ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УПРАВЛЕНИЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ И ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

Кафедра «Экономическая безопасность, учет и право»

Учебное пособие

«Статистика. Теория статистики»

Автор
Сидорина Т.В.

Ростов-на-Дону, 2022

Аннотация

В учебном пособии изложены основные теоретические положения учебного материала по курсу «Статистика. Теория Статистики», иллюстрируемые примерами по каждой теме курса.

Представленное учебное пособие позволит помочь обучающимся наиболее результативно организовать работу по овладению статистическими методами, выполнению курсовой работы и практических заданий.

Учебное пособие подготовлено в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по специальности 38.05.01 Экономическая безопасность (уровень специалитета).

Автор

к.э.н., доцент кафедры «ЭБУиП» Сидорина Т.В.

Оглавление

ВВЕДЕНИЕ

Статистика – это фундаментальная дисциплина, направленная на формирование научного экономического мировоззрения, умения собирать, обрабатывать и анализировать информацию о социально-экономических процессах и явлениях.

Дисциплина «Статистика» является дисциплиной базовой дисциплиной, формирующая профессиональный уровень современного специалиста в области экономики и финансов.

Учебное пособие представлено разделом «Теория статистики».

Для всестороннего анализа происходящих в стране процессов, необходимо овладеть статистической методологией и, прежде всего, основами «Теория статистики» – науки, разрабатывающей понятия и категории статистики, методы отбора, обработки, обобщения и анализа массовых общественных явлений, структурных различий, методы изучения динамики и анализа закономерностей развития, моделирования и прогнозирования конкретных социально-экономических процессов.

Основной принцип, лежащий в основе изучения статистики, состоит, с одной стороны, в повышении уровня фундаментальной статистической подготовки студентов, а с другой – в усилении ее прикладной направленности.

Цель учебного пособия – помочь студентам в усвоении методов общей теории статистики, связанных с проведением статистического наблюдения, сводки, группировки его материалов, с исчислением статистических величин и их анализом.

Учебное пособие состоит из десяти тем курса теории статистики, каждая из которых содержит краткий теоретический материал, примеры решения типовых задач. Учебное пособие подготовлено в соответствии с ФГОС ВО и рассчитано на обучающихся по специальности 38.05.01 Экономическая безопасность.

1 РАЗДЕЛ. ОПИСАТЕЛЬНАЯ СТАТИСТИКА

1 СТАТИСТИКА КАК НАУКА

Термин «статистика» произошел от латинских слов stato (государство) status (положение вещей, политическое состояние) и впервые был употреблен немецким ученым Г. Ахенвалем в 1749 г., в его книге о государствоведении.

В настоящее время термин "статистика" употребляют обычно в трех значениях.

Статистика — это наука, изучающая количественную сторону массовых явлений и процессов в неразрывной связи с их качественной стороной, количественное выражение закономерностей общественного развития в конкретных условиях места и времени.

Статистика — это отрасль практической деятельности по сбору, накоплению, обработке и анализ цифровых данных, характеризующих население, экономику, культуру, образование и другие явления общественной жизни и предназначенную для задач государственного регулирования и управления.

Статистика — это собственно данные (цифровой материал), который обрабатывается определенными методами.

Предметом статистики выступают количественные характеристики и соотношения качественно определенных социально-экономических явлений, закономерности их связей и развития в конкретных условиях места и времени.

Задачи статистики в современных условиях:

1) исследование происходящих в обществе преобразований социальных и экономических процессов на основе системы специальных показателей;

2) обобщение и прогнозирование тенденций развития народного хозяйства и его составляющих;

3) влияние имеющихся резервов эффективности общественного производства;

4) создание единого информационного пространства органов государственной власти;

5) организация статистики отраслей народного хозяйства и общества (прикладной статистики).

Статистика состоит из двух разделов.

1. Теория статистики – наука, разрабатывающая понятия и категории статистики, методы сбора, обработки, обобщения и анализа данных, характеризующих массовые общественные явления, а также методы анализа взаимосвязей и закономерностей развития социально-экономических процессов. Теория статистки — методологическая основа всех отраслевых (прикладных) статистик: экономической; социальной; труда; государственной; финансов.

2. Социально-экономическая статистика – занимающаяся количественной оценкой, обобщением и истолкованием системы показателей, характеризующих уровни, структуру, динамику и взаимосвязи социально-экономических процессов с целью выявления общих закономерностей развития национальной экономики, ее отдельных звеньев (отраслей, секторов, предприятий), а также социальной сферы.

Статистика – это экономическая и методологическая наука.

Она разрабатывает методологические принципы экономического анализа и формирует систему показателей.

Организация и функции статистических служб

В России в 1811 г. при департаменте полиции было образовано статистическое отделение, в 1857 г. — Центральный статистический комитет, губернские и земские Комитеты, с 25 июля 1918 г. — Центральное статистическое управление (ЦСУ).

Система государственной статистики имеет иерархическую структуру, включающую в себя следующие уровни: федеральный, республиканский, краевой, областной, окружной, городской и районный.

Система государственной статистики в России организована в соответствии с административно-территориальным делением страны и подчиняется таким организационным принципам, как:

1) централизованное руководство;
2) единое организационное строение и методология;
3) неразрывная связь с органами государственного управления.

Наивысшим органом управления статистикой в России является Федеральная служба государственной статистики Российской Федерации.

Федеральная служба государственной статистики (Росстат, ранее Госкомстат) является федеральным органом исполнительной власти, осуществляющим функции по формированию официальной статистической информации о социальном, экономическом, демографическом и экологическом положении страны, а также функции по контролю и надзору в области государственной статистической деятельности на территории Российской Федерации.

В настоящее время деятельностью Службы руководит Правительство РФ. (Прежде Росстат находился в ведении Минэкономразвития России.)

Федеральная служба государственной статистики осуществляет свою деятельность непосредственно и через свои территориальные органы. В их состав входят 82 территориальных органа.

Основные категории и понятия статистики

Статистическая совокупность – это множество варьирующих объектов, явлений, объединенных какими-либо общими свойствами и подвергающихся статистическому исследованию.

Единица совокупности – это каждый элемент статистической совокупности называется

Признак – отличительная черта, свойство, качество присущие единице совокупности, и учитываемое при статистическом исследовании называются

Классификация признаков в зависимости от характера которых зависит выбор применяемых в исследовании методов.

1. *По характеру выражения:*

– Описательные (качественные) – признаки, значения которых имеют выражение в форме понятий, наименований. Описательные признаки, в свою очередь, делятся на номинальные и порядковые. Их отличие состоит в том, что номинальные признаки нельзя ранжировать, тогда как с помощью порядковых признаков данные можно ранжировать, упорядочивать.

– Количественные. Количественные признаки – признаки, значения которых выражаются в форме чисел. Например, число студентов в группе, количество машин в автохозяйстве и т.д.

2. *По способу измерения:*

– Первичные – характеризующие абсолютные размеры социально-экономических явлений, выражающиеся в единицах меры протяженности, площади, массы (веса) и т.п., в единицах счета времени, в денежных единицах или в виде числа единиц совокупности.

– Вторичные (или расчетные) – признаки, образующиеся в результате соотношения первичных признаков.

3. *По отношению к характеризуемому объекту:*

– Прямые (непосредственные) – это признаки, выражающие качества, непосредственно присущие единице совокупности.

– Косвенные – отражают качества, не принадлежащие непосредственно объекту статистического исследования, они характеризуют другие совокупности, имеющие отношение к объекту.

4. *По характеру вариации:*

– Альтернативные – признаки, значения которых выражаются только через два значения.

– Дискретные – признаки, выражающиеся в форме числа и принимающие только определенное значение.

– Непрерывные – признаки, значения которых могут изменяться без ограничений.

5. *По отношению ко времени:*

– Моментные – признаки, содержащие величины на определенный момент времени, например, на начало (конец) года, квартала, месяца т.п.

– Интервальные – признаки, у которых значения отражаются за определенный период, например, за год, за квартал, за месяц.

Исследование изменений общественных явлений, выраженных с количественной стороны, и определенных в качестве предмета статистической науки, производится с помощью статистических показателей.

Статистическим показателем называют количественно-качественную характеристику социально-экономических явлений и процессов.

