}184}{\sqrt{2\cdot 2}}}=0{,}3$$

Связь подтверждается $K_ч=0{,}3$.

Значимость коэффициентов Пирсона и Чупрова проверяем по критерию:

$$\chi_p^{\ 2}=n\left[\sum\left(\frac{n_{xy}^2}{n_x\cdot n_y}\right)-1\right]=$$

$$=120\cdot\left[\left(\frac{19^2}{40\cdot 30}+\frac{12^2}{40\cdot 40}+\frac{9^2}{40\cdot 50}+\frac{7^2}{40\cdot 30}+\frac{18^2}{40\cdot 40}+\frac{15^2}{40\cdot 50}+\frac{4^2}{40\cdot 30}+\frac{10^2}{40\cdot 40}+\frac{26^2}{40\cdot 50}\right)-1\right]=$$

$$=120\cdot(1{,}2-1)=24 \qquad \chi_p^2=24>\chi_m^2=9{,}49$$

Следовательно, связь между признаками «накладные расходы» и «себестоимость» - значима.

8.4. Методы собственно-корреляции

Все явления и процессы, характеризующие социально-экономическое развитие и составляющие единую систему национальных счетов, тесно взаимосвязаны и взаимозависимы между собой.

Корреляционная зависимость является частным случаем стохастической зависимости, при которой изменение значений факторных признаков ($x_1, x_2, ..., x_n$) влечет за собой изменение среднего значения результативного признака.

Корреляционная зависимость исследуется с помощью методов корреляционного и регрессионного анализов.

Корреляционный анализ имеет своей задачей количественное определение тесноты связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи). Теснота связи количественно выражается величиной коэффициентов корреляции.

Основной предпосылкой применения корреляционного анализа является необходимость подчинения значений всех факторных признаков и результативного нормальному закону распределения или близость к нему.

Если объем изучаемой совокупности достаточно большой ($n > 50$), то нормальность распределения может быть подтверждена на основе расчета и анализа, например, критерия Пирсона. Если $n < 50$, то закон распределения исходных данных определяется на базе построения и визуального анализа поля корреляции (графически).

Целью регрессионного анализа является оценка функциональной зависимости условного среднего значения результативного признака от факторных признаков. Он заключается в определении аналитического выражения связи.

Основной предпосылкой регрессионного анализа является то, что только результативный признак подчиняется нормальному закону распределения, а факторные признаки могут иметь произвольный закон распределения.

Уравнение регрессии, или статистическая модель связи социально-экономических явлений, выражаемая функцией $\overline{y_x} = f(x)$.

Теоретическая обоснованность моделей взаимосвязи, построенных на основе корреляционно-регрессионного анализа, обеспечивается соблюдением следующих основных условий:

1. Все признаки должны подчиняться нормальному закону распределения.

2. Отдельные наблюдения должны быть независимыми, т.е. между собой.

Практика выработала определенный критерий в определении оптимального числа факторов. Число факторных признаков должно быть в 5-6 раз меньше объема изучаемой совокупности.

По количеству включаемых факторов модели могут быть *однофакторными* и *многофакторными.*

Наиболее разработанной в теории статистики является методология так называемой *парной корреляции,* рассматривающая влияние вариации факторного признака *x*на результативный признак *y* и представляющая собой *однофакторный корреляционный* и *регрессионный анализ.* Овладение теорией и практикой построения и анализа двухмерной модели корреляционного и регрессионного анализа представляет собой исходную основу для изучения многофакторных стохастических связей.

Важнейшим этапом построения регрессионной модели (уравнения регрессии) является установление в анализе исходной информации математической функции. Сложность заключается в том, что из множества функций необходимо найти такую, которая лучше других выражает реально существующие связи между анализируемыми признаками.

По форме зависимости различают:

- *линейную* регрессию, которая выражается уравнением прямой (линейной функцией) $\overline{y_x} = a_0 + a_1 x$;

- *нелинейную* регрессию, которая выражается уравнениями вида:

- *гиперболы* - $\overline{y_x} = a_0 + \frac{a_1}{x}$;

- *параболы второго порядка* - $\overline{y_x} = a_0 + a_1 x + a_2 x^2$; и т. Д.

По направлению связи различают:

- *прямую* регрессию (положительную), возникающую при условии, если с увеличением или уменьшением значений факторного признака значения результативного признака также соответственно увеличиваются или уменьшаются;

- *обратную* (отрицательную) регрессию, появляющуюся при условии, что с увеличением или уменьшением значений факторного признака значения результативного признака соответственно уменьшаются или увеличиваются.

Определить тип уравнений можно, исследуя зависимость графически. В системе координат на оси абсцисс откладываются значения факторного признака, а на оси ординат – результативного. Каждое пересечение линий, проводимых через эти оси, обозначается точкой. При отсутствии тесных связей имеет место беспорядочное расположение точек на графике. Чем сильнее связь между признаками, тем теснее будут группироваться точки вокруг определенной линии, выражающей форму связи. Проиллюстрировать их графическое изображение можно рисунками 8.1 и 8.2.

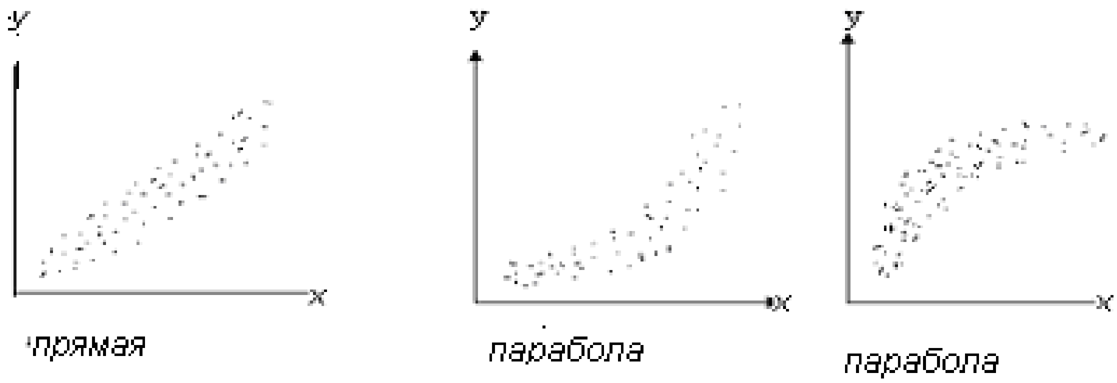


Рис. 8.1. Положительная регрессия.

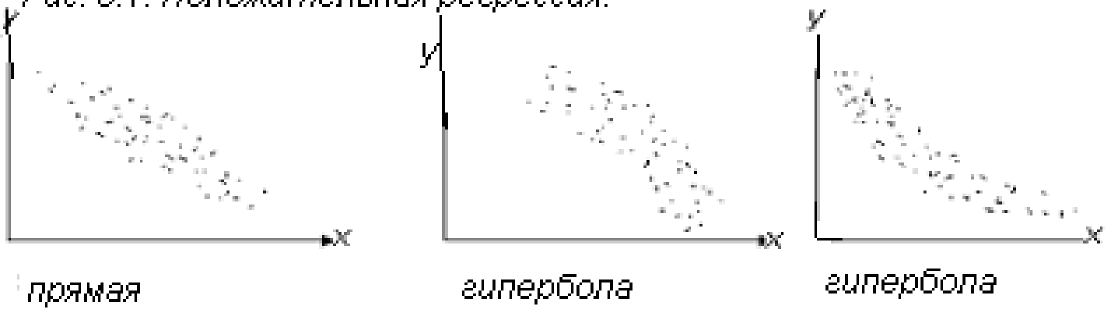


Рис. 8.2. Отрицательная регрессия.

Оценка параметров уравнений регрессии (a_0, a_1, a_2) осуществляется методом наименьшим квадратов, в основе которого лежит предположение о независимости наблюдений исследуемой совокупности.

Система нормальных уравнений для нахождения параметров линейной парной регрессии методом наименьших квадратов имеет следующий вид:

$$\left.\begin{aligned} n \cdot a_0 + a_1 \sum x &= \sum y \\ a_0 \sum x + a_1 \sum x^2 &= \sum xy \end{aligned}\right\},$$

где n – объем исследуемой совокупности (число единиц наблюдения).

В уравнениях регрессии параметр a_0 показывает усередненное влияние на результативный признак неучтенных (не выделенных для исследования) факторов; параметр (a_1) – коэффициент регрессии показывает, на сколько изменяется в среднем значение результативного признака при увеличении факторного на единицу собственного измерения.Параметр a_2 характеризует степень ускорения или замедления кривизны параболы и при $a_2 \rangle 0$ парабола имеет минимум, а при $a_2 \langle 0$ – максимум. Параметр a_1 характеризует крутизну кривой, а параметр a_0вершину кривой.

Коэффициент регрессии применяют для определения *коэффициента эластичности*, который показывает, на сколько процентов в среднем изменяется величина результативного признака *y* при изменении признака-фактора *x* на один процент.

Коэффициент эластичности определяется по формуле:

$$Э_x = a_1 \cdot \frac{\overline{x}}{\overline{y}}$$

Систему нормальных уравнений для нахождения параметров гиперболы можно представить следующим образом:

$$\left.\begin{array}{l} n \cdot a_0 + a_1 \sum \frac{1}{x} = \sum y \\ a_0 \sum \frac{1}{x} + a_1 \sum \frac{1}{x} = \sum \frac{y}{x} \end{array}\right\}$$

Система нормальных уравнений при параболической зависимости имеет следующий вид:

$$\left.\begin{array}{l} n \cdot a_0 + a_1 \sum x + a_2 \sum x^2 = \sum y; \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2 y \end{array}\right\}$$

Решив соответствующие системы уравнений, и найдя значения неизвестных коэффициентов a_0, a_1, a_2, получают уравнение регрессии. Затем определяются теоретические значения $\overline{y_x}$.

Измерение тесноты и направления связи является важной задачей изучения и количественного измерения взаимосвязи. Оценка тесноты связи между признаками предполагает определение меры соответствия вариации результативного признака от одного (при изучении парных зависимостей) или нескольких (множественных) факторов.

В случае наличия между двумя признаками линейной зависимости теснота связи измеряется *линейным коэффициентом корреляции*:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y} = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n} \cdot \frac{\sum y}{n}}{\sqrt{\overline{x^2} - \left(\bar{x}\right)^2} \cdot \sqrt{\overline{y^2} - \left(\bar{y}\right)^2}}$$

Линейный коэффициент корреляции изменяется от -1 до +1: $-1 \leq r \leq +1$.

Знаки коэффициентов регрессии и корреляции совпадают.

Значимость линейного коэффициента корреляции проверяется на основе *t*критерия Стьюдента:

$$t_P = r \cdot \sqrt{\frac{n-2}{1-r^2}}\ .$$

Если расчетное значение $t_p \rangle t_T$ (табличного), то это свидетельствует о значимости линейного коэффициента корреляции.

Пример 8.5.

Имеются данные о стоимости основных производственных фондов и объеме произведенной продукции необходимо определить уравнение связи и тесноту связи между объемом основных фондов и объемом произведенной продукции:

Таблица 8.10

Стоимость основных производственных фондов, млн. руб. (x)	Объем произведенной продукции, млн. руб. (y)	xy	x^2	y^2	$\overline{y_x}$
А	*Б*	1	2	3	4
1	20	20	1	400	19,4
2	25	50	4	625	25,0
3	31	93	9	961	30,6
4	31	124	16	961	36,2
5	40	200	25	1600	41,8
6	56	330	36	3136	47,4
7	52	364	49	2704	53,0
8	60	480	64	3600	58,6
9	60	540	81	3600	64,2
10	70	700	100	4900	69,8
55	445	2907	385	22487	446,0

Связь предполагается линейная, уравнение прямой $y_x = a_0 + a_1 x$.

Решаем систему уравнений методом наименьших квадратов:

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases} \quad \begin{cases} 10a_0 + 55a_1 = 445 \\ 55a_0 = 385a_1 = 2907 \end{cases}$$

$a_0 = 13{,}8$; $a_1 = 5{,}6$; $\bar{y}_x = 13{,}8 + 5{,}6x$

Подставляя вместо x соответствующие значения получаем значения $\overline{y_x}$.

Коэффициент регрессии a_1 свидетельствует о том, что при увеличении объема основных фондов на 1 млн. руб. количество произведенной продукции увеличится на 5,6 млн. руб.

Тесноту связи определяем по линейному коэффициенту корреляции:

$$r=\frac{\overline{xy}-\bar{x}\cdot\bar{y}}{\sigma_x\cdot\sigma_y}=\frac{\overline{xy}-\bar{x}\cdot\bar{y}}{\sqrt{\overline{x^2}-\left(\bar{x}\right)^2}\cdot\sqrt{\overline{y^2}-\left(\bar{y}\right)^2}}=\frac{\frac{2907}{10}-\frac{55}{10}\cdot\frac{446}{10}}{\sqrt{\frac{385}{10}-5{,}5^2}\cdot\sqrt{\frac{22487}{10}-44{,}6^2}}=$$

$$=\frac{45{,}4}{2{,}9\cdot 16{,}1}=\frac{45{,}4}{46.7}=0{,}972$$

Следовательно, связь прямая и очень тесная.

Значимость линейного коэффициента корреляции проверяем по критерию Стьюдента:

$$t_P=r\cdot\sqrt{\frac{n-2}{1-r^2}}=0{,}972\cdot\sqrt{\frac{10-2}{1-0{,}972^2}}=11{,}723.$$

Табличное значение критерия Стьюдента при уровне значимости 0,05 (приложение 3*)* $t_T=2{,}262$, т.е. меньше t_p. Следовательно, с вероятностью 0,995 можно утверждать, что найденная теснота связи значима.

По сгруппированным данным в случае линейной и нелинейной зависимости между двумя признаками для измерения тесноты связи применяют *корреляционное отношение*.

Эмпирическое корреляционное отношение рассчитывается по данным группировки по формуле:

$$\eta_э=\sqrt{\frac{\delta_y^2}{\sigma_y{}^2}}\text{, где}$$

δ_y^2 – межгрупповая дисперсия результативного признака;

$\sigma_y{}^2$ – общая дисперсия результативного признака.

Корреляционное отношение изменяется в пределах от 0 до 1 $\left(0\le\eta_э\le 1\right)$.

Подкоренное выражение корреляционного отношения представляет собой *коэффициент детерминации:*

($\eta^2 = \frac{\delta_y^2}{\sigma_y^2}$), который показывает долю вариации результативного признака под влиянием вариации признака-фактора.

Для оценки значимости уравнения регрессии в целом, особенно при нелинейных зависимостях, используют F-критерий Фишера.

Проверка значимости коэффициента детерминации осуществляется также по F-критерию Фишера, расчетное значение которого:

$$F_p = \frac{\eta^2 \cdot (n-m)}{(1-\eta^2)\cdot(m-1)},$$

где n-число наблюдений, а m-число признаков (при парной корреляции $m=2$). Вычисленные значения F_p сравнивается с критическим (табличным) F_T для принятого уровня значимости и чисел степеней свободы $v_1 = m-1$ и $v_2 = n-m$. Значимость подтверждается, если $F_p \rangle F_T$.

Пример 8.6.

По сельхозпредприятиям имеются данные об урожайности и количестве внесенных минеральных удобрений. Необходимо выявить зависимость урожайности от количества внесенных минеральных удобрений.

Таблица 8.11

№№ с/х предприятия	Внесено удобрений на 1 га, ц (x)	Урожайность ц\га (y)	x^2	x^3	x^4	xy	x^2y	$\bar{y}_x$	y^2
1	2	3	4	5	6	7	8	9	10
1	0,4	14	0,16	0,064	0,0256	5,8	2,24	14,98	196
2	0,5	16	0,25	0,125	0,0625	8,0	4,00	17,11	256
3	0,5	19	0,25	0,125	0,0625	9,5	4,75	17,11	361
…	…	…	…	…	…	…	…	…	…
29	1,4	32	1,96	2,744	3,8416	44,8	62,72	30,02	1024
30	1,5	30	2,25	3,375	5,0625	45,0	67,50	30,77	900
Итого	30,0	750	32,90	38,484	47,0762	791,1	899,95	750,0	20400

Произведем выравнивание по параболе второго порядка:

$$\overline{y_x} = a_0 + a_1 x + a_2 x^2;$$

Решаем систему нормальных уравнений:

$$\begin{cases} n \cdot a_0 + a_1 \sum x + a_2 \sum x^2 = \sum y, \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy, \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2 y \end{cases}$$

$$\begin{cases} 30a_0 + 30a_1 + 32{,}90a_2 = 750{,}00, \\ 30a_0 + 32{,}90a_1 + 38{,}484a_2 = 781{,}10, \\ 32{,}90a_0 + 38{,}484a_1 + 47{,}0762a_2 = 899{,}95 \end{cases}$$

Решение этой системы уравнений методом наименьших квадратов дает следующие значения параметров:

$a_0 = 5{,}086 \qquad a_1 = 27{,}511 \qquad a_2 = -6{,}927$

$\bar{y}_0 = 5{,}086 + 27{,}511x - 6{,}927x^2$

Т.к. связь криволинейная для определения тесноты связи используется корреляционное отношение:

$\eta_Э = \sqrt{\delta_y^2 / \sigma_y^2}$, где общая дисперсия признака y, $\sigma_y^2 = \overline{y^2} - (\bar{y})^2$, а межгрупповая дисперсия $\delta_y^2 = \sum (y_i - \bar{y})^2 / \sum f_i$

Общая дисперсия определятся по исходным (не сгруппированным) данным, т.е. $\sigma_y^2 = \overline{y^2} - (\bar{y})^2 = \dfrac{20400}{30} - \left(\dfrac{750}{30}\right)^2 = 680 - 625 = 55$

Для нахождения межгрупповой дисперсии необходимо произвести группировку предприятий по количеству внесенных минеральных удобрений.

Результат группировки данных по стоимости основных фондов представлен в нижеследующей таблице:

Таблица 8.12

Группы с/х предприятий по количеству внесенных минеральных удобрений на 1 га, ц.	Количество предприятий	Средняя урожайность, ц/га
1	2	3
0,3 – 0,5	4	14
0,5 – 0,7	5	16
0,7 – 0,9	9	32
0,9 – 1,1	6	29
1,1 - 1,3	4	28
1,3 и более	2	20

В данной задаче факторный признак количество минеральных удобрений (x), а результативный – урожайность (y). Межгрупповая дисперсия $\delta_y^2 = \sum(\overline{y}_i - \bar{y})^2 / \sum f_i$; δ_y^2 вычисляем по данным группировки в вышеизложенной таблице $\overline{y_i}$ – средняя урожайность в каждой группе, т.е.

$$\bar{y}_1 = 14; \quad \bar{y}_2 = 16; \bar{y}_3 = 32; \quad \bar{y}_4 = 29; \overline{y_5}. = 28; \quad \overline{y_6} = 20.$$

Общая средняя признака $\bar{y} = 750/30 = 25$, f_i - число предприятий в каждой группе, $\sum f_i = 30$ заводам.

Таблица 8.13

$\bar{y}_i$	f_i	$(\bar{y}_i - \bar{y})$	$(\bar{y}_i - \bar{y})^2$	$(\bar{y}_i - \bar{y})^2 f_i$
1	2	3	4	5
14	4	-11	121	484
16	5	-9	81	405
32	9	7	49	441
29	6	4	16	96
28	4	3	9	36
20	2	-5	25	50
Итого	30	*X*	*x*	1512

Определяем межгрупповую дисперсию $\delta_y^2 = \frac{1512}{30} = 50{,}4$. Теперь можно вычислить корреляционное отношение:

$$\eta_Э = \sqrt{\frac{50{,}4}{55}} = \sqrt{0{,}916} = 0{,}957$$

Это означает, что связь между количеством внесенных минеральных удобрений и урожайностью тесная. Значимость корреляционного отношения проверяем по F критерию Фишера:

$$F_p = \frac{\eta^2 \cdot (n-m)}{(1-\eta^2)\cdot(m-1)} = \frac{0{,}957^2 \cdot (30-2)}{(1-0{,}957^2)\cdot(2-1)} = \frac{25{,}644}{0{,}084} = 305{,}29$$

Расчетное значение F критерия (приложение 3) при уровне значимости 0,05 равно 41,34. Следовательно, с вероятностью 0,995 можно утверждать, что теснота связи между урожайностью и количеством внесенных минеральных удобрений значима.

Для качественной оценки тесноты связи на основе показателя эмпирического корреляционного отношения можно воспользоваться соотношениями Чэддока :

$\eta_Э$	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-0,99
Сила связи	Слабая	Умеренная	Заметная	Тесная	Весьма тесная

ГЛАВА 9. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНАМИКИ

9.1. Понятие о рядах динамики. Виды рядов динамики

Динамикой в статистике называют изменение явления во времени. Для отображения динамики строят *ряды динамики*, которые представляют собой ряды изменяющихся во времени значений статистического показателя, расположенных в хронологическом порядке.

Ряд динамики состоит из двух элементов: *момента или периода времени*, к которым относятся данные, и статистических показателей *(уровней)*. Оба элемента вместе образуют *члены ряда*.

Уровни ряда обычно обозначаются через *« y »,* а периоды или моменты времени, к которым относятся значения показателя, - через *« t »*.

Существуют различные виды рядов динамики. В зависимости от *способа выражения* уровней ряда динамики они подразделяются на ряды динамики: *абсолютных величин* (табл. 9.1, строка 1), *относительных величин* (табл. 9.1, строка 2) *и средних величин* (табл. 9.1, строка 3).

Таблица 9.1

Показатели численности и заработной платы населения Российской Федерации

№ п/п	Показатели	2010 г.	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.
1	2	3	4	5	6	7	8
1.	Численность населения (на 1 января; тысяч человек)	142,8	142,9	143,0	143,3	143,7	146,3
2.	Удельный вес городского населения в общей численности населения (на 1 января; в процентах)	73,5	73,7	73,9	74,0	74,2	74,5

1	2	3	4	5	6	7	8
3.	Среднемесячная номинальная начисленная заработная плата работающих в экономике, (тыс. руб.)	20,9	23,4	26,6	29,8	32,5	34,0

По длительности времени, к которым относятся уровни ряда, ряды динамики делятся на моментные и интервальные.

В *моментных рядах* каждый уровень характеризует явления *на момент времени* (табл. 9.1. строки 1; 2).

В *интервальных рядах* динамики каждый уровень ряда характеризует явление *за период времени* (табл. 9.1. строка 3).

В интервальных рядах динамики абсолютных величин уровни ряда можно суммировать и получить общую величину за ряд следующих друг за другом периодов. В моментных рядах эта сумма не имеет смысла.

Ряды динамики могут быть с равными и неравными интервалами. Понятие интервала в моментных и интервальных рядах различное. *Интервал моментного ряда* – это период времени от одной даты до другой даты, на которые приведены данные. Если это данные о численности населения на 1 января, то интервал равен от 1 января одного года, до 1 января другого года. *Интервал интервального ряда* – это период времени, за который обобщены данные. Если это среднемесячная заработная плата по годам, то интервал равен одному году.

Интервал ряда может быть *равным* и *неравным* как в моментных, так и в интервальных рядах динамики.

С помощью рядов динамики определяют скорость и интенсивность развития явлений, выявляют основную тенденцию их развития, выделяют сезонные колебания, сравнивают развитие во времени отдельных показателей разных стран, выявляют связи между развивающимися во времени явлениями.

9.2. Сопоставимость уровней ряда динамики и рядов динамики

При построении динамических рядов следует помнить, что уровни его должны быть сопоставимы между собой, т.к. для несопоставимых величин невозможно вести расчеты показателей динамики.

Уровни ряда динамики могут быть несопоставимы по следующим причинам:

- *несопоставимость по территории* (изменение границ). В этом случае старые (прежние) данные пересчитывают в новые границы, о чем делается оговорка; Для достижения сопоставимости уровней в этом случае используют прием *смыкания рядов динамики*.

- *несопоставимость вследствие различных единиц измерения и единиц счета.* Нельзя, например, сравнивать производство тканей в погонных метрах и в квадратных метрах.

- *несопоставимость по методологии учета или расчета показателей*. Обычно для достижения сопоставимости прежние показатели пересчитывают по новой методологии, о чем делается оговорка.

- *несопоставимость по кругу охватываемых объектов*, которая возникает вследствие ряда организационных причин, например, перехода объектов из одного подчинения в другое. В этом случае сопоставимость достигается также *смыканием рядов динамики* (табл. 9.2).

Смыканием рядов динамики называют объединение в один ряд (более длинный) двух или нескольких рядов динамики, уровни которых исчислены по разной методологии или в разных территориальных единицах, или охватывающих различное количество объектов. Сопоставимый ряд при этом можно получить в абсолютных величинах и можно в относительных.

Пример 9.1.

По объединению имеются данные динамики объема продукции. Необходимо для сопоставимости рядов динамики между собой произвести смыкание.

Таблица 9.2

Годы	2010 г.	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.
Объем продукции, тыс. руб. По двум предприятиям По трем предприятиям	 1910 -	 1970 -	 2000 -	 2120 2280	 - 2360	 - 2450
Сомкнутый (сопоставимый) ряд абсолютных величин, тыс. руб.	2053	2118	2200	2150	2360	2450
Сомкнутый (сопоставимый) ряд относительных величин, % к 2013 г.	90,1	92,9	94,3	100,0	103,5	107,5

Для получения сомкнутого ряда в *абсолютных* величинах рассчитывают коэффициент пересчета предыдущих уровней $K_n = \frac{\textit{новый уровень}}{\textit{прежний уровень}} = \frac{2280}{2120} = 1{,}075$, на который затем *предыдущие* уровни умножают: $1910 \cdot 1{,}075 = 2053$; $1970 \cdot 1{,}075 = 2118$; $2000 \cdot 1{,}075 = 2150$; а последующие остаются неизменными. Для получения сомкнутого ряда в *относительных* величинах период, в который произошло изменение, принимают за 100 %. Это и будет базой сравнения. Только для предыдущих уровней – это прежний (старый) уровень, а для последующих уровней - новый уровень. Так в табл. 9.2. уровни 2010-2012 гг. определяют деление на уровень 2013 г. равный 2120 тыс. руб. Например 2010 г. $\frac{1910}{2120} \cdot 100 = 90{,}1\,\%$, а 2015 г. $\frac{2450}{2280} \cdot 100 = 107{,}5\,\%$.

Иногда возникает проблема *сопоставимости рядов динамики* между собой: сопоставление тенденции развития явления различных показателей; при параллельном анализе развития во времени одинаковых показателей, но относящихся к различным объектам, например, странам. В этом случае ряды приводят *к одному основанию*, т.е. к одному и тому же периоду или моменту времени, принятому за базу сравнения. В этом случае характер развития выступает более наглядно.

Пример 9.2.

Предположим, имеются данные о производстве сахара в двух регионах, тыс. тонн.

Таблица 9.3

Годы	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.
Регион А	45,5	72,4	95,2	122.0	128,8
Регион Б	56,1	65,1	66,5	65.0	67,0

В абсолютных уровнях не видны особенности производства в регионах.

Приведем ряды к одному основанию, т.е. примем за постоянную базу сравнения уровни 2011 года (100 %) и получим следующие данные в процентах к 2011 году.

Таблица 9.4

Годы	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.
Регион А	100,0	159,1	209,2	268,1	281,3
Регион Б	100,0	116,0	118,5	115,9	119,4

В относительных величинах различный характер развития выступает более наглядно.

9.3. Показатели изменения уровней ряда динамики

Аналитические показатели уровней ряда динамики получаются в результате сравнения уровней ряда между собой. При этом сравниваемый уровень называется *текущим,* а тот, с которым происходит сравнение - *базисным.*

При сравнении каждого последующего уровня с каждым предыдущим получаются *цепные* показатели. При сравнении каждого последующего уровня с одним уровнем (базой) получаются *базисные* показатели. Выбор базы сравнения должен быть обоснован экономически.

К показателям изменения уровней ряда относятся: абсолютный прирост темпа роста, темп прироста, абсолютное значение 1 % прироста.

Абсолютный прирост (Δy) характеризует размер увеличения (или уменьшения) уровня за определенный промежуток времени. Он равен разности сравниваемых уровней и выражает абсолютную скорость изменения:

$$\Delta y = y_n - y_{n-K_Э}$$

где y_n – любой уровень ряда, кроме первого (текущий), а $y_{n-K_Э}$ - базисный уровень.

Если $k=1$, то y_{n-k} – предыдущий уровень и все абсолютные приросты будут цепными. Если $k \neq 1$, то абсолютные приросты будут базисными. Следовательно, цепной абсолютный прирост рассчитывается по формуле $\Delta y_ц = y_n - y_{n-1}$, а базисный – по формуле $\Delta y_б = y_n - y_1$. Абсолютный прирост показывает, на сколько единиц сравниваемый уровень больше (если «+») или меньше (если «-») уровня, принятого за базу сравнения. Между цепными и базисными абсолютными приростами есть взаимосвязь: сумма цепных абсолютных приростов равна базисному абсолютному приросту. $\Delta y_б = \sum \Delta y_ц$

Темп роста (T_p) - показывает во сколько раз текущий уровень ряда больше (или меньше) базисного уровня. Он равен отношению сравниваемых уровней, выраженному в процентах.

При $k=1$ T_p - цепные, а при $k \neq 1$ - базисные. Темп роста равен коэффициенту роста, умноженному на 100. Следовательно, темп роста – это

$$T_p = \frac{y_n}{y_{n-к}} \cdot 100 = K_p \cdot 100.$$

Цепной темп роста $T_{р.ц} = \frac{y_n}{y_{n-1}} \cdot 100$, а базисный $T_{р.б} = \frac{y_n}{y_1} \cdot 100$.

Между цепными и базисными коэффициентами роста существует взаимосвязь: произведение цепных коэффициентов роста равно базисному коэффициенту роста. $К_{р.б} = ПК_{р.ц}$.

Следовательно, $Т_{р.б} = ПК_{р.ц} \cdot 100\%$

Темп прироста ($Т_{пр}$) показывает, на сколько процентов уровень текущий больше (или меньше) базисного уровня:

$$Т_{пр} = \frac{y_n - y_{n-к}}{y_{n-k}} \cdot 100\%$$

Он также может быть цепным и базисным:

цепной: $Т_{пр.ц} = \dfrac{y_n - y_{n-1}}{y_{n-1}} \cdot 100\%$

базисный $Т_{пр.б} = \dfrac{y_n - y_1}{y_1} \cdot 100\%$

Между темпом роста и темпом прироста существует взаимосвязь:

$$Т_{пр} = \frac{y_n - y_{n-k}}{y_{n-k}} \cdot 100\,\% = \frac{y_n}{y_{n-k}} \cdot 100\,\% - 100\,\% = Т_р - 100\,\%$$

Для сравнительного анализа динамики двух пространственных объектов коэффициент опережения (отставания). *Коэффициент опережения (отставания*) представляет собой отношение базисных темпов (коэффициентов) роста двух сравниваемых рядов динамики:

$$К_{оп} = \frac{Т'_{р.б.}}{Т''_{р.б.}}$$

По данным таблицы 9.4. производство сахара в регионе А опережает производство сахара в регионе Б в $\dfrac{281{,}3}{119{,}4} = 2{,}4\,раза$.

Абсолютное значение 1 % прироста $(А1\%)$ получается в результате сравнения абсолютного прироста и темпа прироста за один и тот же промежуток времени:

$$A1\% = \frac{\Delta y_ц}{T_{пр.ц}} \text{ или } (y_n - y_{n-1}) : \frac{y_n - y_{n-1}}{y_{n-1}} \cdot 100\ \% = \frac{y_{n-1}}{100\%},$$

т.е. равно 1 % предыдущего уровня.

Этот показатель имеет смысл лишь для цепных показателей. Он позволяет видеть, что замедление темпов прироста часто не сопровождается уменьшением абсолютных приростов и наоборот.

Рассмотрим расчет всех перечисленных показателей изменения уровней ряда динамики на условном примере 9.3.

Таблица 9.5

Динамика производства газа в регионе за 2011-2015 гг.

Годы	Производство газа, млн.м^3. y	Абсолютный прирост, млн.м^3		Темп роста, %		Темп прироста, %		Абсолютное значение 1% прироста $A1\% = \frac{\Delta y_ц}{T_{пр.ц}}$ или $\frac{y_{n-1}}{100\%}$
		цепной $\Delta y_ц = y_n - y_{n-1}$	базисный $\Delta y_б = y_n - y_1$	Цепной $T_{р.ц} = \frac{y_n}{y_{n-1}} \cdot 100\%$	базисный $T_{р/б} = \frac{y_n}{y_1} \cdot 100\%$	цепной $T_{пр.ц} = \frac{\Delta y_ц}{y_{n-1}} \cdot 100\%$ или $T_{р.ц} - 100\%$	базисный $T_{пр.б} = \frac{\Delta y_ц}{y_1} \cdot 100\%$ или $T_{рбц} - 100\%$	
1	2	3	4	5	6	7	8	9
2011	289	-	-	-	100,0	-	-	-
2012	321	32	32	111,1	111,1	11,1	11,1	2,9
2013	346	25	57	107,8	119,7	7,8	19,7	3,2
2014	372	26	83	107,5	128,7	7,5	28,7	3,4
2015	407	35	118	109,4	140,8	9,4	40,8	3,7
Итого	1735	118	Х	Х	Х	Х	Х	Х

Как видно из таблицы, производство газа за 5 лет увеличилось на 118 млн. м^3 или на 40,8 %. Ежегодные темпы роста до 2015 года снижались, а в 2015 году выросли на 9,4 %. Положительным является то, что содержание

1 % прироста все время увеличивалось.

9.4. Средние характеристики ряда динамики

Средние характеристики ряда динамики охватывают изменение явления за весь период, к которому относится ряд динамики. К средним характеристикам относятся: средний уровень ряда, средний абсолютный прирост, средний темп роста и средний темп прироста.

Средний уровень ряда $(\overline{y})$ показывает, какова средняя величина уровня характерная для всего периода ряда. Средний уровень ряда исчисляется по-разному для интервальных и моментных рядов.

Для интервального ряда с равным интервалом, он определяется по средней арифметической простой, делением суммы уровней ряда на число периодов:

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n},$$

где $\sum_{i=1}^{n} y_i$ - сумма уровней ряда, n - число периодов.

По данным таблицы 9.5 среднее годовое производство газа за 2011-2015 годы составило $\overline{y} = \frac{1735}{5} = 347\ млн.\ м^3$.

Для интервального ряда с неравным интервалом средний уровень ряда определяется по формуле средней арифметической взвешенной:

$$\overline{y} = \frac{\sum y_i t_i}{\sum t_i},$$ где t_i - величина интервала.

Например, имеются сведения об изменении численности работников фирмы за апрель. На 1 апреля в списках фирмы числилось 50 человек. 10 апреля было принято 6 человек, 20 апреля уволен 1 человек, 25 принято 3 человека. До конца месяца изменений не было. Средняя численность работников фирмы составила:

$$\bar{y} = \frac{50\cdot 9 + 56\cdot 10 + 55\cdot 5 + 58\cdot 6}{30} = \frac{1633}{30} = 54 \text{ чел.}$$

Для моментного ряда с равным интервалом средний уровень ряда определяется по формуле средней хронологической:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + y_3 + ... + \frac{1}{2}y_n}{n-1}$$

Пример 9.4.

По данным таблицы 9.1 о численности населения РФ за 2010-2015 гг определим среднюю годовую численность населения за 5 лет.

$$\bar{y} = \frac{\frac{142{,}8}{2} + 142{,}9 + 143{,}0 + 143{,}3 + 143{,}7 + \frac{146{,}3}{2}}{6-1} =$$

$$\frac{717{,}45}{5} = 143{,}49 \text{ тыс.чел.} \approx 144 \text{ млн. чел.}$$

Для моментного ряда с неравным интервалом средний уровень ряда можно определить по формуле средней скользящей взвешенной:

$$\bar{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + ... + (y_{n-1} + y_n)t_{n-1}}{2\sum_{i=1}^{n-1} t_i}$$

№ п/п	Показатели	2010 г.	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.
1	2	3	4	5	6	7	8
1.	Численность населения (на 1 января; тысяч человек)	142,8	142,9	143,0	143,3	143,7	146,3
2.	Удельный вес городского населения в общей численности населения (на 1 января; в процентах)	73,5	73,7	73,9	74,0	74,2	74,5

1	2	3	4	5	6	7	8
3.	Среднемесячная номинальная начисленная заработная плата работающих в экономике, (тыс. руб.)	20,9	23,4	26,6	29,8	32,5	34,0

Пример 9.5. Предположим, имеются сведения о численности населения РФ на 1 января только 2010 года, 2012 года и 2015 года (таблица 9.1). Тогда среднегодовая численность населения за период 2010-2015 гг. составила:

$$\bar{y} = \frac{(142{,}8 + 143{,}0) \cdot 2 + (143{,}0 + 146{,}3) \cdot 3}{2 \cdot 5} = \frac{1439{,}5}{10} = 143{,}95 \text{ } \textit{млн. чел.}$$

Средний абсолютный прирост характеризует скорость развития явления во времени. Его можно определить как среднюю величину из цепных абсолютных приростов:

$$\overline{\Delta y} = \frac{\sum_{i=1}^{m} y_i}{m},$$

где m – число цепных абсолютных приростов.

Либо по данным уровней ряда: $\overline{\Delta y} = \dfrac{y_n - y_1}{n - 1}$, т.к. сумма цепных абсолютных приростов всегда равна базисному абсолютному приросту.

Средний темп роста дает сводную характеристику интенсивности изменения явления за весь период ряда динамики. Он может быть определен по формуле средней геометрической на основании данных о цепных коэффициентах роста:

$$\overline{T_p} = \sqrt[m]{K_{p_1} \cdot K_{p_2} \cdot \ldots \cdot K_{p_m}} \cdot 100\%,$$

где m – число темпов роста.

Либо на основании данных об уровнях ряда:

$\overline{T_p} = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\%$, т.к. произведение цепных коэффициентов роста всегда равно базисному коэффициенту роста.

Эта формула ценна тем, что позволяет определить средний темп роста при отсутствии нескольких или всех промежуточных данных.

Средний темп прироста определяется на основании данных о среднем темпе роста как разность: $\overline{T_{np}} = \overline{T_p} - 100\%$.

Пример 9.6.

По данным таблицы 9.5 средний абсолютный прирост, средний темп роста и прироста будут соответственно равны.

$$\overline{\Delta y} = \frac{\sum_{i=1}^{m} y_i}{m} = \frac{32 + 25 + 26 + 35}{4} = \frac{118}{4} = 29{,}5 \textit{ млн.м}^3 \quad \textit{или}$$

$$\overline{\Delta y} = \frac{y_n - y_1}{n-1} = \frac{407 - 289}{5-1} = \frac{118}{4} = 29{,}5 \textit{ млн.м}^3$$

$$\overline{T_p} = \sqrt[m]{K_{p_1} \cdot K_{p_2} \cdot \ldots \cdot K_{p_m}} \cdot 100\% =$$

$$\sqrt[4]{1{,}111 \cdot 1{,}078 \cdot 1{,}075 \cdot 1{,}094} \cdot 100\% = 108{,}9\%$$

или $\overline{T_p} = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\% = \sqrt[5-1]{\frac{407}{289}} \cdot 100\% = 108{,}9\%$

$$\overline{T_{np}} = \overline{T_p} - 100\% = 108{,}9\% - 100\% = 8{,}9\%$$

9.5. Выявление основной тенденции динамических рядов

Одним из методов анализа и обобщения динамических рядов является выявление его основной тенденции - *тренда.* В статистической практике выявление основной тенденции развития осуществляют двумя способами: сглаживанием и аналитическим выравниванием.

Сглаживание – это механическое выравнивание отдельных членов ряда динамики с использованием фактических значений соседних уровней.

Аналитическое выравнивание - это выравнивание с применением кривой, проведенной между конкретными фактическими уровнями таким образом, чтобы она отражала тенденцию, присущую ряду, освобождая его от незначительных колебаний.

Сглаживание может осуществляться методом *укрупнения интервала*, т.е., например, ряд суточного выпуска продукции заменить рядом ежемесячного выпуска продукции. Таким образом, сглаживаются суточные колебания выпуска.

Сглаживание методом *простой скользящей средней*, заключается в том, что вычисляется средний уровень из трех, пяти, семи и т.д. уровней. Таким образом, вместо каждого уровня ряда берутся средние из окружающих его уровней с обеих сторон. В этой средней сглаживаются случайные отклонения. Она будет скользящей, поскольку период осреднения все время меняется. Из него вычитается один предыдущий и прибавляется один следующий. Например, скользящая средняя из 3-х уровней будет:

$$\overline{y_1} = \frac{y_1 + y_2 + y_3}{3}, \ \overline{y_2} = \frac{y_2 + y_3 + y_4}{3} \text{ и т.д.}$$

Средняя скользящая относится в этом случае ко 2-му, 3-му, 4-му и т.д. периоду.

Если скользящая средняя находится по четному число членов, то для отнесения ее к конкретному периоду необходимо произвести *центрирование,* т.е. найти среднюю из двух смежных скользящих средних.

Недостаток метода простой скользящей средней в том, что сглаженный ряд динамики сокращается (укорачивается) для начала и конца.

Пример 9.7.

Необходимо выявить тенденцию урожайности зерновых культур. Используем простейший способ скользящей средней (трехчленной).

Таблица 9.6

Динамика урожайности зерновых культур за 2001-2015 гг.

Год	Урожайность, ц/га y	Трехчленная скользящая средняя $\bar{y}$
1	2	3
2001	13,8	-
2002	12,1	13,3
2003	14,0	13,1
2004	13,2	14,3
2005	15,6	14,7
2006	15,4	15,0
2007	14,0	15,7
2008	17,6	15,7
2009	15,4	14,6
2010	10,9	14,6
2011	17,5	14,5
2012	15,0	17,0
2012	18,5	15,9
2014	14,2	15,9
2015	14,9	-
Итого	222,0	Х

Рассчитываем трехчленную скользящую среднюю (за 2001-2003 годы):

$$\bar{y} = \frac{13{,}8 + 12{,}1 + 14{,}0}{3} = 13{,}3\ ц/га$$

Этот результат записываем в середину, т.е. против 2002 года.

Затем сдвигаем уровни на один и считаем трехчленную скользящую среднюю за 2002-2004 годы: $\bar{y} = \frac{12{,}1 + 14{,}0 + 13{,}2}{3} = 13{,}1\ ц/га$

Результат записываем против 2003 года и т.д. Как видно из таблицы. Приближенно видно, что урожайность растет, т.е. наблюдается тенденция к росту.

Аналитическое выравнивание предполагает представление уровней данного ряда динамики в виде функции времени: $y = f(t)$.

В практике экономических исследований применяется аналитическое выравнивание по любому рациональному многочлену.