Единица измерения – значение, в котором выражается и с которым сравнивается исследуемая величина. Единицы измерения статистического показателя определяются его содержанием. Например, число родившихся младенцев, рассчитанное на каждую тысячу жителей региона, будет измеряться в промилле, которое обозначается как «‰».

Классификация статистических показателей.

I. Применительно *к содержанию* показатели делятся следующим образом:

1. *По охвату единиц совокупности:*

– индивидуальные – показатели, характеризующие исследуемый процесс по одной единице совокупности.

– сводные – характеризуют общественное явление по группе исследуемых единиц; они делятся на объемные и расчетные.

2. *По временному фактору:*

– плановые – показатели, величина которых отражает уровень изучаемого явления, который должен быть достигнут в соответствии с планом;

– отчетные – показатели, величина которых отражает уровень изучаемого явления, достигнутый в исследуемом периоде (если признак моментный) или за исследуемый период (если признак интервальный);

– базисные – показатели, величина которых принимается в качестве базы для сравнения.

3. *По отношению к характеризуемому свойству:*

– прямые – показатели, которые непосредственно характеризуют изучаемое свойство;

– обратные.

II. Применительно *к форме выражения* статистические показатели делятся на:

– абсолютные – показатели, отражающие свойства явления, выраженные первичными признаками;

– относительные – показатели, получающиеся путем соотношения абсолютных показателей;

– средние – показатели, характеризующие величину изучаемого признака, приходящуюся на единицу совокупности.

Признак определяет качественное содержание объекта исследования;

По характеру влияния друг на друга выделяют признаки:

1. Признак-результат – признак, анализируемый в данном исследовании, рассматривается как следствие взаимодействия других факторов.

2. Признак-фактор – признак, оказывающий влияние на исследуемый признак (признак-результат).

2 СТАТИСТИЧЕСКОЕ ИЗМЕРЕНИЕ И НАБЛЮДЕНИЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Статистическое наблюдение – первый этап любого статистического исследования.

Статистическое наблюдение – это планомерный, научно организованный и систематический сбор данных о явлениях и процессах общественной жизни путем реорганизации заранее намеченных существенных признаков с целью получения в дальнейшем обобщающих характеристик этих явлений и процессов.

Статистическое наблюдение – это сбор данных (фактов, сведений) об изучаемых явлениях.

Его главная задача заключается в получении достоверной информации для выявления закономерностей развития явлений и процессов.

Необходимые свойства статистического наблюдения: массовость, достоверность, систематичность.

Этапы статистического наблюдения:

1. Программно-методологическая подготовка проведения наблюдения.
2. Организационная подготовка проведения наблюдения.
3. Сбор данных статистического наблюдения.
4. Контроль качества данных статистического наблюдения.
5. Выработка выводов и предложений по усовершенствованию статистического наблюдения.

При подготовке к проведению статистического наблюдения решаются программно-методологические и организационные вопросы.

Программно-методологические вопросы включают в себя формулировку задачи наблюдения, определения объекта и единиц наблюдения, а также составление программ наблюдения. Объектом наблюдения называют явления или совокупность явлений, информацию о которых собирают в процессе наблюдения.

Статистическое наблюдение необходимо проводить по строго определенному плану, включающему программно-методологические и организационные вопросы.

Объект статистического наблюдения – это ограниченное в пространстве и во времени определенное целостное множество взаимосвязанных единиц наблюдения, о котором должны быть собраны статистические сведения.

Единица статистического наблюдения – это составной неделимый элемент объекта наблюдения, являющийся основой учета и носителем определенного круга признаков, наличие (или отсутствие) которых у каждой единицы изучаемой совокупности должно быть зафиксировано в процессе статистического наблюдения.

Каждая единица наблюдения характеризуется определенными чертами, свойствами, которые в статистике называются признаками.

Программа наблюдения – перечень вопросов, ответы на которые получают в процессе наблюдения (план, включающий программно-методологические вопросы наблюдения).

Вопросы программы статистического наблюдения различаются по: форме, функциям и содержанию.

Для решения организационных вопросов составляется организационный план статистического наблюдения, определяющий цель, вид, форму, способ наблюдения, место и сроки его проведения.

В результате статистического наблюдения получают первичные данные о единицах совокупности, которые на следующем этапе статистического исследования – этапе сводки – обобщаются в группы, систематизируются.

Для проведения статистического наблюдения разрабатывается **инструментарий наблюдения**, который включает в себя ***формуляр и инструкцию.***

По форме статистическое наблюдение может быть проведено через отчетность и специально организованные наблюдения.

Отчетность – представляет собой перечень показателей, характеризующих деятельность предприятий и организаций, представленный в виде форм отчетности, утвержденных Госкомстатом РФ и обязательно предоставляемые в установленные сроки в районные управления статистики.

Специально организованные наблюдения – представляют собой сбор средств посредством переписи, специальных обследований, проводимых работниками органов статистики.

Исследуемые факты могут относиться либо к какому-либо периоду (добыча угля за год, количество родившихся за месяц, год и т.д.), либо к определенному моменту времени (численность населения на начало года, автомобильный парк на конец года и т.д.). Та дата, на которую регистрируются факты, называется **критическим моментом**.

Период, в течение которого проводится сбор данных, называется **временем наблюдения**.

Статистические данные могут быть получены разными способами: путем непосредственного измерения (подсчета, взвешивания и т.д.), опроса и пр.

Виды статистического наблюдения:

По времени регистрации фактов:

– непрерывное (текущее) (учет кадров);

– периодическое (регистрация по мере надобности) (бухгалтерский учет);

– единовременное (перепись жилого фонда).

По степени охвата единиц совокупности:

– сплошное, при котором регистрации подлежат все без исключения единицы изучаемой совокупности;

– несплошное, в процессе которого регистрации подвергается какая-то часть единиц совокупности.

Несплошное наблюдение осуществляется различными способами:

– основного массива, при котором обследованию подвергается основной массив, та часть единиц, которая вносит наибольший вклад в изучение явления;

– выборочное, при котором из всей массы объекта обследования в случайном порядке отбирается для изучения отдельная часть единиц совокупности, полученные при этом результаты распространяются на весь объект;

– монографическое, при котором детально, подробно обследуются отдельные единицы изучаемого объекта, являющиеся представителями каких-либо новых типов явлений.

Способы статистического наблюдения:

– непосредственное наблюдение характеризуется тем, что признаки единиц наблюдения устанавливают сами регистраторы на основе личного осмотра, подсчета, промера, взвешивания;

– документальный – основан на использовании в качестве источника статистических сведений первичных документов учетного характера;

– опрос, т.е. получение информации от непосредственного носителя признаков (респондента) с его слов.

Виды опроса:

– устный (экспедиционный), при котором регистраторы сами фиксируют ответы в формуляре;

– саморегистрации – формуляры (бланки опросного листа) заполняют сами респонденты;

– анкетный, при котором сбор информации осуществляется на основе рассылки специальных вопросников (анкет) лично респондентам либо публикации в печатных изданиях.

В практике статистики используются как карточные бланки (заполняемы на каждую единицу отдельно), так и списочные (заполняемые для нескольких единиц одновременно).

Собранные в процессе наблюдения сведения должны быть подвергнуты проверке путем логического и арифметического контроля.

3 ОБОБЩЕНИЕ И ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ

Сводка статистических данных

Проверенные материалы первичного учета необходимо систематизировать и обобщить, т.е. вычислить статистические показатели для заполнения форм отчетности и изучения деятельности организации. Второй стадией статистического исследования является статистическая сводка.

Статистическая сводка – это комплекс последовательных операций по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Она состоит в систематизации первичных данных, образовании статистической совокупности и получении итоговых абсолютных обобщающих показателей, группировке данных по количественным и качественным признакам и представлении их в табличной или графической формах.

Результатом сводки является сводная характеристика всей совокупности.

Виды сводок по:

– *глубине и точности обработки:* простая, сложная;
– *форме обработке материала:* централизованная, децентрализованная;
– *технике выполнения:* механизированная, ручная.

Статистическая сводка проводится по определенной **программе**, разрабатываемой исходя из целей и задач статистического исследования, которая предусматривает:

– выбор группировочных признаков;
– определение порядка формирования групп;
– разработка системы показателей для характеристики групп и явления (объекта) в целом;
– подсчет групповых и общих итогов;
– оформление конечных результатов сводки в статистических таблицах.

Статистическую сводку выполняют последовательными этапами.

Первый этап – объединение единиц по интересующему нас признаку в качественно однородные группы – **статистическая группировка.**

Второй этап статистической сводки – подсчет абсолютных значений по каждой группе и совокупности в целом.

Третий этап статистической сводки – запись результатов подсчетов по группам и совокупности в статистические таблицы.

Группировка статистических данных

Группировка – разделение единиц статистической совокупности на группы, однородные по какому-либо одному или нескольким признакам.

Группировка позволяет упорядочить статистическую совокупность, пригодную для дальнейшего статистического анализа.

Значения, которые может принимать тот или иной варьирующий признак, называются вариантами.

По характеру вариантов признаки делятся на атрибутивные и количественные. Если варианты признака не выражаются числами, то этот признак называют атрибутивным. Если варианты выражены в виде чисел, то это количественный признак.

Виды группировок:

– количественные, когда она получена по количественному признаку;
– типологические;
– пространственные, созданные по географическому признаку;
– аналитические, предназначенные для исследования зависимости между явлениями.

Последовательность построения группировки:

1) определение группировочного признака (основания группировки);

2) определение количества групп, на которые необходимо разбить совокупность;

3) установление интервала группировки.

К группировкам примыкают **классификации**, которые представляют собой систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основании их сходства и различия.

От группировок классификации отличаются тем, что в их основу кладется только качественный признак, они стандартны и устойчивы. Классификации устанавливаются органами государственной и международной статистики, они едины для любого исследования и неизменны в течение длительного времени.

Интервал – это значение варьирующего признака, лежащее в определенных границах. Каждый интервал имеет свою величину, верхнюю и нижнюю границы или хотя бы одну из них. **Нижней границей** интервала называется наименьшее значение признака в интервале, а **верхней границей –** наибольшее значение признака в интервале.

Величина интервала представляет собой разность между верхней и нижней границами интервала.

Интервалы группировки в зависимости от их величины бывают равные и неравные. Последние делятся на прогрессивно возрастающие, прогрессивно убывающие, произвольные и специализированные. Возможна группировка с равными и неравными интервалами.

Если вариация признака проявляется в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с **равными интервалами.**

Величина равного интервала:

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

где $R = x_{\max} - x_{\min}$, т.е. размах вариации;

$x_{\max}$ – наибольшее значение варьирующего признака;

$x_{\min}$ – наименьшее значение варьирующего признака.

Определение величины интервала по формуле Стерджесса:

$$h = \frac{x_{\max} - x_{\min}}{1 + 3.322 \lg N},$$

где N – численность единиц совокупности.

Величина неравных интервалов:

а) изменяющихся в арифметической прогрессии:

$$h_{i+1} = h_i + a;$$

б) изменяющихся в геометрической прогрессии:

$$h_{i+1} = h_i + q,$$

где а – константа – число, которое будет положительным при прогрессивно возрастающих интервалах и отрицательным при прогрессивно убывающих интервалах;

q – константа – положительное число, которое при прогрессивно – возрастающих интервалах будет больше 1, а при прогрессивно – убывающих – меньше 1.

Пример 1. Имеются данные о среднегодовой стоимости имущества по 30 организациям, млн р.:

6,1; 3,2; 55; 5,4; 3,1; 7,9; 4,5; 5,8; 6,5; 5,2; 3,8; 5,4; 4,8; 5,6; 4,2; 3,4; 7,6; 6,8; 4,9; 5,2; 6,3; 4,1; 5,6; 7,3; 6,7; 5,4; 5,7; 5,6; 4,3; 5,9.

Постройте интервальный ряд распределения по среднегодовой стоимости имущества, образовав 5 групп с равными интервалами.

Решение

Для построения интервального ряда определяется величина интервала:

$$h = \frac{x_{\max} - x_{\min}}{n} = \frac{7{,}9 - 3{,}1}{5} = 0{,}96$$ млн р.

Таблица 1 – Распределение организаций по среднегодовой стоимости имущества

Группы организаций по среднегодовой стоимости имущества, млн р.	Количество организаций в группе		Стоимость имущества организаций группы, млн р	Средняя стоимость имущества организации группы, млн р.
	в абсолютном выражении	в %		
3,1 – 4,06	4	13,33	13,5	3,375
4,06 – 5,02	6	20,00	26,8	4,467
5,02 – 5,98	12	40,00	66,3	5,525
5,98 – 6,94	5	16,67	32,4	6,48
6,94 – 7,9	3	10,00	22,8	7,6
Итого	30	100,00	161,8	5,393

Группировки, построенные за один и тот же период времени для различных регионов или для одного региона для разных периодов времени, могут оказаться несопоставимы из-за разного количества групп и размеров интервалов в группировках. В этом случае используют перегруппировку данных – **вторичную группировку**, которая представляет собой процедуру образования новых групп на основе ранее осуществленной группировки.

Для **образования новых групп используют методы:**

1) *объединения первоначальных интервалов* – применяется при совпадении границ новых и старых группировок;
2) *долевой перегруппировки данных* применяется при несовпадении границ новых и старых групп и основывается на принципе равномерности распределения единиц наблюдения внутри интервальных групп

Пример 2. Имеются данные о распределении работников организации по размеру заработной платы.

Таблица 2 – Распределение работников по размеру заработной платы

Номер интервала	Заработная плата, р.	Численность работников, чел
1	9000-10000	17
2	10000-11000	20
3	11000-12000	30
4	12000-13000	27
5	13000-14000	32
6	14000 -15000	10
7	15000 и выше	17
Итого		153

Задание. Осуществить перегруппировку данных таблицы 2.2. образовав новые интервалы: «9000-11000», «11000-13000», «13000-15000» и «15000 и выше».

Решение

Таблица 3 – Распределение работников организации по размеру заработной платы (вторичная группировка методом объединения первоначальных интервалов)

Номер интервала	Заработная плата, р.	Численность работников, чел
1	9000-11000	37 (17+20)
2	11000-13000	57 (30+27)
3	13000-15000	42 (32+10)
4	15000 и выше	17
Итого		153

Пример 3. Перегруппировать данные таблицы 2.2 образовав новые интервалы: «9000-10600», «10600-11400», «11400-14900», «14900- и выше».

Решение

Для образования первого интервала «9000-10600», в него включаем все 17 единиц из исходной совокупности («9000-10600»), а также часть единиц из второго исходного интервала («10000-11000»). Новая граница «10000» первого нового интервала разбивает второй интервал на два отрезка «10000-10600» и «10600-11000». Находим, какую долю составляет длина отрезка «10000-10600»» от длины второго интервала «10000-11000»: ($\frac{10600-10000}{11000-10000}=\frac{6}{10}$), значит от 20 единиц, находящихся во втором интервале в исходной группировки следует взять для нового первого интервала 12 единиц ($\frac{6}{10}*20=12$). Тогда первый новый интервал будет содержать 29 единиц (17+12).

Во второй новый интервал «10600-11400», войдут оставшиеся от второго интервала исходной группировки 8 единиц (20-12) и часть единиц из третьего интервала. Для этого находим, какую долю составляет отрезок «10600-11400» от длины третьего интервала «11000-12000» ($\frac{11400-10600}{12000-11000}=\frac{8}{10}$), значит от 30 единиц, находящихся в третьем интервале в исходной группировке следует взять для нового второго интервала 24 единицы

($\frac{8}{10}*30=24$). Тогда второй новый интервал будет содержать 32 единицы (8+24) и т.д.

Результаты вторичной группировки в новых интервалах представим в таблице 4.

Таблица 4 – Распределение работников организации по размеру заработной платы (методом долевой перегруппировки)

Номер интервала	Заработная плата, р.	Численность работников, чел
1	9000-10600	29
2	10600-11400	32
3	11400-14900	74
4	14900 и выше	18
Итого		153

Статистические таблицы и графики

Статистическая таблица – это форма рационального и наглядного изложения результатов сводной обработки материалов статистического наблюдения.

Основные элементы статистической таблицы:

подлежащее – объект исследования (перечень единиц статистической совокупности или их групп по каким-либо признакам);

сказуемое – Система показателей, которыми характеризуется объект исследования, т.е. подлежащее.