Правильно установить тип кривой, тип аналитической зависимости от времени - является одной из трудных задач статистики. К этому следует подходить с большой осторожностью. Аналитическое выравнивание состоит в подборе для данного ряда динамики теоретической кривой, наилучшим образом описывающей эмпирические данные. Это могут быть различные функции: полиномы степени, экспоненты, логистические кривые и другие виды.

Полиномы имеют следующий вид:

полином первой степени $\overline{y_t} = a_0 + a_1 t$ (прямая);

полином второй степени $\overline{y_t} = a_0 + a_1 t + a_2 t^2$ (парабола 2-го порядка);

полином n-ой степени $\overline{y_t} = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n$.

$a_0; a_1; a_2; ... a_n$ параметры полиномов, t – условное обозначение времени. Параметр a_0 трактуется как характеристика средних условий ряда динамики, параметры $a_1; a_2; ... a_n$ как изменения ускорения.

Наиболее приближенный и простой способ определения формы теоретической кривой – графический.

После выбора вида уравнения необходимо определить параметры уравнения. Самый распространенный способ определения параметров уравнения - это метод наименьших квадратов.

Система нормальных уравнений для оценивания параметров прямой:

$\overline{y_t} = a_0 + a_1 t$ примет вид: $\begin{cases} na_0 + a_1 \sum t = \sum y, \\ a_0 \sum t + a_1 \sum t^2 = \sum yt \end{cases}$

В статистической практике применяется упрощенный расчет параметров уравнения, который заключается в переносе начала отсчета времени в середину ряда динамики. Тогда $\sum t = 0$ и система нормальных уравнений упрощается для прямой: $\begin{cases} na_0 = \sum y, \\ a_1 \sum t^2 = \sum yt \end{cases}$

Аналогично для параболы второго порядка система нормальных уравнений будет:

$$\begin{cases} na_0 + a_1 \sum t + a_2 \sum t^2 = \sum y, \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum yt, \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum yt^2 \end{cases}$$

После упрощения:

$$\begin{cases} na_0 + a_2 \sum t^2 = \sum y, \\ a_1 \sum t^2 = \sum yt, \\ a_0 \sum t^2 + a_2 \sum t^4 = \sum yt^2 \end{cases}$$

т.к. суммы всех нечетных степеней t будут равны нулю.

Решив системы относительно неизвестных параметров, получим величины параметров соответствующих уравнений. Подставляя вместо t значения времени, получим теоретические значения $\overline{y}_t$, которые будут отражать тенденцию. Рассмотрим пример 9.8. На основании данных таблицы 9.6 произведем выравнивание по прямой.

Таблица 9.7

Динамика урожайности зерновых культур за 2001-2015 годы

Годы	Урожайность, ц/га	t	t^2	$y \cdot t$	$\overline{y_t}$
2001	13,8	-7	49	-95,9	13,3
2002	12,1	-6	36	-72,6	13,8
2003	14,0	-5	25	-70,0	13,9
2004	13,2	-4	16	-52,8	14,1
2005	15,6	-3	9	-46,8	14,3
2006	15,4	-2	4	-30,8	14,5
2007	14,0	-1	1	-14,0	14,6
2008	17,6	0	0	0	14,8
2009	15,4	1	1	15,4	15,0
2010	10,9	2	4	21,8	15,1
2011	17,5	3	9	52,5	15,3
2012	15,0	4	16	60,0	15,5
2013	18,5	5	25	92,5	15,7
2014	14,2	6	36	85,2	15,8
2015	14,9	7	49	104,3	16,0
Итого	222,0	-	280	48,8	222,0

Для упрощения расчетов переносим отсчет времени (t) в середину ряда, т.е. середина – это 2008 год. Ставим здесь ноль, от которого время назад будет -1 год, -2 год, -3 год и т.д., а вперед будет +1 год, + 2 год и т.д. Тогда $\sum t = 0$.

Для нахождения параметров уравнения прямой: $\overline{y_t} = a_0 + a_1 t$.

Решаем систему уравнений: $\begin{cases} na_0 + a_1 \sum t = \sum y, \\ a_0 \sum t + a_1 \sum t^2 = \sum yt \end{cases}$

Т.к. $\sum t = 0$ уравнение примет вид:

$$\begin{cases} na_0 = \sum y, \\ a_1 \sum t^2 = \sum yt \end{cases}$$

Отсюда $a_0 = \frac{\sum y}{n} = \frac{222}{15} = 14{,}8$ $\quad a_1 = \frac{\sum yt}{\sum t^2} = \frac{48{,}8}{280} = 0{,}17$ и уравнение примет вид $\overline{y_t} = 14{,}8 + 0{,}17t$. Подставляя последовательно вместо t его значения (из таблицы) получим выравненные значения $\overline{y_t}$.

Отразим на графике фактические и выравненные значения урожайности.

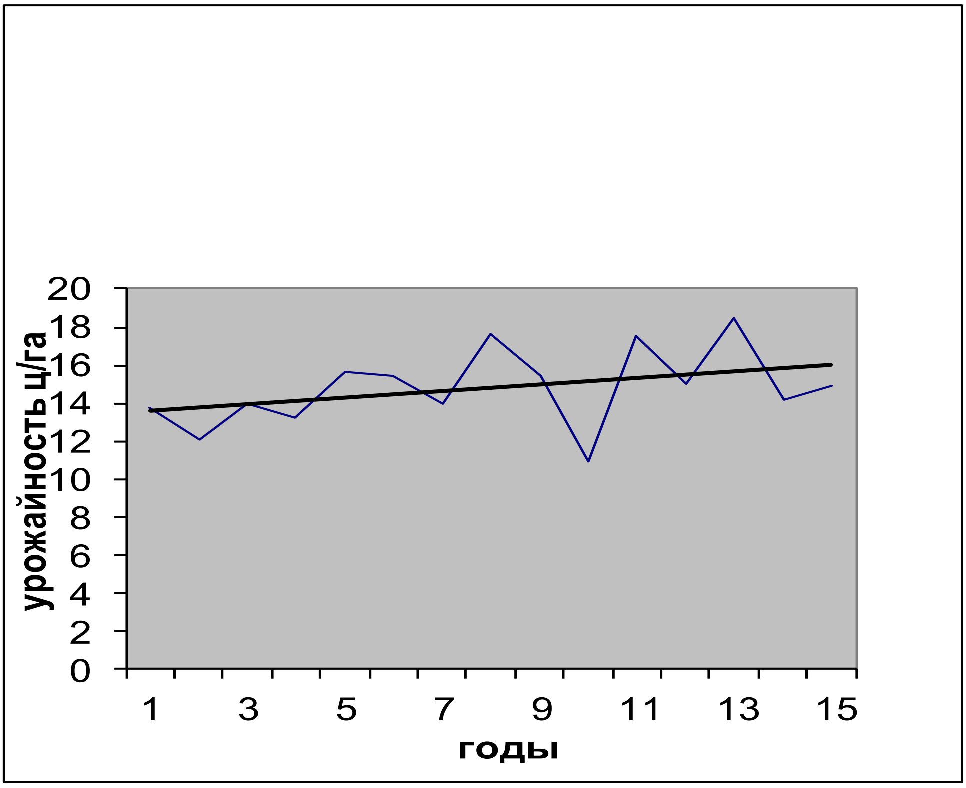


Рис. 9.1. Динамика урожайности зерновых культур за 2001-2015 гг.

9.6. Изучение сезонных колебаний

При анализе квартальных или месячных данных многих социально-экономических явлений обнаруживаются повторяющиеся колебания, которые не изменяются длительный период времени. Они являются результатом действия природно-климатических условий, общих экономических факторов и других экономических факторов, частично регулируемых. В статистике такие колебания называются *сезонными.* Это особый тип динамики.

Сезонность можно понимать как внутригодовую динамику вообще. Сезонность может возникать в производствах, связанных с переработкой сельхозсырья, в торговле из-за сезонного характера спроса на товары и т.д.

Глубину сезонных колебаний измеряют коэффициентом сезонности или индексом сезонности, который представляет собой отношение средней из фактических уровней одноименных месяцев к средней из выровненных данных по тем же месяцам. $K_c = \frac{\overline{y_м}}{\overline{\overline{y_t}}}$.

Следовательно, величина коэффициента сезонности зависит от способа выравнивания. Если это способ *средней арифметической*, то $\bar{y}_t = \frac{\sum \overline{y_м}}{n}$, где $\bar{y}_M = \frac{\sum y}{m}$, m - число лет, n число месяцев = 12. Если $\bar{y}_t$ - это 12 месячная скользящая средняя, то это способ *скользящей средней*. Если $\bar{y}_t$ – получен аналитическим выравниванием – способ *аналитического выравнивания*. Коэффициенты сезонности отклоняются в обе стороны от единицы.

Рассмотрим на примере расчет коэффициентов сезонности методом средней арифметической.

Пример 9.9.

Предположим, имеются данные о количестве расторгнутых браков в городе за 3 года по месяцам.

Таблица 9.8

Месяц	Число расторгнутых браков			Число расторг-нутых браков в среднем за три го-да $(\bar{y}_M)$	Коэффици-ент сезонности $\frac{\overline{y_M}}{\overline{y_t}}$
	2013 г.	2014 г.	2015 г.		
1	2	3	4	5	6
Январь	195	158	144	166	1,220
Февраль	164	141	136	147	1,080
Март	153	153	146	151	1,110
Апрель	136	140	132	136	1,000
Май	136	136	136	136	1,000
Июнь	123	129	125	126	0,926
Июль	126	128	124	126	0,926
Август	121	122	119	121	0,890
Сентябрь	116	118	118	118	0,868
Октябрь	126	130	128	128	0,941
Ноябрь	129	131	135	132	0,971
Декабрь	138	141	139	139	1,022
Средний уровень ряда $(\bar{y}_г)$	139	136	132	$\overline{y_t} = 136$	1,000

Находим средние месячные уровни числа расторгнутых браков по месяцам:

за январь $\bar{y}_M = \dfrac{195+158+144}{3} = 166$

за февраль $\overline{y_M} = \dfrac{164+141+136}{3} = 147$ и т.д.

Средний выравненный уровень способом средней арифметической

$$\overline{y_t} = \frac{\sum \overline{y_M}}{12} = \frac{166+147+151+\ldots+139}{12} = 136$$

или

$$\bar{y}_t = \frac{\sum \overline{y_г}}{m} = \frac{139+136+132}{3} = 136.$$

Тогда коэффициенты сезонности будут:

в январе: $K_c = \frac{\overline{y_м}}{\bar{\bar{y}}_t} = \frac{166}{136} = 1{,}220$,

в феврале: $K_c = \frac{\overline{y_м}}{\bar{\bar{y}}_t} = \frac{147}{136} = 1{,}080$ и т. д.

В приведенном примере наблюдается наименьшее количество расторгнутых браков в августе к зиме оно увеличивается и в январе достигает наибольшей величины.

9.7. Экстраполяция и интерполяция

Исследование динамики и характеристика основной тенденции динамических рядов дают основание для прогнозирования будущих размеров уровня экономического явления.

Статистические методы прогнозирования основаны на предположении, что закономерность развития, действовавшая в прошлом (внутри ряда динамики), сохраняется и в прогнозируемом периоде. Определение прогнозируемых уровней на основании тенденции, сложившейся внутри ряда динамики, называется *экстраполяцией.* Экстраполяция, проводимая в будущее. Называется *перспективной*, а в прошлое *ретроспективной.* Чаще используют перспективную экстраполяцию.

В зависимости от принципов положенных в основу прогноза, и исходных данных можно использовать элементарные методы экстраполяции, которые основаны на показателях *среднего абсолютного прироста, среднего темпа роста,* и экстраполяции на основе *выравнивания рядов по какой-либо функции.*

Прогнозирование *по среднему абсолютному приросту* можно применить в том случае, когда есть уверенность в равномерном изменении уровня (под равномерностью понимается стабильность абсолютных приростов). Тогда перспективную экстраполяцию можно сделать по формуле:

$$\hat{y}_{i+t} = y_i + \overline{\Delta y} \cdot t$$, где

$\hat{y}_{i+t}$ -экстраполируемый уровень;

($i+t$)- номер этого уровня (года);

i-номер последнего уровня (года) исследуемого периода, за который рассчитан $\overline{\Delta y}$;

t -срок прогноза (период упреждения);

$\overline{\Delta y}$ средний абсолютный прирост.

Прогнозирование по среднему темпу роста можно осуществлять в том случае. Когда есть основание считать. Что общая тенденция ряда характеризуется показательной кривой. Тогда перспективный экстраполируемый уровень находится по формуле:

$$\hat{y}_{i+t} = y_i \cdot \overline{K_p}^{\,t}$$, где

y_i -последний уровень ряда динамики;

t - срок прогноза;

$\overline{K_p}$ - средний коэффициент роста.

Если ряду динамики свойственна иная закономерность, значения получатся приближенными.

Наиболее распространенным методом прогнозирования является *аналитическое выражение тренда*. Для определения прогнозируемого уровня в этом случае достаточно выйти за пределы значения независимой переменной времени t. При таком подходе к прогнозированию предполагается, что размер уровня явления формируется под воздействием множества факторов, выделить которые невозможно. В связи с этим ход развития явления связывают не с конкретными факторами, а со временем, т.е. $y = f(t)$.

При анализе рядов динамики иногда приходится прибегать к определению некоторых неизвестных уровней внутри ряда динамики, т.е. к *интерполяции.* Интерполяция может производиться на основе формул экстраполяции. При интерполяции считается, что ни выявленная тенденция, ни ее характер не претерпели существенных изменений в том промежутке времени, уровень которого нам не известен. Рассмотрим расчет экстраполируемого перспективного уровня различными способами.

Пример 9.10.

На основе данных таблицы 9.6 определим будущий уровень 2017 года.

1. По среднему абсолютному приросту:

Средний абсолютный прирост в исследуемом ряду составил

$$\overline{\Delta y} = \frac{y_n - y_1}{n-1} = \frac{14{,}9 - 13{,}8}{15-1} = 0{,}08\ ц/га,$$

а прогнозируемый уровень урожайности в 2017 году:

$$\hat{y}_{i+t} = y_i + \overline{\Delta y} \cdot t = y_{2015+2} = y_{2015} + \overline{\Delta y} \cdot 2 = 14{,}9 + 0{,}08 \cdot 2 = 15{,}1\ ц/га$$

2. По среднему коэффициенту роста:

Средний коэффициент роста в исследуемом ряду составил:

$$\overline{K_p} = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[15-1]{\frac{14{,}9}{13{,}8}} = \sqrt[14]{1{,}0797101} = 1{,}0055,$$

а прогнозируемый уровень урожайности в 2017 году:

$$\hat{y}_{i+t} = y_i \cdot \overline{K_p}^{\,t} = y_{2015+2} = y_{2015} \cdot \overline{K_p}^{\,2} = 14{,}9 \cdot 1{,}0055^2 = 16{,}2\ ц/га$$

3. Методом аналитического выравнивания по прямой уровень 2017 года будет равен:

$$\overline{y_t} = 14{,}8 + 0{,}17t = 14{,}8 + 0{,}17 \cdot 9 = 16{,}3ц/га$$

Результаты расчетов свидетельствуют о том, что прогноз на основе среднего абсолютного прироста дает осторожную оценку, а по среднему коэффициенту роста и аналитическому выравниванию близки.

ГЛАВА 10. ИНДЕКСНЫЙ МЕТОД

10.1. Понятие об индексах

Слово "index" латинское и означает "показатель", "указатель". В статистике *под индексом понимается* обобщающий количественный показатель, выражающий соотношение двух совокупностей, состоящих из элементов, непосредственно не поддающихся суммированию. Например, объем продукции предприятия в натуральном выражении суммировать нельзя (кроме однородной), а для обобщающей характеристики объема это необходимо. Нельзя суммировать цены на отдельные виды продукции и т.д. Для обобщающей характеристики таких совокупностей в динамике, в пространстве и по сравнению с планом применяются индексы. Кроме сводной характеристики явлений индексы позволяют дать оценку роли отдельных факторов в изменении сложного явления. Индексы используются и для выявления структурных сдвигов в народном хозяйстве.

Индексы рассчитываются как для сложного явления (общие или сводные), так и для отдельных его элементов (индивидуальные индексы).

В индексах, характеризующих изменение явления во времени различают базисный и отчетный (текущий) периоды. *Базисный период* – это период времени, к которому относится величина, принятая за базу сравнения. Обозначается он подстрочным знаком *"0"*.*Отчетный период* – это период времени, к которому относится величина, подвергающаяся сравнению. Обозначается он подстрочным знаком *"1"*.

Индивидуальные индексы представляют собой отношение отдельных элементов совокупности. Это обычная относительная величина.

Пример 10.1.

Если цена товара в текущем периоде 30 руб., а в базисном была 25 руб., то индивидуальный индекс будет равен $\frac{30\,руб.}{25\,руб.} = 1{,}2$ или 120 %.

Сводный индекс - характеризует изменение всей сложной совокупности в целом, т.е. состоящей из несуммируемых элементов. Следовательно, чтобы рассчитать такой индекс надо преодолеть несуммарность элементов совокупности. Это достигается введением дополнительного показателя *(соизмерителя).* Сводный индекс состоит из двух элементов: индексируемой величины и веса.

Индексируемая величина – это показатель, для которого рассчитывается индекс. *Вес (соизмеритель*) – это дополнительный показатель, вводимый для целей соизмерения индексируемой величины. В сводном индексе в числителе и знаменателе всегда *сложная совокупность, выраженная суммой произведений индексируемой величины и веса*.

В зависимости от объекта исследования как общие, так и индивидуальные индексы подразделяются на индексы *объемных (количественных) показателей* (физического объема продукции, посевной площади, численности рабочих и др.) и *индексы качественных показателей* (цены, себестоимости, урожайности, производительности труда, заработной платы и др.).

В зависимости от базы сравнения индивидуальные и общие индексы могут быть *цепными и базисными.*

В зависимости от методологии расчета общие индексы имеют *две формы: агрегатную и форму среднего индекса.*

10.2. Агрегатная форма индекса

Агрегатная форма сводного индекса является основной. От нее происходят все остальные сводные индексы.

В дальнейшем изложении будут использованы следующие обозначения:

i – индивидуальный индекс;

I – общий (сводный) индекс;

x – обобщенная характеристика качественного показателя;

d – обобщенная характеристика количественного показателя.

"x" может принимать значения:

p – цена единицы товара (продукции);

z – себестоимость единицы товара (продукции);

y – урожайность отдельной культуры;

f – заработная плата;

w – выработка продукции одним человеком в единицу времени;

t – трудоемкость продукции.

"d" может принимать значения:

q – физический объем товара (продукции);

$П$ – посевная площадь;

T – численность рабочих или работников (затраты труда).

Для построения сводных индексов в агрегатной форме следует помнить следующие правила:

1. В индексе изменяется только индексируемая величина и всегда от отчетного периода (в числителе) к базисной (в знаменателе). Исключение – индекс производительности труда по трудоемкости.

2. Вес (соизмеримость) остается неизменным, т.е. одинаковым в числителе и знаменателе (кроме случая, когда индексируемой величиной является все произведение).

3. В индексах качественных показателей индексируемая величина качественный показатель *("x")*, а весом является количественный показатель *("d")*, который берется неизменным в числителе и знаменателе на уровне отчетного периода *("1")*.

4. В индексах количественных показателей индексируемая величина – количественный показатель *("d")*, а весом является качественный показатель *("x")*, который берется неизменным в числителе и знаменателе на уровне базисного периода *("0")*.

5. При записи сводного индекса на первом месте (первым сомножителем) пишется индексируемая величина, а на втором – вес (правило не строгое, но необходимое во избежание механических ошибок).

6. Изменение изучаемого явления в абсолютном выражении определяется как разность числителя и знаменателя сводного индекса (исключение - индекс производительности труда по трудоемкости).

Тогда *индексы всех качественных индексов* (кроме исключения) в общем виде можно записать в виде формулы:

$$I_x = \frac{\sum x_1 d_1}{\sum x_0 d_1},$$

а изменение в абсолютном выражении как разность:

$$\Delta x = \sum x_1 d_1 - \sum x_0 d_1 .$$

Качественные индексы конкретных показателей:

1. Индекс цен $I_p = \dfrac{\sum p_1 q_1}{\sum p_0 q_1}$,

где $\sum p_1 q_1$ - товарооборот (или стоимость произведенной продукции) отчетного периода;

$\sum p_0 q_1$ - товарооборот (стоимость продукции) отчетного периода в базисных ценах.

$\Delta x = \sum p_1 q_1 - \sum p_0 q_1$ характеризует изменение товарооборота (стоимости продукции) за счет цен "+" - увеличение, "- " уменьшение.

2. Индекс себестоимости $I_z = \dfrac{\sum z_1 q_1}{\sum z_0 q_1}$,

где $\sum z_1 q_1$ - издержки (затраты или себестоимость всей продукции) отчетного периода;

$\sum z_0 q_1$ - издержки (затраты или себестоимость всей продукции) базисного периода в пересчете на фактический объем.

$\Delta z = \sum z_1 q_1 - \sum z_0 q_1$ характеризует экономию, если “- “ от снижения себестоимости или дополнительные издержки (затраты) от роста себестоимости, если "+".

3. Индекс урожайности $I_y = \dfrac{\sum y_1 П_1}{\sum y_0 П_1}$,

где $\sum y_1 П_1$ - валовой сбор отчетного (текущего) периода;

а $\sum y_0 П_1$ - валовой сбор с площади отчетного периода при базисной урожайности.