Подлежащее обычно располагается в левой части таблицы, сказуемое – в верхней части таблицы в виде названия граф.

Подлежащее обычно располагается в левой части таблицы, сказуемое – в верхней части таблицы в виде названия граф.

Таблицы различаются по цели исследования, построению подлежащего и разработке сказуемого.

Вид статистической таблицы в зависимости от построения подлежащего:

простые – в подлежащем даётся перечень каких-нибудь объектов или территориальных единиц);

групповые – (подлежащее, т.е. объект исследования, подразделяется на группы по какому-либо одному признаку);

комбинационные – (подлежащее разделяется на группы по двум и более признакам).

Разработка сказуемого таблицы, может быть простой и сложной.

Простая разработка – сказуемого означает последовательное перечисление показателей, характеризующих подлежащее.

При сложной разработке сказуемого признаки, характеризующее подлежащее, берутся в сочетании, комбинации.

Основа статистической таблицы

Содержание строк	**Наименование граф (верхние заголовки)**				Итоговая **графа**
А	1	2	...	n – 1	n
Наименование строк (боковые заголовки)					
Итоговая строка					

При построении статистических таблиц следует учитывать определенные правила:

– таблица должна быть небольшой по размерам; иметь точно, кратко и ясно сформулированные название, заголовки строк подлежащего и граф сказуемого;

– в таблице должны быть указаны название территории и период, к которому относятся приводимые данные, границы интервалов, единицы измерения и т.д.;

– таблица должна содержать обязательно необходимые итоги (групповые, общие, проверочные);

– при большом количестве строк в подлежащем и граф в сказуемом их можно нумеровать порядковыми номерами. При этом в сказуемом нумеруются только графы, отводимые для вписывания цифр; графы для обозначения подлежащего и единиц его измерения обозначаются прописными буквами «А» и «Б»;

– округление чисел необходимо производить с одинаковой степенью точности для всей графы однородных показателей.

Для наглядности изображения статистических данных применяют графики.

Статистическими графиками называются условные изображения статистических величин и их соотношений в форме различных геометрических образов. Они применяются для характеристики развития явления во времени, отображения структуры явления и структурных сдвигов, для изучения взаимосвязи между явлениями.

Статистические *графики классифицируются:*

По содержанию выделяют графики сравнения в пространстве, во времени, размещения по территориям. Возможны их комбинации.

По способу построения – диаграммы, картограммы и картодиаграммы.

Диаграммы – изображение статистических данных при помощи геометрических фигур, линий, точек.

Картограммы – географическая карта, которая графически характеризует пространственное распределение какого – либо статистического показателя путем различной окраски, штриховки и т.п.

Картодиаграмма – совмещение картограммы с диаграммой, т.е. в отдельных районах условными знаками наносят абсолютные значения статистических показателей.

По характеру геометрического образа – точечные, линейные, плоские, объемные.

Статистические ряды распределения

В результате сводки статистических материалов получают ряды статистических данных, характеризующих либо изменение объемов совокупности в динамике (статистические ряды динамики), либо распределение единиц совокупности по тем или иным варьирующим признакам в статистике (статистические ряды распределения).

Статистический ряд распределения – это упорядоченное распределение единиц совокупности на группы по определенному признаку.

Он характеризует состав (структуру) изучаемого явления, позволяет судить об однородности совокупности, закономерности распределения и границах варьирования единиц совокупности.

В зависимости *от признака положенного в основу ряда распределения* различают атрибутивные и вариационные ряды распределения.

Атрибутивные ряды распределения – это ряды распределения, построенные по качественному признаку, т.е. признаку, не имеющему числового значения, например: **распределение студентов по половому признаку, национальности и т.д.**

Вариационные ряды распределения – это ряды распределения, построенные по количественному признаку, **например: распределение населения по заработной плате, числу детей в семье и т.д.**

Любой вариационный ряд состоит из двух элементов:

- **варианты –** отдельных значений признака, которые он принимает в вариационном ряду, т.е. конкретное значение варьирующего признака.
- **частоты –** численности отдельных вариантов или каждой группы вариационного ряда, т.е. это числа, которые показывают, как часто встречаются те или иные варианты в ряду распределения. Сумма всех частот определяет численность всей совокупности, ее объем.

Частоты, выраженные в долях единицы или в процентах к итогу называются частостями. Соответственно сумма частостей равна 1, или 100%.

В зависимости *от характера вариации признака* различают дискретные и интервальные ряды распределения.

Дискретный вариационный ряд характеризует распределение единиц совокупности по дискретному признаку, принимающему только целые значения.

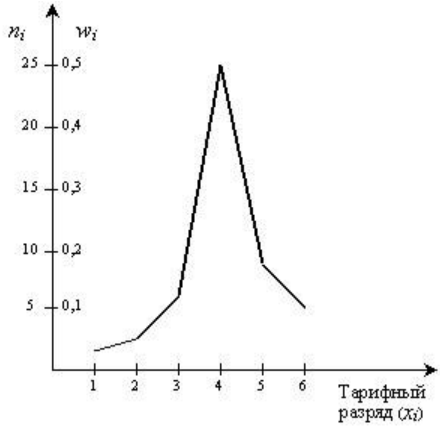
Анализ рядов распределения можно наглядно проводить на основе их графического изображения. Для изображения дискретных вариационных рядов используют полигон. По оси абсцисс – в одинаковом масштабе откладываются расшифрованные значения варьирующего признака; по оси ординат – наносится шкала для величины частот.

Для изображения интервального вариационного ряда применяется гистограмма. На оси абсцисс – в одинаковом масштабе откладываются величины интервала, а частоты отражаются прямоугольниками, построенными на соответствующих интервалах. Высота столбиков должна быть пропорциональна частотам. Если середины верхних строк соединить прямыми, то гистограмма может быть преобразована в полигон распределения.

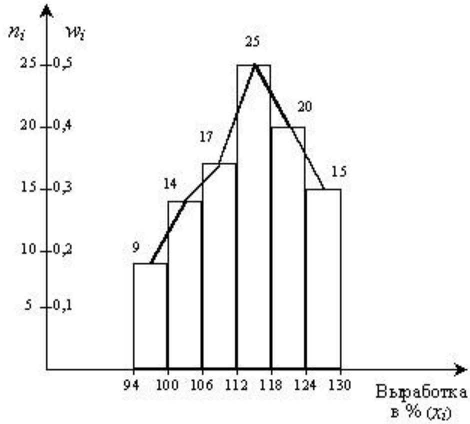
Для изображения вариационных рядов используется *кумулятивная кривая* (кумулята).

Пример 4. Дано распределение рабочих цеха по тарифному разряду. Построить графическое отображение вариационного ряда.

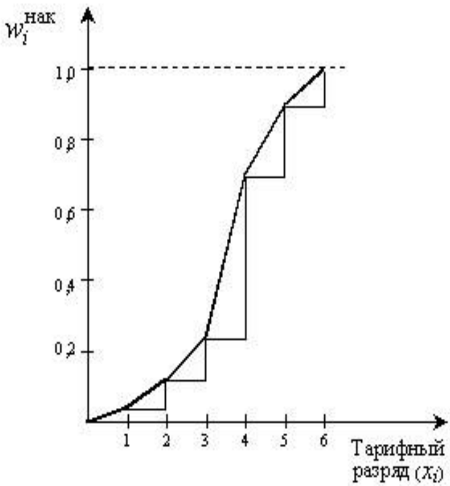
Тарифный разряд, x_i	1	2	3	4	5	6	Сумма
Количество рабочих (частота), n_i	2	3	6	25	9	5	50
Частость, $w_i = n_i/n$	0,04	0,06	0,12	0,5	0,18	0,1	1



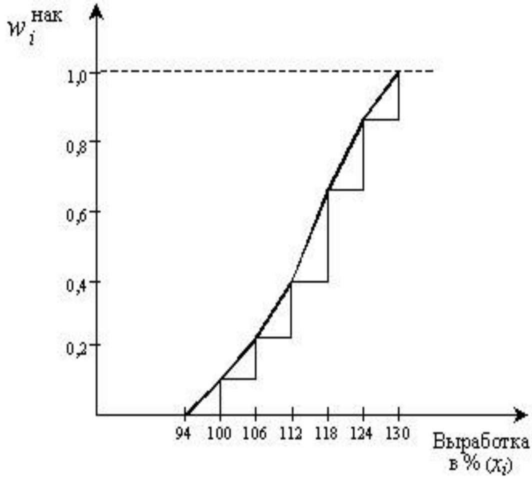
а) Дискретный вариационный ряд,
(полигон)



б) Интервальный вариационный ряд,
(гистограмма, полигон)



а) Дискретный вариационный ряд,
(кумулята)



б) Интервальный вариационный ряд,
(кумулята)

Рис. 1. Графическое отображение вариационных рядов

2 РАЗДЕЛ. АНАЛИТИЧЕСКАЯ СТАТИСТИКА

4 АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

Абсолютные статистические показатели

Статистический показатель – это количественная характеристика изучаемого объекта или свойства. Формами выражения статистических показателей служат абсолютные и относительные величины.