$\Delta y = \sum y_1 П_1 - \sum y_0 П_1$ свидетельствует об увеличении валового сбора, если "+", и об уменьшении валового сбора за счет снижения урожайности, если "-".

4. Индекс заработной платы $I_f = \dfrac{\sum f_1 T_1}{\sum f_0 T_1}$,

где $\sum f_1 T_1$ - фонд оплаты труда отчетного периода;

$\sum f_0 T_1$ - базисный фонд оплаты труда в пересчете на отчетную численность рабочих (работников).

$\Delta f = \sum f_1 T_1 - \sum f_0 T_1$ характеризует экономию фонда оплаты труда за счет снижения уровня зарплаты, если "- " и перерасход фонда оплаты труда за счет роста зарплаты, если "+".

5. Индекс производительности труда по выработке $I_w = \frac{\sum w_1 T_1}{\sum w_0 T_1}$,

где $\sum w_1 T_1$ - количество продукции отчетного периода,

$\sum w_0 T_1$ - объем продукции отчетного периода при базисной производительности труда.

$\Delta w = \sum w_1 T_1 - \sum w_0 T_1$ увеличение объема продукции за счет роста производительности труда, если «+», уменьшение объема продукции за счет снижения производительности труда, если «–».

6. Индекс производительности труда по трудоемкости (исключение).

$$I_w = \frac{\sum t_0 q_1}{\sum t_1 q_1},$$

где $\sum t_0 q_1$ - общие затраты труда базисного периода в пересчете на фактический объем продукции;

$\sum t_1 q_1$ - общие затраты труда на выпуск продукции отчетного периода.

$\Delta t = \sum t_1 q_1 - \sum t_0 q_1$ свидетельствует об экономии труда за счет роста производительности труда, если «–», дополнительных затратах труда за счет снижения его производительности, если «–».

7. Индекс трудоемкости $I_t = \frac{\sum t_1 q_1}{\sum t_0 q_1}$

$\Delta t = \sum t_1 q_1 - \sum t_0 q_1$ изменение затрат за счет изменения трудоемкости.

Индексы количественных показателей в общем виде: $I_d = \dfrac{\sum d_1 x_0}{\sum d_0 x_0}$,

а изменение в абсолютном выражении: $\Delta d = \sum d_1 x_0 - \sum d_0 x_0$.

Конкретные количественные индексы:

1. Индекс физического объема в зависимости от исходной информации может иметь три различных веса.

Если весом является цена, то:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0},$$

где $\sum q_1 p_0$ - отчетный объем продукции (или товарооборота) в базисных ценах;

$\sum q_0 p_0$ - товарооборот (или стоимость произведенной продукции) базисного периода.

$\Delta q = \sum q_1 p_0 - \sum q_0 p_0$ дает представление об увеличении (если "+") или уменьшении (если "-") товарооборота (стоимости продукции) за счет соответственно увеличения или уменьшения физического объема продукции (товара).

Если весом является себестоимость, то:

$$I_q = \frac{\sum q_1 z_0}{\sum q_0 z_0}$$

$\Delta q = \sum q_1 z_0 - \sum q_0 z_0$ свидетельствует об увеличении (если "+") или уменьшении (если "-") издержек (затрат или себестоимости всей продукции) за счет соответственно увеличения или уменьшения физического объема продукции.

Если весом является трудоемкость, то:

$$I_q = \frac{\sum q_1 t_0}{\sum q_0 t_0}.$$

$\Delta q = \sum q_1 t_0 - \sum q_0 t_0$ характеризует изменение затрат труда за счет изменения физического объема продукции.

2. Индекс посевной площади: $I_П = \frac{\sum П_1 y_0}{\sum П_0 y_0}$

$\Delta y = \sum П_1 y_0 - \sum П_0 y_0$ показывает изменение валового сбора за счет изменения размера посевных площадей.

3. Индекс численности рабочих (работников) также может быть рассчитан в двух вариантах в зависимости от веса.

Если весом является выработка, то:

$$I_T = \frac{\sum T_1 w_0}{\sum T_0 w_0}$$

$\Delta T = \sum T_1 w_0 - \sum T_0 w_0$ характеризует изменение объема продукции за счет изменения численности.

Если весом является заработная плата, то:

$$I_T = \frac{\sum T_1 f_0}{\sum T_0 f_0}$$

$\Delta T = \sum T_1 f_0 - \sum T_0 f_0$ показывает экономию (если "-") или перерасход (если "+") фонда оплаты труда за счет соответственно сокращения или увеличения численности рабочих (работников).

Если индексируемой величиной является вся сложная совокупность (товарооборот, валовой сбор, фонд оплаты труда, издержки производства и т.д.), то оба сомножителя в числителе отчетного периода, а в знаменателе базисного периода.

В общем виде: $I_{xd} = \dfrac{\sum x_1 d_1}{\sum x_0 d_0}$.

1. Индекс товарооборота: $I_{pq} = \dfrac{\sum p_1 q_1}{\sum p_0 q_0}$

2. Индекс издержек (затрат) на производство: $I_{zq} = \dfrac{\sum z_1 q_1}{\sum z_0 q_0}$

3. Индекс валового сбора: $I_{УП} = \dfrac{\sum У_1 П_1}{\sum У_0 П_0}$

4. Индекс фонда оплаты труда $I_{fT} = \dfrac{\sum f_1 T_1}{\sum f_0 T_0}$

5. Индекс затрат труда: $I_{tq} = \dfrac{\sum t_1 q_1}{\sum t_0 q_0}$

Разность числителя и знаменателя индекса $\Delta xd = \sum x_1 d_1 - \sum x_0 d_0$ характеризует общее изменение сложной совокупности.

10.3. Взаимосвязь индексов связанных явлений

Между отдельными индексам существуют взаимосвязи, позволяющие на основе одних индексов определять другие. Одной из таких взаимосвязей является взаимосвязь индексов связанных явлений.

Большинство экономических явлений, изучаемых с помощью индексов, связаны между собой. *Между индексами существует точно такая же взаимосвязь, как и между показателями, которые они отражают*. Например, т.к. товарооборот - это произведение цены на количество товара, то и индекс товарооборота равен произведению индексов цен и физического объема товарооборота:

$$I_{pq} = I_p \cdot I_q \text{, т.е. } \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_1}{\sum p_0 q_1} \cdot \frac{\sum q_1 p_0}{\sum q_0 p_0}$$

В абсолютном выражении эта взаимосвязь выглядит в следующем виде: $\Delta pq = \Delta p + \Delta q$

$$\sum p_1 q_1 - \sum p_0 q_0 = (\sum p_1 q_1 - \sum p_0 q_1) + (\sum q_1 p_0 - \sum q_0 p_0)$$

Аналогично запишем взаимосвязь связанных явлений в общем виде:

$$I_{xd} = I_x \cdot I_d \text{, } \frac{\sum x_1 d_1}{\sum x_0 d_0} = \frac{\sum x_1 d_1}{\sum x_0 d_1} \cdot \frac{\sum d_1 x_0}{\sum d_0 x_0}$$

$$\Delta xd = \Delta x + \Delta d$$

$$\sum x_1 d_1 - \sum x_0 d_0 = (\sum x_1 d_1 - \sum x_0 d_1) + (\sum d_1 x_0 - \sum d_0 x_0)$$

$$I_{zq} = I_z \cdot I_q \qquad \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{\sum z_1 q_1}{\sum z_0 q_1} \cdot \frac{\sum z_1 p_0}{\sum z_0 p_0}; \; \Delta zq = \Delta z + \Delta q$$

$$I_{yП} = I_y \cdot I_П \qquad \frac{\sum У_1 П_1}{\sum У_0 П_0} = \frac{\sum У_1 П_1}{\sum У_0 П_1} \cdot \frac{\sum У_1 П_0}{\sum У_0 П_0};$$

$$\Delta yП = \Delta y + \Delta П;$$

$$I_{fT} = I_f \cdot I_T \qquad \frac{\sum f_1 T_1}{\sum f_0 T_0} = \frac{\sum f_1 T_1}{\sum f_0 T_1} \cdot \frac{\sum f_1 T_0}{\sum f_0 T_0};$$

$$\Delta fT = \Delta f + \Delta T;$$

$$I_T = I_t \cdot I_q \,,$$

так как $T = tq$ $\;, \frac{\sum t_1 q_1}{\sum t_0 q_0} = \frac{\sum t_1 q_1}{\sum t_0 q_1} \cdot \frac{\sum t_1 p_0}{\sum t_0 p_0}\,;\; \Delta T = \Delta t + \Delta q\,;$

$$I_q = I_w \cdot I_T \,,\;\; \text{так как} \;\; q = wT\,, \frac{\sum w_1 T_1}{\sum w_0 T_0} = \frac{\sum w_1 T_1}{\sum w_0 T_1} \cdot \frac{\sum T_1 w_0}{\sum T_0 w_0}\,;$$

$$\Delta q = \Delta w + \Delta T\,.$$

Пример 10.2.

Имеются данные о продаже различных товаров одной товарной группы:

Товар	Январь		Февраль	
	количество, q_0	цена за ед., руб., p_0	количество, q_1	цена за ед.,руб., p_1
1	2	3	4	5
A	2500	15	3000	12
$Б$	200	50	250	55

Необходимо определить:

а) индивидуальные индексы цен и физического объема товарооборота**;**

б) общий индекс цен;

в) общий индекс физического объема товарооборота;

г) индекс товарооборота;

д) покажите взаимосвязь исчисленных индексов.

Расчет произведем в табличной форме, проставив необходимые обозначения:

Таблица 10.1

Индивидуальный индекс		Товарооборот		
цен $i_p = \frac{p_1}{p_0}$	физического объема, $i_q = \frac{q_1}{q_0}$	отчетного периода, руб., p_1q_1	базисного периода, руб., p_0q_0	отчетного периода в ценах базисного периода, руб., p_0q_1
1	2	3	4	5
0,800	1,200	36000	37500	45000
1,100	1,250	13750	10000	12500
Итого	x	49750	47500	57500

$$\text{б)}\quad I_p = \frac{\sum p_1q_1}{\sum p_0q_1} = \frac{49750}{57500} = 0{,}865 \quad \text{или } 86{,}5\%$$

$$\Delta_p = \sum p_1q_1 - \sum p_0q_1 = 49750 - 57500 = -7750\ (руб.)$$

$$\text{в)}\quad I_q = \frac{\sum q_1p_0}{\sum q_0p_0} = \frac{57500}{47500} = 1{,}210 \text{ или } 121{,}0\%$$

$$\Delta_q = \sum q_1p_0 - \sum q_0p_0 = 57500 - 47500 = 10000\ (руб.)$$

$$\text{г)}\quad I_{pq} = \frac{\sum p_1q_1}{\sum p_0q_0} = \frac{49750}{47500} = 1{,}047 \text{ или } 104{,}7\%$$

$$\Delta_{pq} = \sum p_1q_1 - \sum p_0q_0 = 49750 - 47500 = 2250(руб)$$

$$\text{д)}\quad I_{pq} = I_p \cdot I_q \qquad 1{,}047 = 0{,}865 \cdot 1{,}210$$

$$\Delta_{pq} = \Delta_p + \Delta_q \qquad 2250 = -7750 + 10000$$

Выводы:

Товарооборот товарной группы в текущем периоде по сравнению с базисным увеличился на 2250 рублей или на 4,7 %. За счет роста физического объема товарооборот увеличился на 10000 рублей или на 21 %, а за счет снижения цен уменьшился на 7750 рублей или на 13,5 %.

10.4. Форма среднего индекса

Сводный индекс может быть исчислен *как средняя величина из индивидуальных индексов.* Форма среднего индекса используется в тех случаях, когда в агрегатной форме индекс на основе имеющейся информации рассчитать невозможно. Однако форму средней для этого нужно выбрать таим образом, чтобы полученный *средний индекс был бы тождественен исходному агрегатному индексу*. В практике статистики в большинстве случаев принято все *количественные индексы* рассчитывать как *средние арифметические*, а все *качественные* как *средние гармонические*.

Выведем *средний арифметический индекс* из агрегатного в общем виде:

$$I_d = \frac{\sum d_1 x_0}{\sum d_0 x_0} = \frac{\sum i_d d_0 x_0}{\sum d_0 x_0},$$

т.к. $i_d = \dfrac{d_1}{d_0}$. Отсюда $d_1 = i_d d_0$.

Аналогично записываются все конкретные количественные индексы:

1. Индекс физического объема продукции:

$$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0} \text{ или } I_q = \frac{\sum i_q q_0 z_0}{\sum q_0 z_0}, \text{ или } I_q = \frac{\sum i_q q_0 t_0}{\sum q_0 t_0}.$$

2. Индекс посевной площади: $I_П = \dfrac{\sum i_П П_0 y_0}{\sum П_0 y_0}$;

3. Индекс численности: $I_T = \dfrac{\sum i_T T_0 f_0}{\sum T_0 f_0}$ или $I_T = \dfrac{\sum i_T T_0 w_0}{\sum T_0 w_0}$;

Выведем средний гармонический индекс из агрегатного в общем виде:

$$I_x = \frac{\sum x_1 d_1}{\sum x_0 d_1} = \frac{\sum x_1 d_1}{\sum \frac{x_1 d_1}{i_x}}, \text{ т.к. } i_x = \frac{x_1}{x_0}. \text{ Отсюда } x_0 = \frac{x_1}{i_x}$$

Аналогично записываются все качественные индексы (кроме исключения):

1. Индекс цен: $I_p = \dfrac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}}$

2. Индекс себестоимости: $I_z = \dfrac{\sum z_1 q_1}{\sum \frac{z_1 q_1}{i_z}}$

3. Индекс урожайности: $I_y = \dfrac{\sum y_1 \varPi_1}{\sum \frac{y_1 \varPi_1}{i_y}}$

4. Индекс заработной платы: $I_f = \dfrac{\sum f_1 T_1}{\sum \frac{f_1 T_1}{i_f}}$

5. Индекс производительности труда по выработке: $I_w = \dfrac{\sum w_1 T_1}{\sum \frac{w_1 T_1}{i_w}}$

Исключение.

Индекс производительности труда по трудоемкости (индекс Струмилина):

$$I_w = \frac{\sum t_0 q_1}{\sum t_1 q_1} = \frac{\sum i_w t_1 q_1}{\sum t_1 q_1} = \frac{\sum i_w T_1}{\sum T_1}, \text{ т.к. } i_w = \frac{t_0}{t_1}.$$

Отсюда: $t_0 = i_w t_1$.

Численные значения индексов производительности труда в обоих случаях будут одинаковыми. Изменение же явления в абсолютном выражении определяется, так же как и в агрегатной форме, разностью числителя и знаменателя индекса (исключение индекс производительности труда по трудоемкости:

$$\Delta t = \sum t_1 q_1 - \sum i_w t_1 q_1 = \sum T_1 - \sum i_w T_1 \ .$$

Пример.10.3.

Предположим, известен товарооборот базисного периода и изменение в %% количества отдельных товаров необходимо определить как изменился в среднем физический объем товарооборота.

Таблица 10.2

Товары	Товарооборот базисного периода, тыс. руб., $p_0 q_0$	Изменение количества проданныхтоваров, % (изм.)	Индивидуальный индекс физического объёма товарооборота, $i_q = \frac{100\% \pm ИЗМ}{100}$	Товарооборот отчетного периода в ценах базисного периода, тыс. руб., $p_0 q_1 = iq \cdot q_0 p_0$
1	2	3	4	5
А	600	+15	1,15	690
Б	200	+10	1,10	220
Итого	800	*х*	*х*	910

$$I_q = \frac{\sum i_q \cdot q_0 p_0}{\sum q_0 p_0} = \frac{910}{800} = 1{,}137 \ \text{или} \ 113{,}7\%$$

Или на $\Delta_q = \sum i_q \cdot q_0 p_0 - \sum q_0 p_0 = 910 - 800 = 110(\text{тыс.руб.})$

(В среднем физический объем товарооборота увеличился на 110 тыс. руб. или на 13,7%).

Пример.10.3.

Предположим известна выручка магазина за отчетный период и изменение цен на отдельные товары. Необходимо определить изменение цен по магазину в целом.

Таблица 10.3

Товары	Товарооборот отчетного периода, тыс. руб., p_1q_1	Изменение цен, % (изм.)	Индивидуальные индексы цен, $i_q = \frac{100\% \pm Изм.}{100}$	Товарооборот отчетного периода в ценах базисного периода, тыс. руб., $p_0q_1 = \frac{p_1q_1}{i_p}$
1	2	3	4	5
А	1800,0	+10	1,10	1636,4
Б	300,0	+25	1,25	240,0
Итого	2100,0	Х	Х	1876,4

$$I_p = \frac{\sum p_1q_1}{\sum \frac{p_1q_1}{i_p}} = \frac{2100,0}{1876,4} = 1,119 \text{ } или \text{ } 111,9\%$$

(средний гармонический индекс)

$$\Delta_p = \sum p_1q_1 - \sum \frac{p_1q_1}{i_p} = 2100,0 - 1876,4 = 223,6 \ (тыс.руб.)_1$$

В целом по магазину цены на товары выросли на 11,9%, что привело к увеличению товарооборота на 223,6 тыс. руб.

Пример 10.4.

Известно изменение производительности труда по цехам предприятия. Необходимо определить изменение производительности труда по предприятию в целом имея информацию о численности рабочих по цехам.

Таблица 10.4

Цеха предприятий	Численность рабочих в отчетном периоде, T_1 *(затраты труда отчетные)*	Индекс динамики производительности труда, i_w	Затраты труда на продукцию отчетного периода по базисной трудоемкости, чел./час $i_w \cdot T_1$
1	2	3	4
№1	122	1,051	128
№2	190	1,033	196
№3	200	1,006	201
Итого	512	X	525

$$I_w = \frac{\sum i_w \cdot T_1}{\sum T_1} = \frac{525}{512} = 1{,}025 \text{ или } 102{,}5\%$$

Средний арифметический индекс (индекс Струмилина)

$$\Delta_T = \sum T_1 - \sum i_w \cdot T_1 = 512 - 525 = -13 \ (\text{чел.})$$

В целом по предприятию производительность труда выросла на 2,5 %, в результате этого затраты труда сократились на 13 человек.

10.5. Базисные и цепные индексы

При изучении динамики явления за ряд последовательных периодов (лет, месяцев т.д.) рассчитывают ряд индексов. Эти индексы показывают изменение явления либо по отношению *к постоянной базе* (базисные индексы), либо по отношению *к переменной базе* (цепные индексы). Цепные и базисные индексы могут быть индивидуальными и общими. Расчет индивидуальных индексов при этом прост. (Для удобства записи отсчет времени начнем с первого периода). Тогда качественные базисные индивидуальные индексы в общем виде:

$$i_1 = \frac{x_2}{x_1};\ i_2 = \frac{x_3}{x_1};\ i_3 = \frac{x_4}{x_1};\ \text{и т.д.}$$

Цепные: $i_1 = \frac{x_2}{x_1}$; $i_2 = \frac{x_3}{x_2}$; $i_3 = \frac{x_4}{x_3}$; и т.д.

Аналогично рассчитываются и количественные базисные и цепные индивидуальные индексы.

Взаимосвязь между ними: произведение цепных индексов равно базисному:

$$\frac{x_2}{x_1} \cdot \frac{x_3}{x_2} \cdot \frac{x_4}{x_3} = \frac{x_4}{x_1}$$

При построении базисных и цепных общих индексов возникает проблема весов. Веса при этом могут быть *постоянными (*т.е. одинаковыми во всех индексах) и могут быть *переменными* (т.е. изменяющимися от индекса к индексу).

В большинстве случаев принято все индексы (базисные и цепные) *количественных* показателей записывать с *постоянными* весами. В общем виде это выглядит так:

а) базисные индексы

$$Id_{2/1} = \frac{\Sigma d_2 x_1}{\Sigma d_1 x_1};\ Id_{3/1} = \frac{\Sigma d_3 x_1}{\Sigma d_1 x_1};\ Id_{4/1} = \frac{\Sigma d_4 x_1}{\Sigma d_1 x_1};\ \text{и т.д.}$$

б) цепные индексы

$$Id_{2/1} = \frac{\Sigma d_2 x_1}{\Sigma d_1 x_1};\ Id_{3/2} = \frac{\Sigma d_3 x_1}{\Sigma d_2 x_1};\ Id_{4/3} = \frac{\Sigma d_4 x_1}{\Sigma d_3 x_1};\ \text{и т.д.}$$

Взаимосвязь между ними *в этом случае* сохраняется: *произведение цепных индексов равно базисному индексу*:

$$\frac{\Sigma d_2 x_1}{\Sigma d_1 x_1} \cdot \frac{\Sigma d_3 x_1}{\Sigma d_2 x_1} \cdot \frac{\Sigma d_4 x_1}{\Sigma d_3 x_1} = \frac{\Sigma d_4 x_1}{\Sigma d_1 x_1}.$$

Базисные и цепные индексы *качественных* показателей в большинстве случаев записываются *с переменными* весами. В общем виде:

а) базисные индексы

$$Id_{2/1}=\frac{\Sigma x_2 d_2}{\Sigma x_1 d_2};\ Id_{3/1}=\frac{\Sigma x_3 d_3}{\Sigma x_1 d_3};\ Id_{4/1}=\frac{\Sigma x_4 d_4}{\Sigma x_1 d_4}\ \text{и т.д.}$$

б) цепные индексы:

$$Id_{2/1}=\frac{\Sigma x_2 d_2}{\Sigma x_1 d_2};\ Id_{3/2}=\frac{\Sigma x_3 d_3}{\Sigma x_2 d_3};\ Id_{4/3}=\frac{\Sigma x_4 d_4}{\Sigma x_4 d_4};\ \text{и т.д.}$$

Между базисными и цепными индексами *с переменными* весами вышеуказанная взаимосвязь *отсутствует.*

Формулы базисных и цепных индексов конкретных показателей приведены в Приложении 2.