В теории статистики понятие «статистический показатель» характеризует качественную характеристику признака, количественную характеристику качественно определенного социально-экономического явления и количественную оценку свойств изучаемого явления.

Результаты статистического наблюдения регистрируются, прежде всего в форме первичных абсолютных величин.

Абсолютная величина – именованное число, имеющее определенную размерность (индивидуальную или общие) и единицу измерения (натуральную, условно-натуральную, стоимостную, трудовую), которая отражает технические или потребительские свойства продукта.

Индивидуальные абсолютные показатели характеризуют отдельный объект или единицу совокупности. Их получают непосредственно в процессе статистического наблюдения: первичного учета, переписных листах и других бланках обследования. На их основе группируются первичные материалы статистического наблюдения и строятся вариационные ряды.

Общие абсолютные показатели получаются в результате суммирования значений индивидуальных абсолютных величин. Они характеризуют объем совокупности или объем признака всего явления или отдельных его частей и являются результатом сводки и группировки первичных статистических данных.

Абсолютные показатели измеряются в различных единицах.

Натуральные единицы измерения бывают простыми, составными и условными.

Простые натуральные единицы измерения — это штуки, километры, килограммы, тонны, метры, литры, мили, дюймы и т. д. В простых натуральных единицах также измеряется объем статистической совокупности или объем отдельной ее части (количество предприятий, из них количество малых предприятий и т.д.).

Составные натуральные единицы измерения имеют расчетные показатели, получаемые как произведение двух или нескольких показателей, имеющих простые единицы измерения, например: грузооборот — в тонно-километрах (масса перевезенных грузов умножается на расстояние перевозки).

Условные натуральные единицы измерения широко используют в анализе производственной деятельности, когда требуется найти итоговое значение (сумму) однотипных показателей, которые напрямую несопоставимы, но характеризуют одни и те же свойства объектов.

Перевод в условные единицы измерения осуществляется на основе специальных коэффициентов, рассчитываемых как отношение потребительских свойств отдельных разновидностей продукта к эталонному значению. Например, при производстве консервов их общий объем пересчитывается в условные консервные банки объемом 353,4 $см^3$; мыла — в условное мыло с 40%-ным содержанием жирных кислот и т. д.

Стоимостные единицы измерения: рубли, доллары, евро, валюта других стран. Они позволяют суммировать либо сравнивать показатели, которые не сопоставимы в натуральных единицах измерения, например, определить общий объем производства различных видов продукции, общий объем всех затрат, связанных с производством продукции. При этом можно сравнивать только сопоставимые во временном аспекте показатели, с учетом цен и инфляции, например, размер заработной платы в 2005 и 2013 г. Для сравнения необходимо использовать корректировочный коэффициент, учитывающий инфляцию.

Стоимостные единицы измерения наибольшее распространение получили при анализе социально-экономических явлений.

Трудовые единицы измерения: человеко-дни (число работников предприятия умножается на количество отработанных за период дней) или человеко-часы (число работников предприятия умножается на среднюю продолжительность одного рабочего дня и количество рабочих дней в периоде), которые используются для учета затрат труда на предприятиях.

Пример 1. За отчетный период сельскохозяйственным предприятием было получено молока.

Таблица 1

Степень жирности молока, %	Количество, литров
1	3000
2	2400
3	1600
4	800

Определить общее количество произведенного молока в условно-натуральных единицах измерения. За условную единицу измерения принять молоко 2 % жирности.

Решение.

Для определения общего количества молока, необходимо исчислить коэффициент перевода. Коэффициент перевода в условное молоко 2 % -й жирности:

молоко 1 % – жирности: 1/2 = 0,5;

молоко 3 % – жирности: 3/2=1,5;

молоко 4 % – жирности: 4/2= 2.

Определим количество молока в условно натуральных единицах измерения (табл. 2).

Таблица 2

Общий объем перевода молока

Степень жирности молока, %	Количество, литров	Коэффициент перевода	Количество молока в условно-натуральных единицах измерения, литров
1	3000	0,5	1500
2	2400	1	2400
3	1600	1,5	2400
4	800	2	1600
Итого	-	-	7900

Т.о. общий объем произведенного молока в 2 %-м исчислении составит 7900 литров.

Относительные статистические показатели

Для обобщающей характеристики массовых социально-экономических явлений наряду с абсолютными величинами используют относительные показатели.

Относительная величина – это результат деления двух абсолютных величин, характеризующий их количественное соотношение.

Числитель – сравниваемая (текущая) величина.

Знаменатель – база (основание) сравнения. Если основание принимается за единицу, то относительная величина выражается в виде коэффициента и показывает, во сколько раз, сравниваемый относительный показатель больше базисного или какую долю от базисной она составляет.

Если база приравнена к 100, то относительная величина выражается в процентах (%), если к 1000 – в промилле (‰), т. е. величина в знаменателе принимается за 1000 единиц, если к 10000 – в продецимилле.

Выбор формы относительной величины зависит от ее абсолютного значения. Если сравниваемая величина больше базы сравнения в 2 раза и более, то выбирают форму коэффициента. Если относительная величина близка к единице, как правило, ее выражают в процентах.

В зависимости от целей статистического анализа различают следующие виды показателей в форме относительных величин:

- планового задания;
- выполнения плана;
- динамики;
- структуры;
- координации;
- сравнения интенсивности.

Относительная величина планового задания (ОВПЗ)– отношение уровня, запланированного на предстоящий период ($X_{пл}$) к уровню показателя, достигнутого в предыдущем периоде (X_0):

$$ОВПЗ = \frac{X_{пл}}{X_0}$$

ОВПЗ характеризует изменение исследуемого явления в плановом периоде по отношению к уровню, достигнутому в предыдущем периоде.

Относительная величина выполнения плана (ОВВП)– отношение фактически достигнутого уровня в текущем периоде ($X_ф$) к уровню планируемого показателя на этот же период ($X_{пл}$):

$$ОВВП = \frac{X_ф}{X_{пл}}$$

На основе рассчитанного ОПВП судят о степени выполнения плана в текущем периоде.

В теории статистики для расчета относительного показателя реализации плана необходимы следующие данные: показатели отчетного периода и показатели плана.

Относительная величина динамики (ОВД) характеризует изменение уровня развития какого-либо явления во времени. Она получается путем деления текущего уровня признака на предшествующий или базисный уровень признака и показывает во сколько раз изменился уровень явления, или какую долю от базисного уровня он составляет.

ОВД может быть выражена в коэффициентах или процентах и рассчитана с использованием переменной базы сравнения – цепная и постоянной базы сравнения – базисная:

базисная:

$$ОВД^{баз} = \frac{X_i}{X_0}.$$

цепная:

$$ОВД^{цепн} = \frac{X_i}{X_{i-1}}.$$

Относительная величина динамики называется темпами роста.

Между цепными и базисными относительными величинами динамики существует взаимосвязь: произведение всех цепных ОПД (взятых в виде коэффициентов) дает базисный ОПД последнего периода.

Между относительными величинами: **планового задания, выполнения плана и динамики** существует взаимосвязь:

$$ОВД = ОВПЗ * ОВВП.$$

Уравнение выполняется, если используемые в нем показатели измерены в виде простого кратного отношения.

Пример 1. Определить относительную величину планового задания и выполнения плана, если выручка торговой фирмы во втором квартале 12 млн р. На третий квартал выручка планировалась в размере 15 млн р., а фактически получено 16 млн р.