Пример 10.5.

Имеются данные о себестоимости изделий на одном из предприятий за I квартал.

Таблица 10.5

Изделия	Себестоимость 1 изделия, тыс. руб.			Выпуск продукции в натуральном выражении, шт.		
	Январь Z_1	Февраль Z_2	Март Z_3	Январь q_1	Февраль q_2	Март q_3
А	4,3	4,5	4,2	250	240	260
Б	3,0	2,9	3,2	900	940	710

Необходимо проанализировать динамику себестоимости продукции и ее физического объема.

Построим расчетную таблицу 10.5.1.

Таблица 10.5.1

Изделия	Затраты на выпуск (издержки)					
	z_1q_1	z_1q_2	z_1q_3	z_2q_2	z_2q_3	z_3q_3
А	1075	1032	1118	1080	1170	1092
Б	2700	2820	2130	2726	2059	2272
Итого	3775	3852	3248	3806	3229	3364

Сопоставим себестоимость каждого месяца с предыдущим (цепные индексы себестоимости):

февраль с январем

$$I_{2/1} \frac{\sum z_2 q_2}{\sum z_1 q_2} = \frac{3806}{3852} = 0{,}988 \text{ или } 98{,}8\%$$

Март с февралем:

$$I_{3/2} = \frac{\sum z_3 q_3}{\sum z_2 q_3} = \frac{3364}{3229} = 1{,}042 \text{ или } 104{,}2\%$$

Сопоставим себестоимость марта с январем (базисный индекс себестоимости):

$$I_{3/1} = \frac{\sum z_3 q_3}{\sum z_1 q_3} = \frac{3364}{3248} = 1{,}036 \text{ или } 103{,}6\%$$

Следовательно, себестоимость продукции по сравнению с январем выросла на 3,6 %, хотя в феврале по сравнению с январем себестоимость снизилась на 1,2 % , но марте по сравнению с февралем наблюдался рост на 4,2.

Физический объем продукции в феврале по сравнению с январем увеличился на 2 %:

$$I_{2/1} = \frac{\sum q_2 z_1}{\sum q_1 z_1} = \frac{3852}{3775} = 1{,}020 \text{ или } 102{,}0\%$$

В марте по сравнению с февралем снизился на17,7 %:

$$I_{3/2} = \frac{\sum q_3 z_1}{\sum q_2 z_1} = \frac{3248}{3852} = 0{,}843 \text{ или } 84{,}3\%$$

В марте по сравнению с январем физический объем продукции снизился на 14 %.

$$I_{3/1} = \frac{\sum q_3 z_1}{\sum q_1 z_1} = \frac{3248}{3775} = 0{,}860 \text{ или } 86\%$$

Взаимосвязь:

$$0{,}86 = 1{,}02 \cdot 0{,}843$$

10.6. Индексы средних показателей

С помощью данных индексов изучается динамика среднего уровня качественного показателя. Качественный показатель при этом характеризует одно и тоже явление (цену, себестоимость продукции, заработную плату, производительность труда и т.п.), которое наблюдается на разных участках. Средний уровень качественного признака зависит не только от самих осредняемых величин, но и от состава (структуры) совокупности, которая определяется по объемному признаку.

Поэтому изменение средней во времени зависит от изменения собственно значений признака и от изменения структуры совокупности.

Методику расчета индекса среднего уровня покажем на примере индексов себестоимости переменного, постоянного составов и структурных сдвигов.

Пример 10.6.

Предположим, что одна и та же продукция " A " производится на двух предприятиях с различной себестоимостью. В этом случае для характеристики динамики себестоимости индекс может быть рассчитан как индекс переменного состава и индекс постоянного (фиксированного) состава.

Таблица 10.6

Себестоимость и количество продукции «А»,

производимой на двух предприятиях

Предприятия	Базисный период		Отчетный период		Индивидуальные индексы себестоимости (по каждому предприятию) $i_z = \frac{z_1}{z_0}$	Затраты на выпуск продукции "А", руб.		
	Себестоимость 1 шт., руб. (z_0)	Количество прод. шт. (q_0)	Себестоимость 1 шт., руб. (z_1)	Количество прод. шт. (q_1)		Базисные z_0q_0	Отчетные z_1q_1	Базисные в пересчете на фактический объем z_0q_1
№1	15	5000	11	20000	0,733	75000	220000	300000
№2	18	10000	13	15000	0,722	180000	195000	270000
Итого	x	15000	x	35000	x	255000	415000	570000

Тогда индекс себестоимости *переменного состава* будет равен:

$$I_{n/c} = \overline{Z}_1 : \overline{Z}_0 = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{415000}{35000} : \frac{255000}{15000} = 11{,}9\,руб. : 17{,}0\,руб. = 0{,}7$$

или 70 %.

Таким образом, по двум предприятиям себестоимость продукции "А" снизилась на 30 %, в то время как снижение себестоимости по первому предприятию 26,7 %, а по второму 27,8 %.

Причина такого расхождения кроется в сущности индекса. Индекс *переменного состава* характеризует изменение *средней* себестоимости ($\overline{Z}$). *На величине средней каждого периода отражается не только изменение себестоимости, но и изменение удельного веса каждого предприятия в об-*

щем объеме производства $\left(\frac{q}{\sum q}\right)$. Следовательно, на индексе переменного состава сказывается влияние сразу двух факторов.

Для того чтобы выявить влияние каждого фактора в отдельности на величину индекса переменного состава, следует рассчитать еще 2 индекса: индекс постоянного (фиксированного) состава и индекс структурных сдвигов.

Индекс *постоянного (фиксированного) состава* - это тоже отношение двух средних уровней себестоимости, но *при условии неизменной структуры* (удельного веса предприятий в общем объеме производства продукции "A").

$$I_{ф/с} = \overline{Z_1} : \overline{Z_0}(усл) = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{415000}{35000} : \frac{570000}{35000} = 11{,}9 руб. : 16{,}3 руб. = = 0{,}73$$

или 73 %.

Этот индекс учитывает изменение *только* самой *себестоимости:* она снизилась на 27 %.

Для выявления влияния структурных сдвигов рассчитываем *индекс структурных сдвигов.* Это тоже отношение двух средних уровней себестоимости, но в них *исключено влияние себестоимости*:

$$I_{с/с} = \overline{Z_0}(усл) : \overline{Z_0} = \frac{\sum z_0 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \frac{415000}{35000} : \frac{255000}{15000} = 16{,}3 руб. : 17{,}0 руб. = = 0{,}96$$

или 96 %.

Следовательно, *в результате изменений в структуре* выпуска, а именно увеличения доли первого предприятия, где себестоимость ниже, произошло дополнительное снижение средней себестоимости на 4%.

Взаимосвязь этих индексов: $I_{п/с} = I_{ф/с} \cdot I_{с/с}$. В приведенном примере $0{,}7 = 0{,}73 \cdot 0{,}96$.

Аналогично рассчитываются все подобные индексы. Следует помнить, что эти индексы могут быть рассчитаны *только для качественных показателей* (цены, себестоимости, урожайности, заработной платы, производительности труда). В общем виде:

1. Индекс *переменного* состава: $I_{п/с} = \dfrac{\bar{x}_1}{\bar{x}_0} = \dfrac{\sum x_1 d_1}{\sum d_1} : \dfrac{\sum x_0 d_0}{\sum d_0}$;

2. Индекс *фиксированного* состава: $I_{ф/с} = \overline{x_1} : \overline{x_0}(усл) = \dfrac{\sum x_1 d_1}{\sum d_1} : \dfrac{\sum x_0 d_1}{\sum d_1}$;

3. Индекс *структурных сдвигов*: $I_{с/с} = \overline{x_0}(усл) : \overline{x_0} = \dfrac{\sum x_0 d_1}{\sum d_1} : \dfrac{\sum x_0 d_0}{\sum d_0}$

Исключением является индекс производительности труда по трудоемкости:

$$I_{п/с} = \overline{t_0} : \overline{t_1} = \frac{\sum t_0 q_0}{\sum q_0} : \frac{\sum t_1 q_1}{\sum q_1}$$

$$I_{ф/с} = \overline{t_0}(усл) : \overline{t_1} = \frac{\sum t_0 q_1}{\sum q_1} : \frac{\sum t_1 q_1}{\sum q_1}$$

$$I_{с/с} = I_{п/с} : I_{ф/с}.$$

Индексы конкретных средних показателей приведены в Приложении 2.

10.7. Территориальные индексы

Индексы, отражающие изменение явления *в пространстве*, т.е. городам, экономическим районам, республикам и т.д. называются *территориальными.*

Трудность построения территориальных индексов заключается в выборе веса, т.к. при сопоставлении по территориям понятия «базисный период» и «отчетный период» имеют условное значение.

Существует несколько различных методов сопоставления уровней признака по территориям, т.е. в пространстве. Рассмотрим один из них – это метод стандартных весов. Этот метод заключается в том, что значения индексируемой величины умножаются не на количество, например, продукции сравниваемого предприятия, а на количество продукции, произведенной на обоих сравниваемых предприятиях. Следовательно, если мы сопоставим себестоимость предприятия *« A »* с себестоимостью предприятия *« B »,* то формула территориального индекса себестоимости примет вид:

$$I_z = \frac{\sum z_A(q_A + q_B)}{\sum z_B(q_A + q_B)}$$

Если же сравнивать себестоимость продукции предприятия *« B »* с себестоимостью продукции предприятия *« A »,* то территориальный индекс себестоимости примет вид:

$$I_z = \frac{\sum z_B(q_A + q_B)}{\sum z_A(q_A + q_B)}$$

Однако это не изменит результатов выводов.

Пример 10.7.

Сравним себестоимость продукции и физический объем продукции двух предприятий. Причем предприятие « A » будем сравнивать с предприятием « B ».

Таблица 10.7

Изделия	Предприятие *« A »*		Предприятие *« B »*	
	Себестоимость 1 изделия, руб. (Z_A)	Количество изделий,шт. (q_A)	Себестоимость 1 изделия,руб. (Z_B)	Количество изделий,шт. (q_B)
Изделие 1	50	1200	60	300
Изделие 2	20	500	10	300
Изделие 3	70	300	60	1400

Территориальный индекс себестоимости при сравнении предприятия *« A »* с предприятием *« B »* будет равен:

$$I_z = \frac{\sum z_A(q_A + q_B)}{\sum z_B(q_A + q_B)} =$$

$$= \frac{50(1200+300)+20(500+300)+70(300+1400)}{60(1200+300)+10(500+300)+60(300+1400)} = \frac{210000}{200000} = 1{,}05 \textit{ или } 105\%$$

Следовательно, себестоимость продукции на предприятии *« A »* выше, чем на предприятии *« B »* на 5 %.

При сравнении себестоимости продукции предприятия *« B »* с себестоимостью продукции предприятия *« A »* территориальный индекс будет равен:

$$I_z = \frac{\sum z_B(q_A + q_B)}{\sum z_A(q_A + q_B)} = \frac{200000}{210000} = 0{,}95 \textit{ или } 95\%$$

т.е. себестоимость продукции на предприятии *« B »* ниже, чем на предприятии *« A »* на 5 %.

Так решается вопрос построения *территориальных качественных показателей,* которые в общем виде можно записать так:

$$I_X = \frac{\sum x_A(d_A + d_B)}{\sum x_Б(d_A + d_B)}$$

Сравнение в пространстве *количественных показателей* также представляет трудность. Например, при сравнении физического объема продукции из-за различной структуры продукции на предприятиях.

Расчет такого территориального индекса по одной из наиболее употребляемых формул, например, по данным вышеприведенного примера будет:

$$I_q = \frac{\sum q_A \cdot \overline{z_{AB}}}{\sum q_B \cdot \overline{z_{AB}}}$$

где:

$$\overline{z_{A,Б}} = \frac{z_A q_A + z_B q_B}{q_A + q_B}$$

Тогда средняя между предприятиями себестоимость

изделия 1:

$$\overline{z}_1 = \frac{5 \cdot 1200 + 6 \cdot 300}{1200 + 300} = 5{,}2 руб.$$

изделия 2:

$$\overline{z}_2 = \frac{20 \cdot 500 + 10 \cdot 300}{500 + 300} = 16{,}2 руб.$$

изделия 3:

$$\overline{z}_3 = \frac{70 \cdot 300 + 60 \cdot 1400}{300 + 1400} = 61{,}8 руб.$$

Индекс физического объема продукции предприятия *« A »* в сравнении с предприятием *« B »* буде равен:

$$I_q = \frac{\sum q_A \cdot \overline{z_{AB}}}{\sum q_B \cdot \overline{z_{AB}}} =$$

$$= \frac{1200 \cdot 52 + 500 \cdot 16{,}2 + 300 \cdot 61{,}8}{300 \cdot 52 + 300 \cdot 16{,}2 + 1400 \cdot 61{,}8} = \frac{89040}{106980} = 0{,}832 \quad \textit{или}\ 83{,}2\%$$

Таким образом, физический объем продукции предприятия $«A»$ меньше чем на предприятии $«B»$ на 16,8 %.

Все *количественные* индексы *территориальные* в общем виде можно представить формулой $I_d = \frac{\sum d_A \cdot \overline{x_{AB}}}{\sum d_B \cdot \overline{x_{AB}}}$

Территориальные индексы конкретных показателей приведены в Приложении 2.

Приложение 1

Комбинации статистических графиков

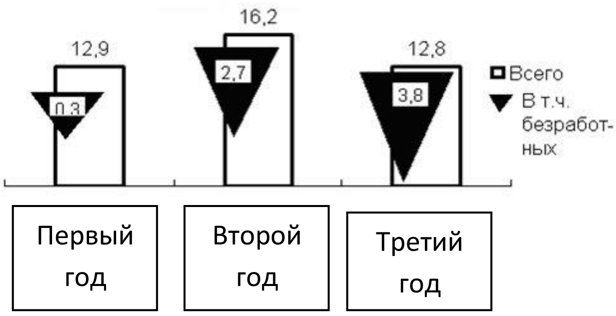


Рис. 1. Численность незанятого трудоспособного населения, ищущего работу в РТ на начало года, тыс. чел.

(Сочетание столбиковой диаграммы с плоскостной треугольной)

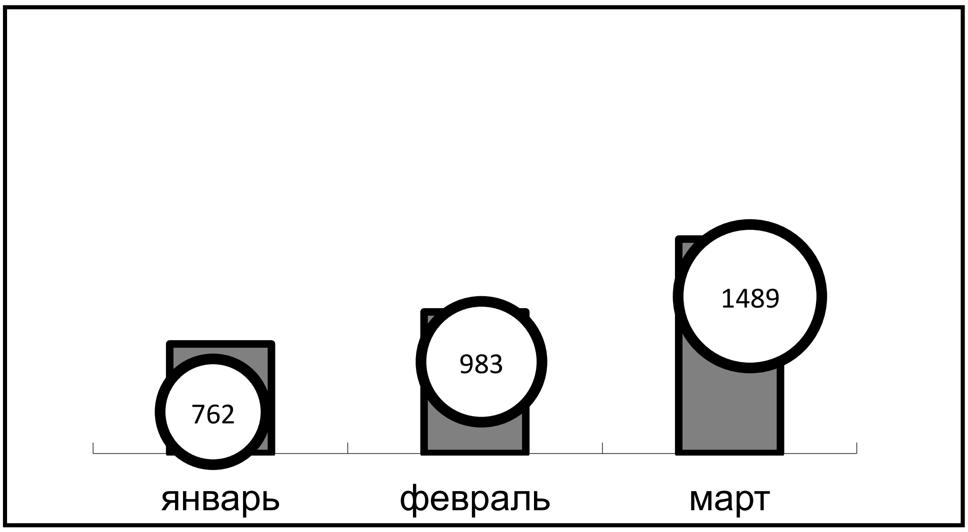


Рис. 2. Розничный товарооборот РФ в I квартале

в текущих ценах, млрд. руб.

(Комбинация столбиковой диаграммы с плоскостной круговой)

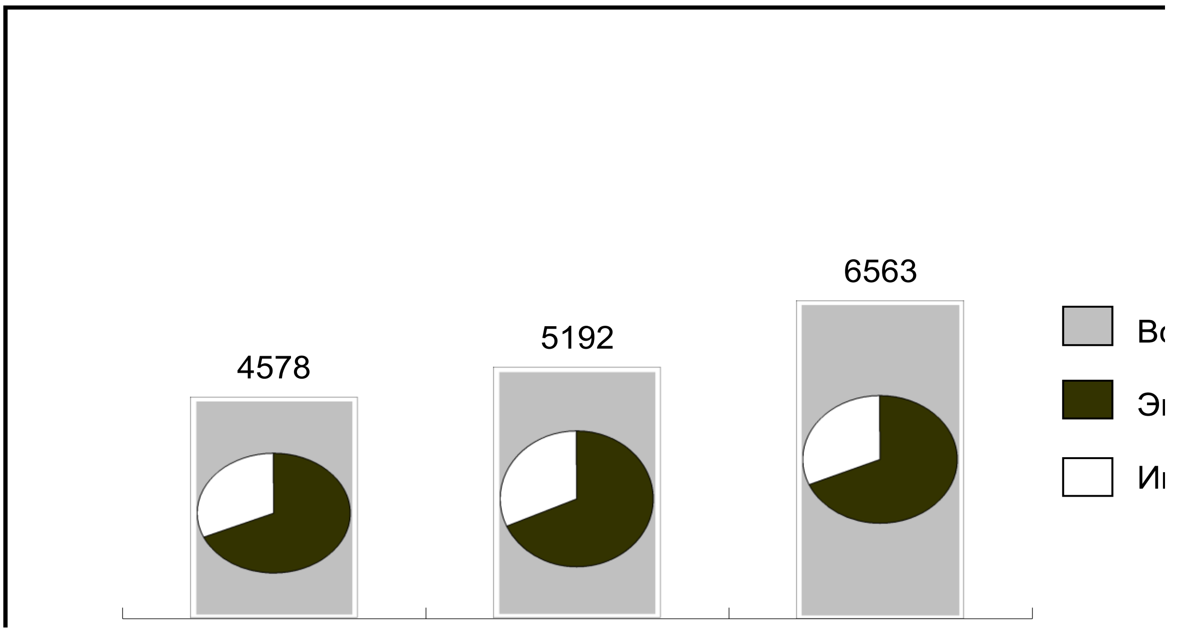


Рис. 3. Внешнеторговый оборот Российской Федерации во втором квартале. (Комбинация столбиковой диаграммы с секторной)

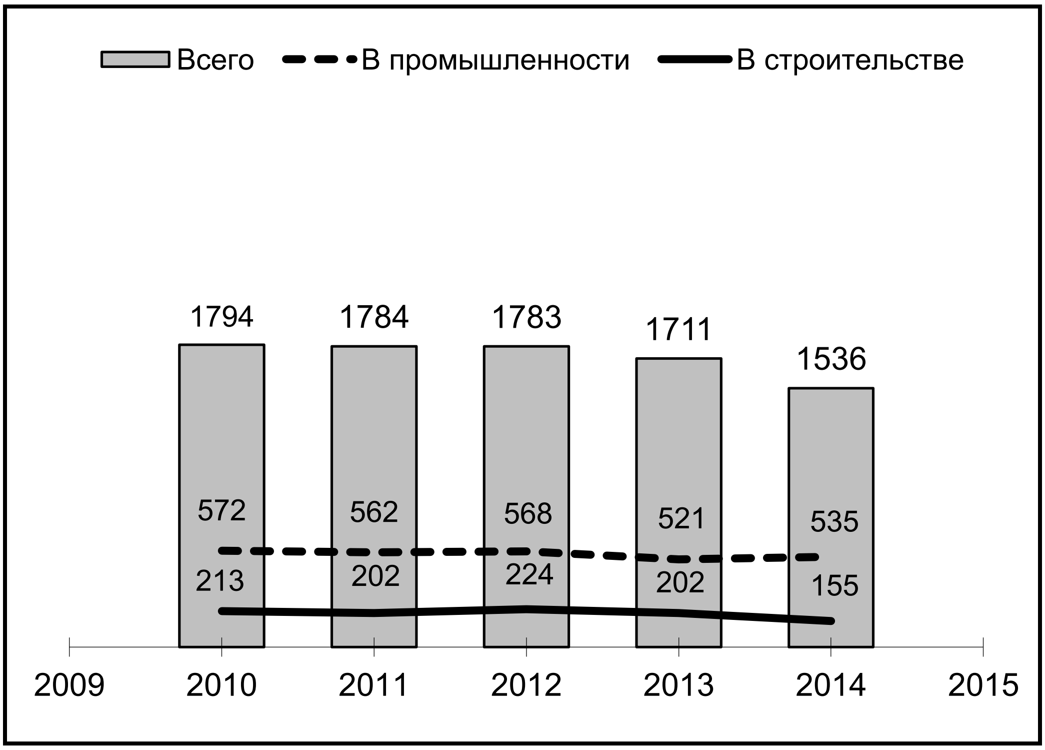


Рис. 4. Динамика средней списочной работающих в РТ, тыс. чел. (Комбинация столбиковой диаграммы и линейного графика)

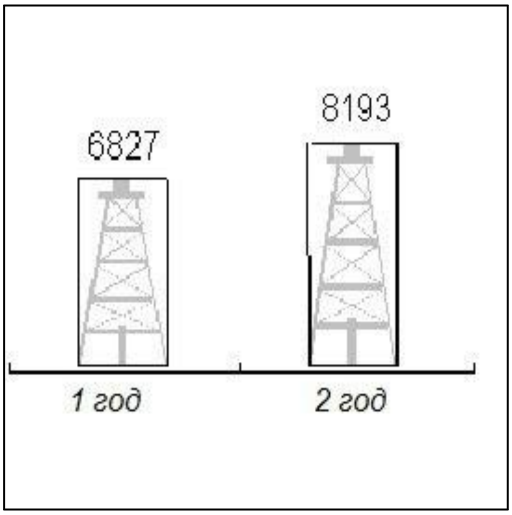


Рис. 5. Динамика экспорта нефти из России, млрд. долл.

(Комбинация столбиковой и фигурной диаграмм)

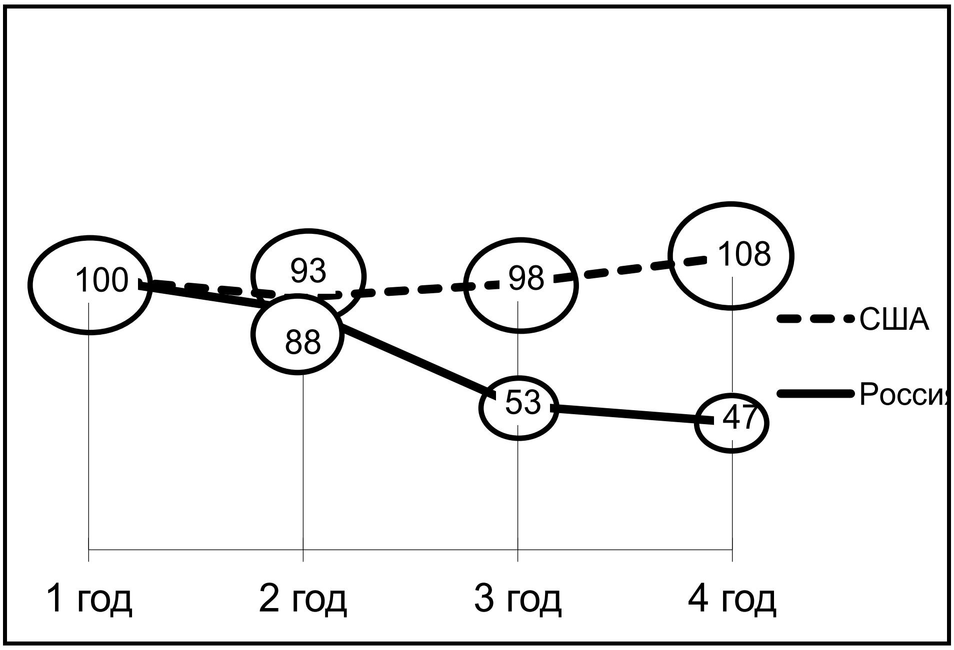


Рис. 6. Динамика капитальных вложений

в России и США в % к 1 году

(Сочетание круговой диаграммы и линейного графика)

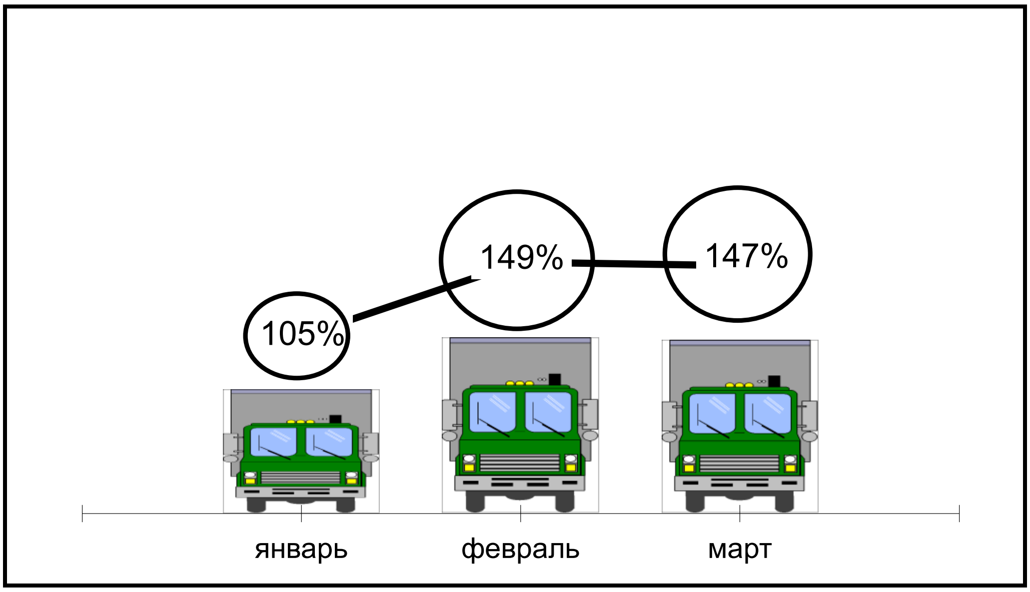


Рис. 7. Темпы роста тарифов автомобильного транспорта в РФ, в % к предыдущему месяцу

(Комбинация круговой и фигурной диаграммы)

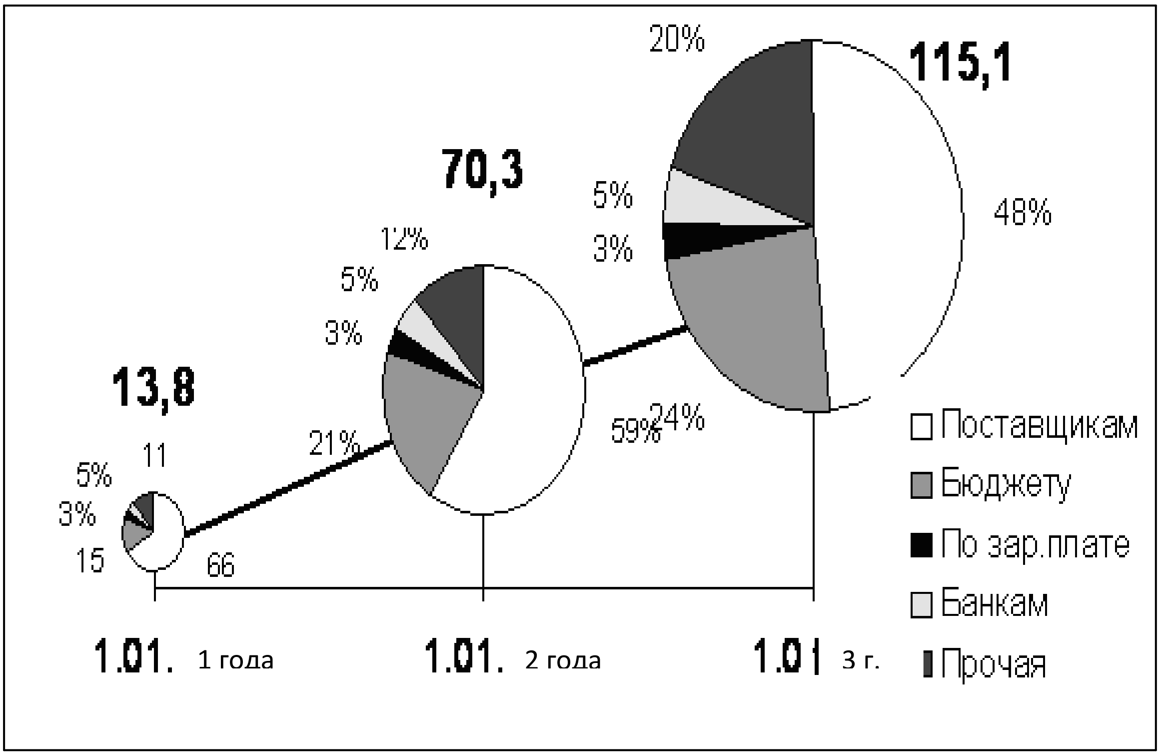


Рис. 8. Динамика объема и структуры просроченных обязательств предприятий РФ.

(Комбинация линейного графика, круговой и секторной диаграмм)

ПРИЛОЖЕНИЕ 2

Формулы индексов

в агрегатной форме и форме среднего индекса

Таблица 1

№ № п/п	Название индекса	Индивидуальный индекс	Агрегатная форма индекса	Форма среднего индекса	Изменение в абсолютном выражении
	А	1	2	3	4
1.	Индекс цен	$i_p = \dfrac{p_1}{p_0}$	$I_p = \dfrac{\sum p_1q_1}{\sum p_0q_1}$	$I_p = \dfrac{\sum p_1q_1}{\sum \dfrac{p_1q_1}{i_p}}$	$\sum p_1q_1 - \sum p_0q_1$ или $\sum p_1q_1 - \sum \dfrac{p_1q_1}{i_p}$
2.	Индекс себестоимости	$i_z = \dfrac{z_1}{z_0}$	$I_z = \dfrac{\sum z_1q_1}{\sum z_0q_1}$	$I_z = \dfrac{\sum z_1q_1}{\sum \dfrac{z_1q_1}{i_z}}$	$\sum Z_1q_1 - \sum Z_0q_1$ или $\sum Z_1q_1 - \sum \dfrac{Z_1q_1}{i_z}$
3.	Индекс урожайности	$i_y = \dfrac{y_1}{y_0}$	$I_y = \dfrac{\sum y_1П_1}{\sum y_0П_1}$	$I_y = \dfrac{\sum y_1П_1}{\sum \dfrac{y_1П_1}{i_y}}$	$\sum У_1П_1 - \sum У_0П_1$ или $\sum У_1П_1 - \sum \dfrac{У_1П_1}{i_y}$
4.	Индекс заработной платы	$i_f = \dfrac{f_1}{f_0}$	$I_f = \dfrac{\sum f_1T_1}{\sum f_0T_1}$	$I_f = \dfrac{\sum f_1T_1}{\sum \dfrac{f_1T_1}{i_f}}$	$\sum f_1T_1 - \sum f_0T_1$ или $\sum f_1T_1 - \sum \dfrac{f_1T_1}{i_f}$

	А	1	2	3	4
5.	Индекс производительности труда по выработке	$i_w = \frac{w_1}{w_0}$	$I_w = \frac{\sum w_1 T_1}{\sum w_0 T_1}$	$I_w = \frac{\sum w_1 T_1}{\sum \frac{w_1 T_1}{i_w}}$	$\sum W_1 T_1 - \sum W_0 T_1$ или $\sum W_1 T_1 - \sum \frac{W_1 T_1}{i_w}$
6.	Индекс производительности труда по трудоемкости	$i_w = \frac{t_0}{t_1}$	$I_w = \frac{\sum t_0 q_1}{\sum t_1 q_1}$	$I_w = \frac{\sum i_w t_1 q_1}{\sum t_1 q_1} = \frac{\sum i_w T_1}{\sum T_1}$	$\sum t_1 q_1 - \sum t_0 q_1$ или $\sum t_1 q_1 - \sum i_w t_1 q_1 = \sum T_1 - \sum i_w T_1$
7.	Индекс трудоемкости	$i_t = \frac{t_1}{t_0}$	$I_t = \frac{\sum t_1 q_1}{\sum t_0 q_1}$	$I_t = \frac{\sum t_1 q_1}{\sum \frac{t_1 q_1}{i_t}}$	$\sum t_1 q_1 - \sum t_0 q_1$ или $\sum t_1 q_1 - \sum \frac{t_1 q_1}{i_t}$
8.	Индекс физического объема продукции	$i_q = \frac{q_1}{q_0}$	$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$ $I_q = \frac{\sum q_1 z_0}{\sum q_0 z_0}$ $I_q = \frac{\sum q_1 t_0}{\sum q_0 t_0}$	$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0}$ $I_q = \frac{\sum i_q q_0 z_0}{\sum q_0 z_0}$ $I_q = \frac{\sum i_q q_0 t_0}{\sum q_0 t_0}$	$\sum q_1 p_0 - \sum q_0 p_0$ или $\sum i_q q_0 p_0 - \sum q_0 p_0$ $\sum q_1 Z_0 - \sum q_0 Z_0$ или $\sum i_q q_0 Z_0 - \sum q_0 Z_0$ $\sum q_1 t_0 - \sum q_0 t_0$ или $\sum i_q q_0 t_0 - \sum q_0 t_0$
9.	Индекс посевной площади	$i_П = \frac{П_1}{П_0}$	$I_П = \frac{\sum П_1 У_0}{\sum П_0 У_0}$	$I_П = \frac{\sum i_П П_0 У_0}{\sum П_0 У_0}$	$\sum П_1 У_0 - \sum П_0 У_0$ или $\sum i_П П_0 У_0 - \sum П_0 У_0$

	A	1	2	3	4
10.	Индекс численности рабочих	$i_T = \frac{T_1}{T_0}$	$I_T \frac{\sum T_1 W_0}{\sum T_0 W_0}$ $I_T \frac{\sum T_1 f_0}{\sum T_0 f_0}$	$I_T = \frac{\sum t_T T_0 W_0}{\sum T_0 W_0}$ $I_T = \frac{\sum i_T T_0 f_0}{\sum T_0 f_0}$	$\sum T_1 W_0 - \sum T_0 W_0$ или $\sum T_1 f_0 - \sum T_0 f_0$ или $\sum i_T T_0 f_0 - \sum T_0 f_0$
11.	Индекс товарооборота	$i_{pq} = \frac{p_1 q_1}{p_0 q_0}$	$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}$	-	$\sum p_1 q_1 - \sum p_0 q_1$
12.	Индекс издержек (затрат) на производство	$i_{zq} = \frac{z_1 q_1}{z_0 q_0}$	$I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0}$	-	$\sum Z_1 q_1 - \sum Z_0 q_0$
13.	Индекс валового сбора	$i_{УП} = \frac{У_1 П_1}{У_0 П_0}$	$I_{УП} = \frac{\sum У_1 П_1}{\sum У_0 П_0}$	-	$\sum У_1 П_1 - \sum У_0 П_0$
14.	Индекс затрат труда	$i_{tq} = \frac{t_1 q_1}{t_0 q_0} = \frac{T_1}{T_0}$	$I_{tq} = \frac{\sum t_1 q_1}{\sum t_0 q_0} =$ $= \frac{\sum T_1}{\sum T_0}$	-	$\sum t_1 q_1 - \sum t_0 q_0 =$ $= \sum T_1 - \sum T_0$
15.	Индекс фонда оплаты труда	$i_{fT} = \frac{f_1 T_1}{f_0 T_0}$	$I_{fT} = \frac{\sum f_1 T_1}{\sum f_0 T_0}$	-	$\sum f_1 T_1 - \sum f_0 T_0$

Формулы базисных и цепных индексов

Таблица 2

№№ п/п	Название индекса	Цепные индексы		Базисные индексы	
		Индивиду-альные	Общие	Индивиду-альные	Общие
	А	1	2	3	4
1.	Индекс цен	$i_{2/1} = \frac{p_2}{p_1}$ $i_{3/2} = \frac{p_3}{p_2}$ и т. д.	$I_{2/1} = \frac{\sum p_2 q_2}{\sum p_1 q_2}$ $I_{3/2} = \frac{\sum p_3 q_3}{\sum p_2 q_3}$ и т. д.	$i_{2/1} = \frac{p_2}{p_1}$ $i_{3/1} = \frac{p_3}{p_1}$ и т. д.	$I_{2/1} = \frac{\sum p_2 q_2}{\sum p_1 q_2}$ $I_{3/1} = \frac{\sum p_3 q_3}{\sum p_1 q_3}$ и т. д.
2.	Индекс себесто-имости	$i_{2/1} = \frac{z_2}{z_1}$ $i_{3/2} = \frac{z_3}{z_2}$ и т.д.	$I_{2/1} \frac{\sum z_2 q_2}{\sum z_1 q_2}$ $I_{3/2} = \frac{\sum z_3 q_3}{\sum z_2 q_3}$ и т.д.	$i_{2/1} = \frac{z_2}{z_1}$ $i_{3/1} = \frac{z_3}{z_1}$ и т.д.	$I_{2/1} \frac{\sum z_2 q_2}{\sum z_1 q_2}$ $I_{3/1} = \frac{\sum z_3 q_3}{\sum z_1 q_3}$ и т.д.

	А	1	2	3	4
3.	Индекс урожайности	$i_{2/1}=\frac{У_2}{У_1}$ $i_{3/2}=\frac{У_3}{У_2}$ и т.д.	$I_{2/1}=\frac{\sum У_2 П_2}{\sum У_1 П_2}$ $I_{3/2}=\frac{\sum У_3 П_3}{\sum У \; П_3}$ и т.д.	$i_{2/1}=\frac{У_2}{У_1}$ $i_{3/1}=\frac{У_3}{У_1}$ и т.д.	$I_{2/1}=\frac{\sum У_2 П_2}{\sum У_1 П_2}$ $I_{3/1}=\frac{\sum У_3 П_3}{\sum У_1 П_3}$ и т.д.
4.	Индекс заработной платы	$i_{2/1}=\frac{f_2}{f_1}$ $i_{3/2}=\frac{f_3}{f_2}$ и т.д.	$I_{2/1}=\frac{\sum f_2 T_2}{\sum f_1 T_2}$ $I_{3/2}=\frac{\sum f_3 T_3}{\sum f_2 T_3}$ и т.д.	$i_{2/1}=\frac{f_2}{f_1}$ $i_{3/1}=\frac{f_3}{f_1}$ и т.д.	$I_{2/1}=\frac{\sum f_2 T_2}{\sum f_1 T_2}$ $I_{3/1}=\frac{\sum f_3 T_3}{\sum f_1 T_3}$ и т.д.
5.	Индекс производительности труда по выработке	$i_{2/1}=\frac{w_2}{w_1}$ $i_{3/2}=\frac{w_3}{w_2}$ и т.д.	$I_{2/1}=\frac{\sum w_2 T_2}{\sum w_1 T_2}$ $I_{3/2}=\frac{\sum w_3 T_3}{\sum w_2 T_3}$ и т.д.	$i_{2/1}=\frac{w_2}{w_1}$ $i_{3/1}=\frac{w_3}{w_1}$ и т.д.	$I_{2/1}=\frac{\sum w_2 T_2}{\sum w_1 T_2}$ $I_{3/1}=\frac{\sum w_3 T_3}{\sum w_1 T_3}$ и т.д.

	А	1	2	3	4
6.	Индекс производительности труда по трудоемкости	$i_{2/1}=\frac{t_1}{t_2}$ $i_{3/2}=\frac{t_2}{t_3}$ и т.д.	$I_{2/1}=\frac{\sum t_1q_2}{\sum t_2q_2}$ $I_{3/2}=\frac{\sum t_2q_3}{\sum t_3q_3}$ и т.д.	$i_{2/1}=\frac{t_1}{t_2}$ $i_{3/1}=\frac{t_1}{t_3}$ и т.д.	$I_{2/1}=\frac{\sum t_1q_2}{\sum t_2q_2}$ $I_{3/1}=\frac{\sum t_1q_3}{\sum t_3q_3}$ и т.д.
7.	Индекс трудоемкости	$i_{2/1}=\frac{t_2}{t_1}$ $i_{3/2}=\frac{t_3}{t_2}$ и т.д.	$I_{2/1}=\frac{\sum t_2q_2}{\sum t_1q_2}$ $I_{3/2}=\frac{\sum t_3q_3}{\sum t_2q_3}$ и т. д.	$i_{2/1}=\frac{t_2}{t_1}$ $i_{3/1}=\frac{t_3}{t_1}$ и т.д.	$I_{2/1}=\frac{\sum t_2q_2}{\sum t_1q_2}$ $I_{3/1}=\frac{\sum t_3q_3}{\sum t_1q_3}$ и т. д.

	A	1	2	3	4
8.	Индекс физи-ческого объема продук-ции	$i_{2/1} = \frac{q_2}{q_1}$ $i_{3/2} = \frac{q_3}{q_2}$ и т.д.	$I_{2/1} = \frac{\sum q_2 p_1}{\sum q_1 p_1}$ $I_{3/2} = \frac{\sum q_3 p_1}{\sum q_2 p_1}$ *и т.д.* $I_{2/1} = \frac{\sum q_2 z_1}{\sum q_1 z_1}$ $i_{3/2} = \frac{\sum q_3 z_1}{\sum q_2 z_1}$ *и т.д.* $I_{2/1} = \frac{\sum q_2 t_1}{\sum q_1 t_1}$ $i_{3/2} = \frac{\sum q_3 t_1}{\sum q_2 t_1}$ *и т.д.*	$i_{2/1} = \frac{q_2}{q_1}$ $i_{3/1} = \frac{q_3}{q_1}$ и т.д.	$I_{2/1} = \frac{\sum q_2 t_1}{\sum q_1 t_1}$ $I_{3/1} = \frac{\sum q_3 t_1}{\sum q_1 t_1}$ *и т.д.* $I_{2/1} = \frac{\sum q_2 p_1}{\sum q_1 p_1}$ $I_{3/1} = \frac{\sum q_3 p_1}{\sum q_1 p_1}$ *и т.д.* $I_{2/1} = \frac{\sum q_2 z_1}{\sum q_1 z_1}$ $I_{3/1} = \frac{\sum q_3 z_1}{\sum q_1 z_1}$ *и т.д.*

	А	1	2	3	4
9.	Индекс посевной площади	$i_{2/1} = \frac{П_2}{П_1}$ $i_{3/2} = \frac{П_3}{П_2}$ и т.д.	$I_{2/1} = \frac{\sum П_2 У_1}{\sum П_1 У_1}$ $I_{3/2} = \frac{\sum П_3 У_1}{\sum П_2 У_1}$ и т.д.	$i_{2/1} = \frac{П_2}{П_1}$ $i_{3/1} = \frac{П_3}{П_1}$ и т.д.	$I_{2/1} = \frac{\sum П_2 У_1}{\sum П_1 У_1}$ $I_{3/1} = \frac{\sum П_3 У_1}{\sum П_1 У_1}$ и т.д.
10.	Индекс численности рабочих	$i_1 = \frac{T_2}{T_1}$ $i_2 = \frac{T_3}{T_2}$ и т.д.	$I_{2/1} = \frac{\sum T_2 w_1}{\sum T_1 w_1}$ $I_{3/2} = \frac{\sum T_3 w_1}{\sum T_2 w_1}$ и т.д. $I_{2/1} = \frac{\sum T_2 f_1}{\sum T_1 f_1}$ $I_{3/2} = \frac{\sum T_3 f_1}{\sum T_2 f_1}$ и т.д.	$i_{2/1} = \frac{T_2}{T_1}$ $I_{3/1} = \frac{T_3}{T_1}$ и т.д.	$I_{2/1} = \frac{\sum T_2 f_1}{\sum T f_{11}}$ $I_{3/1} = \frac{\sum T_3 f_1}{\sum T_1 f_1}$ и т.д.