Решение

Относительная величина планового задания:

$$ОВПЗ = \frac{X_{пл}}{X_0} = \frac{15}{12} = 1{,}25$$ или 125,00 %

Относительная величина выполнения плана:

$$ОВВП = \frac{X_ф}{X_{пл}} = \frac{16}{15} = 1{,}0667$$ или 106,67 %.

$$ОВД = \frac{X_ф}{X_0} = \frac{16}{12} = 1{,}3333$$ или 133,33%

Следовательно, в текущем периоде по плану предполагалось увеличение выручки в 1,25 раза или на 25%. Фактически план был перевыполнен на 6,67%. По сравнению со вторым кварталом выручка выросла в 1,3333 раза, или на 33,33%

Проверим взаимосвязь:

$$ОВД = ОВПЗ * ОВВП$$

1,3333= 1,25 *1,0667.

Относительная величина структуры (ОВС) – отношение части единиц совокупности (f_i) ко всему объему совокупности ($\sum f_i$):

$$ОВС = \frac{f_i}{\sum f_i}.$$

ОВС характеризует долю (удельный вес) составных частей целого в ее общем итоге и обычно выражается в виде коэффициента (доли единиц) или процента.

Пример 3. Доходы от услуг связи в 2010 году составили 1355549,9 млн руб. объем реализации услуг населению 750523,6 млн руб.

Рассчитать относительную величину структуры.

Решение

Относительная величина структуры:

$$OBC_{CTP.} = \frac{f_i}{\sum f_i} = \frac{75052,36}{135554,99} = 0,5537$$ или 55,37 %.

Доля объема реализации услуг населению в общем объеме доходов составила 55,37 %.

Относительная величина координации – соотношение численности каждых групп, входящих в совокупность:

$$OBK = \frac{f_i}{f_g}$$

Они характеризуют соотношение частей изучаемой статистической совокупности, которое показывает, во сколько раз сравниваемая часть является больше или меньше части, принимаемой за основание (базу) сравнения или сколько единиц одной части приходится на 1, 10, 100, 1000 единиц другой части.

Пример 4. Доходы в 2010 году от местной телефонной связи в городской местности составили 145854,1 млн руб., а от местной телефонной связи в сельской местности 12328,5 млн руб.

Рассчитать относительные величины координации, приняв за основу доходы от местной телефонной связи.

ОВС координации:

$$I_K = \frac{f_i}{f_j} = \frac{145854,1}{12328,5} = 11,83$$

Относительная величина сравнения:

$$OBCP = \frac{X_A}{X_B}.$$

ОВСР характеризует отношение одноименных абсолютных показателей, соответствующих одному и тому же периоду или моменту времени, но к различным объектам или территориям. При этом сравниваемые величины должны иметь одну и ту же методологию расчета. Выражается в коэффициентах или в процентах.

Пример 5. Средний размер ссуды выданной жителям Ростовской области составил 432 тыс р., а Краснодарского края – 507 тыс.р.

Рассчитать относительные величины сравнения среднего размера ссуды.

Решение

Относительная величина сравнения среднего размера ссуды выданной жителям Ростовской области со средним размером ссуды выданной жителям Краснодарского края:

$$OBC = \frac{X_A}{X_B} = \frac{432}{507} = 0{,}85 .$$

Таким образом, средний размер выданной жителям Ростовской области на 15% меньше аналогичного показателя в Краснодарском крае.

Относительная величина сравнения среднего размера ссуды, выданной жителям Краснодарского края с размером ссуды выданной жителям Ростовской области:

$$OBC_C = \frac{X_A}{X_B} = \frac{507}{432} = 1{,}17.$$

Таким образом, средний размер ссуды, выданной жителям Краснодарского края в 1,17 раза выше аналогичного показателя в Ростовской области.

Относительные величины интенсивности – сопоставление разноименных признаков одной совокупности, а также объектов двух связанных между собой совокупностей, например, демографические коэффициенты рождаемости, смертности, естественного прироста, брачности и др.

Так, коэффициент рождаемости рассчитывается как отношение числа родившихся за год к среднегодовой численности населения. Выражается в промилле и характеризует уровень явления (число родившихся) в расчете на 1000 жителей.

Является именованным показателем и может выражаться в кратных отношениях, процентах, промилле и других формах.

Относительные величины интенсивности позволяют выяснить, как часто встречается то или иное явление, какова степень его распространенности и насыщенности.

Относительные величины интенсивности могут быть и именованными числами. Например, плотность населения имеет единицу измерения чел./км^2.

При сопоставлении абсолютных величин и рассчитанных на их основе относительных величин необходимо наличие одинаковой: методологии расчета, круга объектов, единиц измерения периодов времени или даты.

5 СРЕДНИЕ ВЕЛИЧИНЫ

Наиболее распространенной формой статистических показателей является средняя величина.

Средняя величина – это обобщающий показатель, характеризующий типичный уровень варьирующего количественного признака в конкретных условиях, месте и времени.

Она показывает уровень признака в расчете на единицу совокупности. С помощью средних величин проводится сравнение различных совокупностей по варьирующим признакам, изучаются закономерности развития явлений и процессов общественной жизни.

Для характеристики любой совокупности, описания ее типических черт и качественных особенностей используется система средних показателей.

Выбор средней определяется экономическим содержанием определяемого показателя и исходных данных.

Виды средних величин:

1) **степенные средние:** арифметическая, гармоническая, квадратическая, геометрическая и др.

2) **структурные средние:** мода, медиана, квартили, децили и др.

Элементы степенной средней:

– варианта (x) – признак, для которого исчисляется средняя величина;

– число единиц (n) –количество вариант в исследуемой совокупности;

– веса, частоты (f) – показатели повторяемости вариант в исследуемой совокупности.

Главным условием, при котором можно использовать степенные средние в статистическом анализе, является *однородность совокупности*, которая не должна содержать исходных данных, резко различающихся по своему количественному значению (в литературе они носят название аномальных наблюдений). Для того, чтобы средняя не теряла бы своего смысла при ее расчете необходимо резко выделяющиеся наблюдения либо исключить из анализа и тем самым сделать совокупность однородной, либо разбить совокупность на однородные группы и вычислить средние значения по каждой группе и анализировать не общую среднюю, а групповые средние значения.

В зависимости от того, как образуется общий объем варьирующего признака, выбирают ту или иную среднюю:

Если объем признака базируется как сумма вариантов, то используется средняя арифметическая, если как сумма обратных значений признака, то – средняя гармоническая, если как произведение вариантов, то – средняя геометрическая.

Степенные средние

►**Средняя арифметическая** $(\bar{x})$ **простая** – равна сумме отдельных значений признака (x), деленной на общее число этих признаков (n):

$$\bar{x} = \frac{\sum x}{n}$$

Средняя арифметическая простая применяется в тех случаях, когда варианты представлены индивидуально в виде перечня в любом порядке.

►**Средняя арифметическая** $(\bar{x})$ **взвешенная:**

$$\bar{x} = \frac{\sum x * f}{\sum f}.$$

Средняя арифметическая $(\bar{x})$ взвешенная применяется в случае неоднократного повторения вариантов.

В качестве весов могут быть использованы относительные величины, выраженные в процентах (d):

$$\bar{x} = \frac{\sum x \times d}{\sum d}$$

Если относительные величины выражены коэффициентами, то

$$\sum d = 1, \rightarrow \bar{x} = \sum xd.$$

Если результаты наблюдения представляют в *виде интервального ряда распределения, т*огда при расчете средней в качестве x_i берут середины интервалов. Если первый и последний интервалы открыты (не имеют одной из границ), то их условно «закрывают», принимая за величины данного интервала величину примыкающего интервала, т. е. первый закрывают исходя из величины второго, а последний — по величине предпоследнего.

Средняя арифметическая величина обладает рядом математических свойств:

1. если все индивидуальные значения признака уменьшить или увеличить в n раз, то величина средней арифметической также уменьшится или увеличится в n раз;
2. если из всех значений признака вычесть постоянную величину (А), то средняя арифметическая уменьшится на эту величину (А).
3. если уменьшить или увеличить частоту (f_i) каждого значения признака в *m* раз, то величина средней арифметической не изменится;
4. средняя величина, умноженная на численность всей совокупности, будет равна сумме произведений каждого варианта на его численность;
5. сумма отклонений значений признака от его средней арифметической равна нулю.