Формулы индексов средних показателей

Таблица 3

№ п/п	Название индекса	Переменного состава	Постоянного (фиксированного) состава	Структурных сдвигов
	А	1	2	3
1.	Индекс себестоимости	$I_{п/с} = \dfrac{\overline{z_1}}{\overline{z_0}} = \dfrac{\sum z_1 q_1}{\sum q_1} : \dfrac{\sum z_0 q_0}{\sum q_0}$	$I_{ф/с} = \dfrac{\overline{z_1}}{\overline{z_{усл}}} = \dfrac{\sum z_1 q_1}{\sum q_1} : \dfrac{\sum z_0 q_1}{\sum q_1}$	$I_{с/с} = \dfrac{\overline{z_{усл}}}{\overline{z_0}} = \dfrac{\sum z_0 q_1}{\sum q_1} : \dfrac{\sum z_0 q_0}{\sum q_0}$
2.	Индекс урожайности	$I_{п/с} = \dfrac{\overline{y_1}}{\overline{y_0}} = \dfrac{\sum y_1 П_1}{\sum П_1} : \dfrac{\sum y_0 П_0}{\sum П_0}$	$I_{ф/с} = \dfrac{\overline{y_1}}{\overline{y_{усл}}} = \dfrac{\sum y_1 П_1}{\sum П_1} : \dfrac{\sum y_0 П_1}{\sum П_1}$	$I_{с/с} = \dfrac{\overline{y_{усл}}}{\overline{y_0}} = \dfrac{\sum y_0 П_1}{\sum П_1} : \dfrac{\sum y_0 П_0}{\sum П_0}$
3.	Индекс заработной платы	$I_{п/с} = \dfrac{\overline{f_1}}{\overline{f_0}} = \dfrac{\sum f_1 T_1}{\sum T_1} : \dfrac{\sum f_0 T_0}{\sum T_0}$	$I_{ф/с} = \dfrac{\overline{f_1}}{\overline{f_{усл}}} = \dfrac{\sum f_1 T_{11}}{\sum T_1} : \dfrac{\sum f_0 T_1}{\sum T_1}$	$I_{с/с} = \dfrac{\overline{f_{усл}}}{\overline{f_0}} = \dfrac{\sum f_0 T_1}{\sum T_1} : \dfrac{\sum f_0 T_0}{\sum T_0}$

1	2	3	4	5
4.	Индекс производительности труда по выработке	$I_{п/с} \frac{\overline{W_1}}{\overline{W_0}} =$ $= \frac{\sum W_1 T_1}{\sum T_1} : \frac{\sum W_0 T_0}{\sum T_0}$	$I_{ф/с} = \frac{\overline{W_1}}{\overline{W_{усл}}} =$ $= \frac{\sum W_1 T_1}{\sum T_1} : \frac{\sum W_0 T_1}{\sum T_1}$	$I_{с/с} = \frac{\overline{W_{усл}}}{\overline{W_0}} =$ $= \frac{\sum W_0 T_1}{\sum T_1} : \frac{\sum W_0 T_0}{\sum T_0}$
5.	Индекс производительности труда по трудоемкости	$I_{п/с} = \frac{\overline{t_0}}{\overline{t_1}} =$ $\frac{\sum t_0 q_0}{\sum q_0} : \frac{\sum t_1 q_1}{\sum q_1}$	$I_{ф/с} = \frac{\overline{t_{усл}}}{\overline{t_1}} =$ $= \frac{\sum t_0 q_1}{\sum q_1} : \frac{\sum t_1 q_1}{\sum q_1}$	$I_{с/с} = I_{п/с} : I_{ф/с}$

Формулы территориальных индексов

(при сравнении территории (объекта) А с территорией (объектом) В

Таблица 4

№ п/п	Название индекса	Формула
1.	Индекс цен	$I_p = \frac{\sum p_A \cdot (q_A + q_B)}{\sum p_B \cdot (q_A + q_B)}$
2.	Индекс себестоимости	$I_p = \frac{\sum z_A \cdot (q_A + q_B)}{\sum z_B \cdot (q_A + q_B)}$
3.	Индекс заработной платы	$I_f = \frac{\sum f_A \cdot (T_A + T_B)}{\sum f_B \cdot (T_A + T_B)}$
4.	Индекс производительности труда по выработке	$I_w = \frac{\sum w_A \cdot (T_A + T_B)}{\sum w_B \cdot (T_A + T_B)}$
5.	Индекс урожайности	$I_У = \frac{\sum У_A \cdot (П_A + П_B)}{\sum У_B \cdot (П_A + П_B)}$
6.	Индекс физического объема продукции	$I_q = \frac{\sum q_A \cdot \overline{p_{AB}}}{\sum q_B \cdot \overline{p_{AB}}}$ или $I_q = \frac{\sum q_A \cdot \overline{z_{AB}}}{\sum q_B \cdot \overline{z_{AB}}}$
7.	Индекс численности работников	$I_T = \frac{\sum T_A \cdot \overline{f_{AB}}}{\sum T_B \cdot \overline{f_{AB}}}$ или $I_T = \frac{\sum T_A \cdot \overline{w_{AB}}}{\sum T_B \cdot \overline{w_{AB}}}$
8.	Индекс посевной площади	$I_П = \frac{\sum П_A \cdot \overline{У_{AB}}}{\sum П_B \cdot \overline{У_{AB}}}$

Приложение 3

Статистико-математические таблицы

Таблица 1

Значения t-критерия Стьюдента

Число степеней свободы $n-1$	Уровень значимости			Число степеней свободы	Уровень значимости		
	0,10	0,05	0,01		0,10	0,05	0,01
1	2	3	4	1	2	3	4
1	6,314	12,706	63,657	18	1,734	2,101	2,878
2	2,920	4,303	9,925	19	1,729	2,093	2,861
3	2,353	3,182	5,841	20	1,725	2,086	2,845
4	2,132	2,776	4,604	21	1,721	2,080	2,831
5	2,015	2,571	4,032	22	1,717	2,074	2,819
6	1,943	2,447	3,707	23	1,714	2,069	2,807
7	1,895	2/365	3,499	24	1,711	2,064	2,797
8	1,859	2,306	3,355	25	1,708	2,059	2,787
9	1,833	2,262	3,249	26	1,706	2,055	2,779
10	1,812	2,228	3,169	27	1,703	2,052	2,771
11	1,796	2,201	3,106	28	1,701	2,048	2,763
12	1,782	2,179	3,054	29	1,699	2,045	2,756
13	1,771	2,160	3,012	30	1,697	2,042	2,750
14	1,761	2,145	2,977	40	1,684	2,02	2,704
15	1,753	2,131	2,947	60	1,671	2,000	2,660
16	1,746	2,120	2,921	120	1,658	1,980	2,617
17	1,740	2,110	2,898				

Таблица 2

Значения χ^2 - критерия Пирсона

Число степеней свободы $K=(m_1 1)\cdot(m_2-1)$	Уровень значимости			Число степеней свободы	Уровень значимости		
	0,10	0,05	0,01		0,10	0,05	0,01
1	2	3	4	1	2	3	4
1	2,71	3,84	6,63	21	29,62	32,67	38,93
2	4,61	5,99	9,21	22	30,81	33,92	40,29
3	6,25	7,81	11,34	23	32,01	35,17	41,64
1	2	3	4	1	2	3	4
4	7,78	9,49	13,28	24	33,20	36,42	42,98
5	9,24	11,07	15,09	25	34,38	37,65	44,31
6	10,64	12,59	16,81	26	35,56	38,89	45,64
7	12,02	14,07	18,48	27	36,74	40,11	46,96
8	13,36	15,51	20,09	28	37,92	41,34	48,28
9	14,68	16,92	21,67	29	39,09	42,56	49,59
10	15,99	18,31	23,21	30	40,26	43,77	50,89
11	17,28	19,68	24,72	40	51,80	55,76	63,69
12	18,55	21,03	26,22	50	63,17	67,50	76,15
13	19,81	22,36	27,69	60	74,40	79,08	88,38
14	21,06	23,68	29,14	70	85,53	90,53	100,42
15	22,31	25,00	30,58	80	96,58	101,88	112,33
16	23,54	26,30	32,00	90	107,56	113,14	124,12
17	24,77	27,59	33,41	100	118,50	124,34	135,81
18	25,99	28,87	34,81				
19	27,20	30,14	36,19				
20	28,41	31,41	37,57				

Таблица 3

Значения F -критерия Фишера при уровне значимости 0,05

v_2 \ v_1	1	2	3	v_2 \ v_1	1	2	3
1	161,00	200,00	216,00	18	4,41	3,55	3,16
2	18,51	19,00	19,16	19	4,38	3,52	3,13
3	10,13	9,55	9,28	20	4,35	3,49	3,10
4	7,71	6,94	6,59	21	4,32	3,47	3,07
5	6,61	5,79	5,41	22	4,30	3,44	3,05
6	5,99	5,14	4,76	23	4,28	3,42	3,03
7	5,59	4,74	4,35	24	4,26	3,40	3,01
8	5,32	4,46	4,07	25	4,24	3,88	2,99
9	5,12	4,26	3,86	26	4,22	3,37	2,98
10	4,96	4,10	3,71	27	4,21	3,35	2,96
11	4,84	3,98	3,59	28	4,20	3,34	2,95
12	4,75	3,88	3,49	29	4,18	3,33	2,93
13	4,67	3,80	3,41	30	4,17	3,32	2,92
14	4,60	3,74	3,34	40	4,08	3,23	2,84
15	4,54	3,68	3,29	50	4,03	3,18	2,79
16	4,49	3,63	3,24	60	4,00	3,15	2,76
17	4,45	3,59	3,20	100	3,94	3,09	2,70

$V_1 = m - 1$; $V_2 = n - m$; n -число наблюдений; m -число признаков.

Таблица 4

Распределение вероятности в малых выборках в зависимости от коэффициента доверия t и объема выборки n

n / t	4	5	6	7	8	9	10	15	20	∞
0,5	348	356	362	366	368	370	372	376	378	383
1,0	608	626	636	644	650	654	656	666	670	683
1,5	770	792	806	816	832	828	832	846	850	865
2,0	860	884	908	908	914	920	924	936	940	954
2,5	933	946	955	959	963	966	968	975	978	988
3,0	942	960	970	970	980	938	984	992	992	997

При $n = \infty$ в таблице даны вероятности нормального распределения. Для определения вероятности соответствующие табличные значения следует разделить на 1000.

ГЛОССАРИЙ

А

Абсолютное значение одного процента прироста – отношение абсолютного прироста к темпу прироста.

Абсолютные статистические величины – форма количественного выражения статистических показателей, непосредственно характеризующая размеры (абсолютные) социально-экономических явлений, их признаков в единицах меры протяженности, площади, массы (веса) и т.п., в единицах счета времени, в денежных единицах или в виде числа элементов (единиц), составляющих данное массовое явление, называемое статистической совокупностью.

Анализ регрессионный – оценка функциональной зависимости условного среднего значения результативного признака от факторных признаков, заключающаяся в определении аналитического выражения связи.

Анализ статистический – заключительная стадия статистического исследования, проводимый на уровне народного хозяйства, отраслей, объединений, предприятий с целью исследования характерных особенностей структуры, связи явлений, тенденции, закономерности развития социально-экономических явлений с использованием специфических экономико-статистических и математико-статистических методов.

Анкетный способ наблюдения – способ собирания статистических данных, состоящий в рассылке, раздаче определенному кругу лиц анкет (вопросников), возврат которых в заполненном виде является делом добровольным.

Б

Базисная величина – величина показателя, с которой сопоставляется какая либо другая сравниваемая (текущая, отчетная) величина.

Балансовый метод – метод статистического изучения процесса воспроизводства, заключающийся в сопоставлении систем показателей, отражающих состояние взаимозависимых элементов общественного воспроизводства.

В

Вариант – значение признака у единицы совокупности, отличное от значений его у других единиц.

Вариационный ряд – расположение случайной выборки с функцией распределения $F(x)$ в порядке их возрастания.

Вариация **–** колеблемость, изменение величины признака в статистической совокупности.

Веса (в статистике) – числа, в виде абсолютных величин или относительных величин, определяющие значимость (весомость, «вес») того или иного варианта признака в данной статистической совокупности, используемые для вычисления обобщающих показателей – средних величин, индексов, темпов роста.

Взаимосвязь индексов – связь между определенными индексами, обусловленная как реальными связями социально-экономических явлений, так и математическими свойствами индексов.

Взаимосвязь показателей – связь между статистическими показателями, отражающая объективно существующие взаимосвязи общественных явлений.

Взвешивание (в статистике) – способ вычисления статистических обобщающих показателей (средних величин, показателей вариации, индексов), заключающийся в том, что в расчет принимаются веса, значимость величины каждого варианта признака в совокупном итоге.

Время наблюдения – время, по состоянию на которое или за которое регистрируются сведения в процессе статистического наблюдения.

Выборка, выборочная совокупность – это совокупность, ограниченного числа наблюдений случайной величины.

Г

Геометрические знаки – это знаки, с помощью которых формируется понятие об отображаемых на графиках явлениях.

Гистограмма – способ графического изображения интервальных распределений, которая строится в прямоугольной системе координат, где на оси абсцис откладываются отрезки, изображающие интервалы значений варьирующего признака.

Графа – вертикальная полоса статистической таблицы.

График временного ряда – способ графического изображения изменения явлений во времени.

График контрольно-плановый показывает ход выполнения плана с помощью сетки со шкалой процентов, где каждому отрезку времени соответствует 5 вертикальных полос по 20 % прочеркивание всех пяти полос горизонтальной линией соответствует выполнению плана на 100 %.

Группировка – процесс образования групп единиц совокупности, однородных в каком-либо существенном отношении, а также имеющих одинаковые или близкие значения группировочного признака.

Группировка аналитическая – группировка, выявляющая взаимосвязи между изучаемыми признаками.

Группировка вторичная - прием, используемый в статистическом исследовании для образования новых групп на основе ранее произведенной группировки.

Группировка комбинационная – группировка, в которой расчленение статистической совокупности на группы производится по двум и более признакам, взятым в сочетании (комбинации).

Группировка простая – группировка, в которой объединение единиц совокупности в группы производится по одному признаку.

Группировка структурная – группировка, выявляющая состав (строение) однородной в количественном отношении статистической совокупности по определенным признакам.

Группировка типологическая – группировка, с помощью которой в изучаемой совокупности явлений выделяются однокачественные в существенном отношении группы, прежде всего классы и социально-экономические типы.

Д

Диаграмма – графическое изображение статистических данных, наглядно показывающих соотношение между сравниваемыми величинами.

Диаграмма балансовая – разновидность диаграммы, характеризующая балансовые соотношения в какой-либо области, наиболее распространенное изображение которой как ломаная линия, где вверх и вниз от горизонтальной оси размещаются графические образы противоположных явлений: поступления и использования, экспорта и импорта и т.п.

Диаграмма динамики структуры – графическое изображение изменения удельных весов и соотношений составных частей явлений.

Диаграмма изобразительная – графический способ, использующий наряду с геометрическими фигурами упрощенные изображения явлений, рисунки, художественные элементы с помощью знаков-символов или фигур -масштабных знаков.

Диаграмма ленточная *(полосовая)* – разновидность диаграммы, на которой величины изображаются в виде полос одинаковой ширины, располагаемых горизонтально.

Диаграмма плоскостная – разновидность диаграммы, изображающая размеры явлений площадями геометрических фигур.

Диаграмма секторная – разновидность диаграммы структурной, является распространенной формой сопоставления различных частей целого при помощи площадей, образуемых секторами круга.

Диаграмма спиральная *(график радиальный)* – вид графика, построенного в полярных координатах, использующийся для изображения явлений. Периодически изменяющихся во времени (преимущественно сезонных колебаний, где время отсчитывается по часовой стрелке по окружности, а значению показателя отвечает расстояние точки от центра.

Диаграмма сравнения – группа диаграмм, применяемых для сопоставления величин с использованием столбиковых, полосовых и плоскостных диаграмм.

Диаграмма столбиковая – разновидность диаграммы, которая изображает статистические величины в форме прямоугольников-столбиков, равных по величине основания и размещенных вертикально рядом или на одинаковом расстоянии друг от друга.

Диаграмма структурная – разновидность диаграммы, показывающей состав (структуру) целого, разделенного на части, в виде столбиковой, секторной и полосовой диаграмм.

Динамика *(в статистике)* – движение (изменение) явлений во времени.

Дискретные ряды распределения являются характеристикой дискретной случайной величины, варианты которой имеют значения целых чисел, т.е. между ними не может быть никаких промежуточных значений.

Дисперсия – средний квадрат отклонений значений признака от их средней арифметической величины.

Дисперсия межгрупповая – средний квадрат отклонений средних величин признака в каждой группе от средней общей для всей статистической совокупности в целом, измеряющая степень колеблемости (вариацию) признака в совокупности за счет фактора, положенного в основание группировки.

Дисперсия средняя из групповых дисперсий – дисперсия исчисляемая как средняя арифметическая величина из дисперсий, рассчитанных по каждой группе, на которые разбита статистическая совокупность и характеризующая степень колеблемости (вариации) признака во всей совокупности в целом за счет действия на него всех прочих факторов, кроме положенного в основание группировки.

Доля выборки – отношение численности единицы совокупности выборочной к численности их в генеральной совокупности, характеризующей степень охвата единиц генеральной совокупности статистическим наблюдением.

Доля выборочная – относительная численность выборочной совокупности, обладающих данным признаком или данным значением его.

Доля генеральная – относительная численность единиц, обладающих данным признаком или данным значением его во всей генеральной совокупности.

Достоверность информации – степень адекватности отображения информацией описываемых ею явлений, событий или процессов.

Е

Единица наблюдения – первичный элемент объекта статистического наблюдения, являющийся носителем регистрируемых при наблюдении признаков.

Единица совокупности – индивидуальный составной элемент статистической совокупности.

З

Закон больших чисел – общий принцип, в силу которого количественные закономерности, присущие массовым общественным явлениям, отчетливо проявляются лишь в достаточно большом числе наблюдений.

Знак Варзара – плоскостная диаграмма в виде прямоугольника, отображающая одновременно три величины: одна изображается основанием прямоугольника, другая – его высотой и третья, равная их произведению.

Значимость – статистическая существенность связи или ее характеристик (параметров, коэффициентов) утверждается при отклонении нулевой гипотезы об отсутствии связи (о равенстве нулю параметров, измеряющих связь) с вероятностью ошибки.

И

Индекс – относительный статистический показатель, характеризующий соотношение во времени или в пространстве сложных социально-экономических явлений, состоящих из элементов непосредственно не поддающихся суммированию.

Индекс агрегатный – основная форма индекса сводного, характеризующая относительные изменения индексируемой величины в текущем периоде по сравнению с базисным, числители и знаменатели которой представляют собой суммы произведений индексируемой величины и ее веса за два сравниваемых периода.

Индекс индивидуальный – относительная величина, характеризующая изменение во времени отдельных элементов сложного социально-экономического явления.

Индекс переменного состава – индекс, выражающий соотношение средних уровней изучаемого явления, относящихся к разным периодам времени или разным территориям, выражающий изменение не только индексируемой величины, но и весов.

Индекс постоянного (фиксированного) состава – индекс, вычисленный с весами фиксируемыми на уровне отчетного периода, и показывающий изменение только индексируемой величины.

Индекс средний – индекс, вычисленный как средняя величина из индивидуальных индексов.

Индекс структурных сдвигов – индекс, характеризующий влияние структурных сдвигов (изменение структуры изучаемого явления) на динамику среднего уровня этого явления.

Индекс территориальный характеризует соотношение социально-экономических явлений в пространстве (экономическим районам, областям, городам и т.п., для построения которого чаще всего используют метод стандартных весов.

Индексный метод *(в статистике*) – метод статистического исследования, основанный на построении в анализе индексов, позволяющих соизмерять сложные социально-экономические явления, который широко используется для изучения динамики явления, для сопоставления в пространстве, сопоставления фактических и плановых заданий, позволяет выявлять и измерять влияние факторов на изменение изучаемого явления.

Индексы базисные – система (ряд) последовательно вычисленных индексов одного и того же явления с постоянной базой сравнения.

Индексы сезонности – показатели интенсивности сезонных колебаний, определяемый в общем виде как отношение каждого уровня ряда динамики в виде помесячных данных к некоторому теоретическому или среднему уровню, принимаемому в качестве базы сравнения, и выраженные в процентах.