На свойствах средней арифметической базируется один из методов ее расчета – способ моментов, или метод отсчета от условного нуля, который используется в случае вариационных рядов, с равными интервала. Согласно этому методу среднюю арифметическую взвешенную можно вычислить по формуле:

$$\bar{x} = i * m_1 + A;$$

$$m_1 = \frac{\sum(\frac{X_i - A}{i}) * f_i}{\sum f_i};$$

где m_1 – величина момента первого порядка;

i – величина интервала;

А – значение середины интервала, находящегося в центре ряда (если количество интервалов нечетное), или середина интервала с наибольшей частотой также из центра ряда (при четном количестве интервалов в центре ряда будут находиться два интервала).

Пример 2. Рассчитаем средний размер складских помещений по площади способом моментов.

Группы складских помещений по площади, тыс. м^2	Середина интервала (x_i)	Количество (f_i)	$(x_i - A)$ A=17,5	$(\frac{x_i - A}{i})$	$(\frac{x_i - A}{i} f_i)$
5-10	7,5	12	-10	-2	-24
10-15	12,5	21	-5	-1	-21
15-20	17,5	17	0	0	0
20-25	22,5	9	5	1	9
25 и более	27,5	7	10	2	14
Сумма	-	66	-	-	-22

$$m_1 = \frac{\sum(\frac{X_i - A}{i}) * f_i}{\sum f_i} = -0,333 \text{ тыс.м}^2$$

$$\bar{x} = i * m_1 + A = 5*(-0,333) + 17,5 = 15,84 \text{ тыс.м}^2$$

Для того чтобы средняя не теряла бы своего смысла, при расчете средней следует аномальные, резко выделяющиеся наблюдения либо исключить из анализа и тем самым сделать совокупность однородной, либо разбить совокупность на однородные группы и вычислить средние значения по каждой группе и анализировать не общую среднюю, а групповые средние значения.

Модифицированной формой средней арифметической является с*редняя гармоническая* величина, которая может в простой и взвешенной форме.

►**Средняя гармоническая простая:**

$$\bar{x} = \frac{n}{\sum \frac{1}{x}}$$

►**Средняя гармоническая взвешенная:**

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

Средняя гармоническая $(\bar{x}_{гарм})$ применяется, когда статистическая информация представлена как произведение частот по отдельным вариантам совокупности.

Пример 1. Имеются следующие данные о заработанной плате рабочих по цехам завода:

№ цеха	Март		Апрель	
	Численность рабочих, чел.	Средняя месячная зарплата, р.	Средняя месячная зарплата, р.	Фонд зарплаты, р.
1	25	6200	6400	153600
2	17	4950	5300	95400
3	19	5340	5620	112400

Определить среднюю месячную заработную плата работников завода.

Решение

Средняя месячная заработная плата за март (по средней арифметической взвешенной):

$$\bar{З}п_0 = \frac{\sum Зп_0 C_0}{\sum C_0} = \frac{6200 \cdot 25 + 4950 \cdot 17 + 5340 \cdot 19}{25 + 17 + 19} = \frac{340610}{61} = 5583{,}77 \text{ руб.}$$

Средняя месячная заработная плата за апрель (по средней гармонической взвешенной):

$$\bar{Зп}_1 = \frac{\sum ФЗП_1}{\sum \frac{ФЗП_1}{Зп_1}} = \frac{153600+95400+112400}{\frac{153600}{6400}+\frac{95400}{5300}+\frac{112400}{5620}} = 5829{,}03 \text{ руб.}$$

►**Средняя геометрическая** $(\bar{x}_{геом})$ исчисляется извлечением корня степени m из произведений цепных темпов роста.

$$\bar{T}_p = \sqrt[m]{k_{p2/1} * k_{p3/2} * \ldots * k_{n/n-1}} = \sqrt[m]{\prod k_{pi/i-1}}$$

Средняя геометрическая применяется в тех случаях, когда индивидуальные значения признака представляют собой относительные величины динамики, построенные в виде цепных величин. Она используется для анализа динамики явлений и позволяет определить средний коэффициент роста.

При расчете по одним и тем же данным между числовыми значениями средних, исчисленных по разным формулам, всегда сохраняется неравенство, которое отражено в правиле мажорантности средних:

$$\bar{x}_{гарм} \le \bar{x}_{геом} \le \bar{x}_{арифм} \le \bar{x}_{квадр}.$$

Оно представляется показателем степени; чем больше показатель степени в формуле средней , тем больше его величина.

►**Средняя хронологическая:**

$$\bar{x} = \frac{\frac{x_1}{2} + x_2 + \ldots + x_{n-1} + \frac{x_n}{2}}{n-1}.$$

Средняя хронологическая предназначена для определения средней величины моментного ряда.

Пример 3. Остатки запасов на складе, тыс. руб., на:

1.01 –120,

1.04 –150,

1.07 –138,
1.10 – 96,
1.01 следующего года – 145.
Определить средний размер запасов.

$$\bar{x} = \frac{\frac{120}{2} + 150 + 138 + 97 + \frac{110}{2}}{4} = 125 тыс.руб.$$

Структурные средние

Разновидностью средней являются мода и медиана. Эти величины также используются в качестве характеристик вариационного ряда.

Мода (Мо) – варианта, встречающаяся в ряду распределения чаще всего, т.е. варианта, которой соответствует наибольшая частота.

Мода – значение признака, которое наиболее часто встречается в изучаемой совокупности.

Для дискретных рядов распределения модой является вариант с наибольшей частотой.

Пример 4. Продолжительность телефонных разговора 10 человек в течение дня составила: 5, 19, **25**, 16, **25**, 18, 31, 9, 12, 28 минут.

Модальная величина телефонного разговора составила **25** минут.

В интервальном ряду распределения с равными интервалами наибольшая частота указывает не на модальную варианту, а на содержащий моду интервал. Вычисление моды производится по следующей формуле:

$$Mo = x_o + i\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)},$$

где x_0 – нижняя граница модального интервала;
i – величина модального интервала;
f_2 – частота модального интервала;
f_1 – частота интервала, предшествующего модальному;
f_3 – частота интервала, следующего за модальным.

Для расчета моды определяют модальный интервал по наибольшей частоте.

Медиана – это значение признака, которое находится в середине ранжированного ряда и делит этот ряд на две равные по численности части.

Для ее определения необходимо расположить в порядке возрастания или убывания все варианты. Серединная варианта и будет являться медианой.

Если ряд состоит из четного числа членов, то за медиану принимают среднюю арифметическую из двух серединных значений.

Расчет медианы для интервального ряда распределения производится по формуле:

$$Me = x_0 + i\frac{\frac{1}{2}\sum f - S_{Me-1}}{f_{Me}},$$

где x_0 – нижняя граница медианного интервала;

$\sum f$ – сумма частот;

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

f_{Me} – частота медианного интервала.

Для определения медианного интервала необходимо рассчитать сумму накопленных частот. Накопленные частоты показывают, сколько единиц совокупности имеют значение признака не больше, чем данное значение и исчисляются путем последовательного прибавления к частоте первого интервала частот последующих интервалов.

Для расчета медианы определяют медианный интервал путем определения порядкового номера медианы на основе подсчета суммы частот по итогам до числа, превышающего половину объема совокупности.

Пример 5. Имеются данные о распределении организаций города по размеру активов:

Группы организаций по размеру активов, тыс. руб.	Число организаций (f)	Накопленные частоты (S)
450 – 550	15	15
550 – 650	42	57
650 – 750	36	93
750 – 850	75	168
850 – 950	84	252 (**$S_{Ме-1}$**)
950 – 1050 (медианный интервал)	(**$f_{Ме}$**) 105 (**f_1**)	357(**$S_{Ме}$**)
1050 – 1150 (модальный интервал)	128 (**f_2**)	485
1150 – 1250	103 (**f_3**)	588

Определить модальный и медианный размер активов.