Индексы цепные – система (ряд) индексов одного и того же явления, вычисленных с меняющейся от индекса к индексу базой сравнения.

Интервалы группировок – обозначение групп «от» – «до», образованных группировкой по количественному признаку, величина которого определяется разностью верхних и нижних границ.

Интерполяция графическая – отыскание промежуточных значений величины по некоторым известным ее значениям с помощью графика.

К

Картограмма – контурная географическая карта, на которой штриховкой различной густоты, точками или окраской различной степени насыщенности показана сравнительная интенсивность какого-либо показателя в пределах каждой единицы нанесенного на карту территориального деления.

Картодиаграмма – вид картограммы, на которой с помощью диаграммных фигур изображены величины какого-либо статистического показателя в пределах каждой единицы нанесенного на карту территориального деления.

Колеблемость ряда динамики – отклонения уровня ряда динамики от их средней величины или от сглаженных их значений, вызванные (в частности) случайными причинами.

Контроль логический (*при статистическом наблюдении*) – сопоставление ответов на взаимосвязанные вопросы статистического формуляра с целью выявления логически несовместимых ответов.

Контроль счетный – проверка правильности арифметических итогов, расчета показателей, содержащихся в статистическом формуляре.

Корреляционное отношение эмпирическое – показатель тесноты связи между взаимосвязанными явлениями (их признаками), равный корню квадратному из отношения межгрупповой к общей дисперсии результативного признака и применяемый для измерения тесноты связи при криволинейной зависимости.

Корреляционное поле – точечный график в прямоугольной системе координат, в котором на оси абсцисс откладывается масштаб для одного признака (x), а на оси ординат – для другого (y).

Корреляционный анализ – метод исследования взаимозависимости признаков в генеральной совокупности, являющихся случайными величинами, имеющими нормальное распределение.

Корреляция – зависимость между случайными величинами, не имеющая строгого функционального характера, при которой изменение одной из случайных величин приводит к изменению математического ожидания другой.

Корреспондентский способ наблюдения – способ организации статистического наблюдения, при котором статистические данные об изучаемых явлениях сообщаются осведомленными, компетентными лицами, не состоящими в штате статистического органа-корреспондентами.

Коэффициент ассоциации – показатель оценки тесноты связи между двумя альтернативными признаками и использующийся при нечисловой информации.

Коэффициент вариации – один из показателей вариации, который является относительной мерой вариации и представляет собой отношение среднего квадратического отклонения к средней величине варьирующего признака и выраженного в процентах.

Коэффициент детерминации – квадрат коэффициента корреляции, который характеризует долю вариации результативного признака под влиянием вариации признака-фактора.

Коэффициент контингенции – показатель сходства используется для изучения зависимости между альтернативными признаками.

Коэффициент корреляции – числовая характеристика совместного распределения двух случайных величин, выражающая их взаимозависимость и измеряющая степень линейной зависимости.

Коэффициент уравнения регрессии – параметр уравнения регрессии, который показывает на сколько изменяется в среднем значение результативного признака при увеличении факторного на единицу собственного измерения.

Коэффициент Фехнера или коэффициент совпадения знаков – это метод исследования взаимосвязей, который основан на применении первых степеней отклонений от средних значений признаков двух связанных рядов показателей.

Коэффициент эластичности – коэффициент в уравнении функциональной зависимости между факторными и результативным признаками, показывающий на сколько процентов изменится результативный признак при увеличении факторного признака на 1 %.

Критическим моментом статистического наблюдения (как правило, переписи) – это момент времени, по состоянию на который производится регистрация сведений.

Кумулята – графическое изображение функции распределения вероятностей, графическое изображение статистического ряда накопленных частот (частостей), где по оси абсцисс откладываются варианты, а по оси ординат – накопленные частоты (частости), показывающие сколько единиц совокупности имеют значения признака, не превосходящие данное значение.

Л

Лаг, *временной лаг, лаг запаздывания* – промежуток времени, за который изменение аргумента приведет к изменению результативного показателя.

М

Макет статистической таблицы – таблица статистическая, не содержащая цифровых данных.

Малая выборка – выборочное наблюдение, численность единиц которого не превышает 30.

Медиана – это численное значение признака у той единицы совокупности, которая находится в середине ранжированного ряда.

Метод аналитических группировок – один из основных методов статистического исследования, заключающийся в расчленении совокупностей, изучаемых статистикой, на группы по определенным существенным признакам.

Метод моментов – один из общих методов оценивания неизвестных параметров распределения, суть которого заключается в том, что некоторые моменты генеральной совокупности как функции неизвестных параметров приравниваются к соответствующим выборочным моментам, после чего система уравнений решается относительно неизвестных параметров.

Метод наименьших квадратов – статистический метод нахождения оценок параметров генеральной совокупности.

Метод основного массива – метод несплошного статистического наблюдения, при котором обследованию подвергаются наиболее крупные, существенные единицы наблюдения.

Метод приведения параллельных рядов – метод статистического исследования, заключающийся в приведении и анализе рядов статистических показателей о взаимосвязанных явлениях.

Метод скользящих средних – *прием,* используемый для анализа рядов динамики с целью выявления основной тенденции изменения их уровней и состоит в замене фактических данных средними арифметическими из нескольких уровней рядов динамики (трех, пяти, семи и т.д. – интервал скольжения), найденных способом скольжения, т.е. постепенным исключением из принятого интервала скольжения первого уровня и включением последующего.

Метод статистики – совокупность приемов, правил и методов статистического исследования социально-экономических явлений, сбор сведений, их обработка, вычисление обобщающих показателей и анализ данных, пользуясь которыми статистика исследует свой предмет.

Методы оценки непараметрические предназначены для оценки характеристик генеральной совокупности с неизвестной формой распределения, в которых в качестве оценок используются порядковые статистики, ранги и выборочные доли вариантов или групп значений изучаемых признаков.

Методы статистики непараметрические – методы математической статистики, не предполагающие знания функционального вида распределений.

Мода – наиболее типичное значение случайной величины, т.е. наиболее часто встречающаяся варианта признака в данной совокупности.

Монографическое обследование – подробное описание отдельных единиц объекта статистического наблюдения, характерных в каком-либо отношении.

Н

Наблюдение выборочное – обследование отобранного в случайном порядке определенного числа единиц генеральной совокупности с целью получения ее обобщающих характеристик.

Наблюдение единовременное – статистическое наблюдение, организуемое в одноразовом порядке или проводимое время от времени без соблюдения строгой периодичности его повторения.

Наблюдение специально организованное – организационная форма статистического наблюдения, которое проводится в целях собирания статистических данных, отсутствующих в статистической отчетности или в целях проверки данных отчетности, а также для решения самостоятельных научно-практических задач.

Наблюдение сплошное – наблюдение всех без исключения единиц изучаемого объекта статистического наблюдения.

Наблюдение статистическое – планомерный, научно организованный сбор данных о явлениях и процессах общественной жизни путем регистрации по заранее разработанной программе наблюдения их существенных признаков.

О

Общая теория статистики – отрасль статистической науки, рассматривающая общие категории, понятия, принципы и ее методы.

Объект статистического наблюдения – совокупность явлений, предметов и т.п. (называемая в статистике совокупностью), подвергаемых статистическому наблюдению.

Объем выборки – число единиц, образующих выборочную совокупность.

Организационный план статистического наблюдения – это составная часть общего плана статистического наблюдения, в которой излагается порядок его организации и проведения – это документ, в котором содержится перечень подготовительных работ и проведения статистического наблюдения с указанием конкретных сроков их проведения.

Отклонение среднее линейное – один из показателей вариации, представляющий собой среднее значение абсолютных отклонений вариантов признака от их средней величины.

Относительная величина (в статистике) – статистическая величина, являющаяся мерой количественного соотношения статистических показателей и отображающая относительные размеры социально-экономических явлений.

Относительная величина выполнения плана – соотношение величины показателя, достигнутой за какое-то время, и величины его, установленной планом.

Относительная величина динамики – соотношение величины показателя за данное время и величины его за какое-либо аналогичное предшествующее время.

Относительная величина интенсивности – соотношение размеров двух качественно различных явлений, характеризующее развитие изучаемого явления или процесса в определенной среде.

Относительная величина координации – соотношение размеров частей некоторого целого между собой.

Относительная величина планового задания – соотношение величины показателя, устанавливаемой на планируемый период к величине его, достигнутой к планируемому периоду.

Относительная величина сравнения – соотношение величин одноименных показателей, относящихся к разным объектам или разным территориям.

Относительная величина структуры – соотношение величины части какого-либо целого и величины этого целого.

Относительная величина уровня экономического развития – соотношение величины важнейших экономических показателей и численности населения, т.е. на душу населения.

Отчетность статистическая – форма статистического наблюдения, при которой соответствующие органы получают от предприятий, организаций и учреждений необходимые им статистические данные в виде установленных в законном порядке отчетных документов за подписями лиц, ответственных за представление и достоверность сообщаемых сведений.

Ошибка регистрации – расхождение между зафиксированными при статистическом наблюдении значениями признака у единицы наблюдения и его действительным значением.

Ошибка репрезентативности – расхождение между значениями изучаемого признака выборочной совокупности и совокупности генеральной.

П

Перепись – один из видов специально организованного наблюдения, проводимого с целью исчисления численности и состава объекта статистического наблюдения по ряду характерных для него признаков не собираемых в порядке статистической отчетности.

Период (срок) наблюдения – это время, в течение которого осуществляется регистрация единиц наблюдения по установленной программе.

Период базисный – период времени или момент, с данными которого сопоставляются данные другого периода (текущего, отчетного).

Период отчетный (текущий) – период времени, за который представляется статистическая отчетность, или период времени, данные за который сопоставляются с данными за другой период (базисный).

Подлежащее таблицы – это единицы статистической совокупности или их группы, характеризуемые цифровыми данными, сказуемым таблицы.

Показатели вариации – показатели, отображающие размеры вариации (степень колеблемости) признака.

Показатель качественный – величина какого-либо показателя в расчете на единицу другого показателя, характеризующая уровень признака, явления.

Показатель объемный (количественный) – показатель, представляющий собой объем признака или объем совокупности.

Поле графика – эта пространство, в котором размещаются геометрические знаки, образующие график.

Полигон распределения – одна из форм графического изображения вариационных рядов как дискретных, так и интервальных (многоугольник распределения), где на оси «x» в прямоугольной системе координат отмечают значения варьирующего признака, а на оси «y» соответствующие им частоты (частости) или «X» – возможные значения дискретной случайной величины, а « y » соответствующие вероятности их появления.

Предмет статистики – размеры и количественные соотношения массовых общественных явлений в конкретных условиях места и времени.

Признак *(в статистике)* – отличительная черта, свойство, качество, присущие единице совокупности (как и единице наблюдения), изучаемые статистикой.

Признак атрибутивный – признак, отдельные значения которого выражаются в виде понятий, наименований.

Признак варьирующий – признак, принимающий в пределах статистической совокупности разные значения у единиц, ее составляющих.

Признак группировочный – признак, принимаемый за основу образования групп в процессе статистической группировки.

Признак количественный – признак, отдельные значения которого имеют количественное выражение.

Признак осередняемый – признак, средняя величина которого исчисляется.

Признак результативный – признак зависимый, т.е. изменяющий свое значение под влиянием другого, связанного с ним и действующего на него факторного признака.

Признак факторный – признак – фактор, оказывающий влияние на другой, связанный с ним признак результативный, и обусловливающий его изменение – вариацию.

Признаки альтернативные – признаки, которыми одни единицы совокупности обладают, а другие – нет, т. е. с противоположными (взаимоисключающими) характеристиками.

Прирост абсолютный – разность двух уровней ряда динамики.

Программа наблюдения – это перечень признаков единицы наблюдения, регистрируемых в процессе статистического наблюдения.

Пространственные ориентиры – размещение знаков на поле графика.

Р

Размах вариации **–** один из показателей вариации, характеризующий пределы колеблемости (вариацию) индивидуальных значений (или вариантов) признака в статистической совокупности и определяемый как разность между максимальным и минимальным значениями признака.

Ранг – это порядковый номер значений признака, расположенных в порядке возрастания или убывания их величины.

Ранжирование – это процедура упорядочения объектов изучения, которая выполняется на основе предпочтения.

Ряд динамики, хронологический ряд – ряд последовательно расположенных в хронологическом порядке значений показателя, который в своих изменениях отражает ход развития изучаемого явления.

С

Сводка – второй этап статистического исследования, состоящий в систематизации, обработке и подсчете групповых и общих итогов, расчете производных величин.

Сводка децентрализованная – способ организации сводки статистических данных, состоящий в том, что обработка первичных данных, полученных в результате наблюдения, производится на местах.

Сводка централизованная – способ организации сводки статистических данных, при котором все первичные материалы, полученные в результате статистического наблюдения, сосредотачиваются в центральном органе, где и подвергаются сводке.

Сглаживание (в статистике) – метод исследования рядов статистических данных о социально-экономических явлениях и заключающийся в нахождении расчетных значений их показателей (уровней) и замене ими фактических с целью выявления закономерностей развития процессов, отображаемых этими рядами.

Сезонность – понятие, характеризующее регулярно повторяющиеся изменения явлений в динамике, связанные со сменой времен года (зима, весна, лето, осень), явлениями природы (период дождей, период созревания растений), выполнением определенных работ и занятий (сезон охоты, лечебный сезон, сезон уборки урожая), а также с обычаями, традициями и праздниками (увеличение свадеб осенью, повышение спроса на цветы в праздничные дни и т.п.).

Сказуемое таблицы – цифровые данные, характеризующие подлежащее, т.е. это система показателей, которыми характеризуется объект изучения.

Случайная величина – переменная величина, принимающая одно из возможных значений в зависимости от случайных обстоятельств.

Смыкание рядов динамики – один из приемов объединения двух или более рядов динамики, характеризующих изменение одного и того же явления, в один (более длинный) и применяемый в случаях, когда уровни рядов динамики несопоставимы в связи с территориальными, ведомственными, организационными изменениями, методологии исчисления показателей и т.п.

Совокупность выборочная – совокупность единиц, отобранных по определенным правилам из генеральной совокупности для статистического наблюдения.

Совокупность генеральная – вся совокупность реально существующих объектов, из которых тем или иным способом извлекается выборочная совокупность.

Совокупность неоднородная *(качественно неоднородная)* – статистическая совокупность, в которой элементы (единицы), ее составляющие, относятся к различным типам изучаемого явления.

Совокупность однородная – статистическая совокупность, в которой ее составные элементы (единицы) сходны между собой по существенным для данного исследования признакам и относятся к одному и тому же типу явления.

Совокупность статистическая – множество объектов, явлений, объединенных какими-либо общими свойствами (признаками) и подвергающихся статистическому исследованию.

Среднее квадратическое отклонение – это обобщающая характеристика размеров вариации признака в совокупности; оно показывает, на сколько в среднем отклоняются конкретные варианты от их среднего значения; является абсолютной мерой колеблемости признака и выражается в тех же единицах, что и варианты.

Среднее линейное отклонение представляет собой среднюю арифметическую величину из абсолютных отклонений отдельных вариантов от средней арифметической.

Средняя арифметическая величина – одна из форм средней величины, определяемая как частное от деления суммы вариантов признак на их число (простая средняя арифметическая) или суммы взвешенных вариантов признак на сумму весов (взвешенная средняя арифметическая).

Средняя величина *(в статистике)* – это обобщенная количественная характеристика признака в статистической совокупности, выражающая характерную, типичную величину одного варьирующего признака у единиц совокупности, образующихся в данных условиях места и времени под влиянием всей совокупности факторов.

Средняя гармоническая – одна из форм средней величины, представляющая обратную величину средней арифметической из обратных значений признака.

Средняя геометрическая – одна из форм средней величины, вычисляемая как корень n степени из произведения отдельных вариантов признака.

Средняя хронологическая – средняя величина из уровней ряда динамики, вычисляемая для моментных рядов с равным интервалом.

Статистическая наука – это отрасль знаний, изучающая явления общественной жизни с их количественной стороны в неразрывной связи с их качественным содержанием в конкретных условиях места и времени.

Статистическая практика – это деятельность по сбору, накоплению, обработке и анализу цифрового материала.

Статистическое исследование – процесс познания (изучения) социально-экономических явлений посредством системы статистических методов и количественных характеристик – системы показателей.

Стохастическая форма связи – форма связи, при которой каждому значению одного признака (факторного) соответствует целый ряд значений другого признака (результативного), стохастическая связь проявляется не в каждом отдельном случае, а лишь в среднем для совокупности явлений данного вида.

Т

Таблица групповая – вид статистической таблицы, подлежащее которой состоит из групп единиц изучаемой совокупности по одному признаку.

Таблица комбинационная – вид статистической таблицы, подлежащее которой состоит из групп и подгрупп, образованных по двум и более признакам.

Таблица простая – вид статистической таблицы, в подлежащем которой содержится название (перечень) характеризуемых при помощи статистических данных объектов или единиц совокупности.

Таблица разработочная – вспомогательная таблица, предназначенная для получения подробных данных при обработке информации, собранной в результате наблюдения.

Таблица сопряженности признаков – табличное представление совместного распределения двух качественных признаков.

Таблица статистическая – форма рационального, наглядного изложения статистических данных о явлениях и процессах общественной жизни.

Темп прироста – относительный показатель динамики, выражаемый в процентах и представляющий собой отношение абсолютного прироста к уровню динамики, по сравнению с которым он рассчитан.

Темп роста – относительный показатель динамики, выражаемый в процентах и представляющий собой отношение уровней ряда динамики.

Теснота связи – качественная характеристика степени зависимости между случайными величинами (признаками).

Точность статистического наблюдения – близость данных об объекте статистического наблюдения, полученных в результате статистического наблюдения, к их действительным значениям.

Тренд – изменение, определяющее общее направление развития, основную тенденцию временных рядов (рядов динамики).

У

Уравнения регрессии – различные способы аппроксимации истинной регрессионной зависимости.

Уровень значимости – одна из характеристик качества критерия статистической проверки гипотез.

Уровень ряда динамики – числовое значение показателя в ряду динамики.

Ф

Федеральная служба государственной статистики (Росстат) – центральный уполномоченный федеральный орган исполнительной власти, осуществляющий функции по формированию официальной статистической информации о социальном, экономическом, демографическом и экологическом положении страны.

Формуляр статистический – бланк (опросной лист, переписной лист, форма отчетности, анкета), содержащий вопросы программы наблюдения и место для ответов на них.

Функциональная связь – это связь, где каждому значению одной переменной (аргументу) соответствует одно вполне определенное значение другой переменной (функции).

Ч

Частость **–** относительная величина, определяющая долю частот отдельных вариантов или интервалов в общей сумме частот, сумма которых равна единице.

Частота **–** абсолютное число, показывающее, сколько раз (как часто) встречается в совокупности то или иное значение признака.

Ш

Шкала графика – совокупность линий, отметок и поставленных у некоторых из них чисел отсчета или других символов, соответствующих ряду последовательных значений изображаемых величин.

Шкала двойная – две системы последовательных числовых значений, соответствующих явлениям или процессам, изображаемым на графике.

Шкала масштабная – линия с нанесенными на нее масштабными отметками и их числовыми значениями.

Э

Экспедиционный способ наблюдения **–** способ статистического наблюдения, осуществляемого специально подготовленными для этой цели лицами.

Экспликация графика – это словесное объяснение содержания графика и смыслового значения каждого знака на графике.

Экстраполяция – нахождение значений функции за пределами ее области определения с использованием информации о поведении данной функции в некоторых точках, принадлежащих области ее определения.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Аскеров П.Ф., Пахунова Р.Н., Пахунов А.В. Общая и прикладная статистика: Учебник для студентов высшего профессионального образования / Под общ. ред. Р.Н. Пахуновой. – М.: НИЦ ИНФРА-М, 2013. – 272 с. + Доп. материалы [Электронный ресурс; Режим доступа http://www.znanium.com]. – (Высшее образование. Бакалавриат). DOI 10.12737/748 (www.doi.org).

ISBN 978-5-16-006669-1 (print)

ISBN 978-5-16-100304-6 (online)

(http://znanium.com/bookread2.php?book=404310).

2. Годин А.М. Статистика: Учебник / А.М. Годин. – 10-е изд., перераб. и испр. – М.: Издательско-торговая корпорация «Дашков и К°», 2013. – 452 с. – ISBN 978-5-394-01494-9.

(http://znanium.com/bookread2.php?book=430372).

3. Елисеева И.И. Статистика: Учебник. – Москва: Проспект, 2012.

4. Едронова В.Н., Малафеева М.В. Общая теория статистики: учебник. – 2-е изд., перераб. и доп. – М.: Магистр, 2015. – 608 с.

ISBN 978-5-9776-0011-8 (в пер.).

ISBN 978-5-16-010210-8.

Агентство CIP РГБ

(http://znanium.com/bookread2.php?book=474554).

5. Назаров М.Г. Общая теория статистики: Учебник. – ОМЕГА-Л, 2010.

6. Орехов С.А. Статистика: Учебник. – ЭКСМО, 2010.

7. Теория статистики: Учебник / Под ред. проф. Г.Л. Громыко. – 3-е изд., перераб. и доп. – М.: ИНФРА-М, 2016. – 476 с. (Высшее образование: Бакалавриат).

ISBN 978-5-16-004857-4 (print)

ISBN 978-5-16-104508-4 (online)

(http://znanium.com/bookread2.php?book=547988).

СОДЕРЖАНИЕ