Решение

Для расчета моды определяют модальный интервал по наибольшей частоте. Наибольшую частоту (128) имеет интервал (1050 – 1150). Отсюда

$$Mo = x_o + i\frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)} = 1050 + 100\frac{128 - 105}{(128 - 105) + (128 - 103)} = 1097{,}92 \text{ тыс.р.}$$

Наиболее часто встречаются организации с величиной активов в размере 1097,92 тыс. р.

$$Me = x_0 + i\frac{\frac{1}{2}\sum f - S_{Me-1}}{f_{ме}} = 950 + 100\frac{294 - 252}{105} = 990 \text{ тыс.р.}$$

Из организаций половина имеет размер активов менее 990 тыс. р., а другая половина организаций – более.

Мода и медиана могут быть определены графически.

Мода определяется **по гистограмме распределения**. Для этого строят на оси абсцисс ряд сомкнутых прямоугольников, у каждого из которых основанием служит величина интервала признака, а высотой – частота каждого интервала. Правую вершину прямоугольника, имеюще- го наибольшую высоту, соединяют с правым верхним углом предыдущего прямоугольника, а левую вершину прямоугольника имеющего наибольшую высоту – с левым верхним углом последующего прямоугольника. Абсцисса точки пересечения этих прямых и будет модой.

Медиана определяется **по кумуляте**, которую строят по накопленным частотам. Для этого из верхней границы каждого интервала на оси абсцисс восстанавливают перпендикуляр, соответствующий по высоте накопленной частоте с начала ряда по данный интервал. Последовательное соединение вершин перпендикуляров позволяет получить кривую, называемую кумулятой.

Из точки на оси ординат, соответствующей половине всех частот (порядковому номеру медианы), проводят прямую, параллельную оси абсцисс, до пересечения ее с кумулятой, из которой опускают перпендикуляр на ось абсцисс. Абсцисса точки пересечения является медианой.

По аналогии с нахождением медианы в вариационных рядах можно найти значение признака у любой ранжированной в определенном порядке единицы.

6 ПОКАЗАТЕЛИ ВАРИАЦИИ

Вариацией признака называется наличие различий в численных значениях признаков у отдельных единиц совокупности.

Для характеристики размеров колеблемости признаков в статистике используют ряд показателей.

Абсолютные показатели вариации

Размах вариации показывает крайние отклонения признака – это разность между максимальным и минимальным значениями признака:

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

Размах вариации характеризует пределы изменения варьирующего признака.

Среднелинейное отклонение – средняя арифметическая величина из отклонений вариантов признака от их средней.

невзвешенное:

$$\bar{d} = \frac{\Sigma|x_i - \bar{x}|}{n}$$

взвешенное:

$$\bar{d} = \frac{\Sigma|x_i - \bar{x}| f_i}{\Sigma f_i}.$$

Дисперсия σ^2 – средний квадрат отклонений индивидуальных значений признака от их средней величины.

Дисперсия вариационного признака:

невзвешенная:

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

взвешенная:

$$\sigma^2 = \frac{\Sigma(x_i - \bar{x})^2 f_i}{\Sigma f_i}$$

Для определения колеблемости признака в совокупности применяется **среднее квадратическое отклонение,** которое представляет собой корень второй степени из среднего квадрата отклонений отдельных значений признака от их средней:

невзвешенное:

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

взвешенное:

$$\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 f_i}{\Sigma f_i}}$$

Среднее квадратическое отклонение – величина именованная, имеет размерность осредняемого признака и показывает, на сколько в среднем отклоняются конкретные варианты признака от его среднего значения.

В зарубежной практике среднее квадратическое отклонение называется стандартным отклонением и применяется в различных стандартах.

Относительные показатели вариации

Для оценки меры вариации и ее значимости используют коэффициент вариации:

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

Значение коэффициента вариации изменяется от 0 до 1, и чем ближе он к нулю, тем типичнее найденная средняя величина для изучаемой статистической совокупности. Он дает характеристику однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33%.

Чем больше его величина, тем больше разброс значений признака вокруг средней, тем менее однородна совокупность по составу.

Линейный коэффициент вариации:

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}}100 \%$$

Коэффициент осцилляции:

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

Коэффициент осцилляции отражает относительную колеблемость крайних значений признака вокруг средней.

Относительные показатели вариации используются для сравнения колеблимости:

1) различных признаков в одной и той же совокупности;
2) одного и того же признака в нескольких совокупностях.

Выяснение общего характера совокупности предполагает оценку не только ее однородности, но и симметричности, для чего применяют **коэффициент асимметрии Пирсона,** как отношение разности степенной средней величины и моды к среднему отклонению от средней величины.

Расчетное значение этого коэффициента может изменяться от –3 до +3, а критериальным значением служит 0.

Нулевое значение возможно при одинаковых значениях степенной средней и моды, что свидетельствует о *симметричном или нормальном* распределении величин в совокупности

При *положительном коэффициенте* асимметрии степенная средняя больше моды, что говорит о *правосторонней асимметрии*, а при *отрицательном коэффициенте*, когда степенная средняя меньше моды, имеется *левосторонняя асимметрия* статистической совокупности.

Пример 1. Имеются выборочные данные.

Таблица 1

Размер активов, тыс. руб.	Число организаций (f)	Середина интервала ($\bar{x}_i$)	$\bar{x}_i$ f	$\bar{x}_i - \bar{x}$	$(\bar{x}_i - \bar{x})^2$	$(\bar{x}_i - \bar{x})^2$ f
450 – 550	15	500	7500	-454	433116	6496740
550 – 650	42	600	25200	-354	125316	5263272
650 – 750	36	700	25200	-254	64516	2322576
750 – 850	75	800	60000	-154	23716	1778700
850 – 950	84	900	75600	-54	2916	244944
950 – 1050	105	1000	105000	46	2116	222180
1050 – 1150	128	1100	140800	146	21316	2728448
1150 –1250	95	1200	114000	246	60516	5749020
Итого	580	-	553300	-	733528	24805880

Определить:

1) средний размер активов;
2) размах вариации;
3) дисперсию;
4) среднее квадратическое отклонение;
5) коэффициент вариации.

Решение.

1. Средний размер активов:

$$\bar{x} = 553300/580 = 954 \text{ тыс.руб.}$$

2. Размах вариации:

R= 1250-450= 800 тыс.руб.

3. Дисперсия:

$$\sigma^2 = 24805880/580 = 42768.$$

4. Среднее квадратическое отклонение:

σ =√42768 = 206,8 тыс.руб.

5. Коэффициент вариации:

V= 206,8/954*100= 21,68 %.

Альтернативный признак – косвенный признак, имеющий две взаимоисключающие разновидности.

Альтернативные признаки принимают всего два значения:

1 – наличие признака;

0 – отсутствие признака.

Дисперсия альтернативного признака:

$$\sigma_p^2 = pq,$$

где p – доля единиц в совокупности, обладающих данным признаком;

q – доля единиц, не обладающих данным признаком.

Среднеквадратическое отклонение альтернативного признака:

$$\sigma_p = \sqrt{pq}.$$

Показатели дисперсии

Вся совокупность может быть разбита на однородные группы по одному признаку-фактору.

Вариация признака в целом по совокупности зависит от:

- вариации признака внутри каждой группы,
- вариации групповых средних, т.е. от межгрупповой вариации признака.

Для статистической совокупности, сгруппированной по изучаемому признаку, возможно вычисление трех видов дисперсий:

- общей,
- частных (внутригрупповых),
- межгрупповой.

Общая дисперсия – измеряет вариацию признака по всей совокупности от общей средней под влиянием всех факторов, обусловивших эту вариацию:

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

Межгрупповая дисперсия характеризует вариацию групповых средних от общей средней:

$$\delta^2 = \frac{\Sigma(\bar{x}_i-\bar{x})^2 f_i}{\Sigma f_i},$$

где $\bar{x}_i$ – групповые средние;

$\bar{x}$ – общая средняя.

Внутригрупповая (частная) дисперсия отражает случайную вариацию, вариацию признака в группах от групповой средней:

$$\sigma^2 = \frac{\Sigma(x-\bar{x}_i)^2 f_i}{\Sigma f_i}.$$

Средняя из внутригрупповых (частных) дисперсии:

$$\bar{\sigma}_i^{\,2} = \frac{\Sigma\sigma_i^2 n_i}{\Sigma n_i},$$

где $\sigma_i^{\ 2}$ – групповые дисперсии;

n_i – число в группах.

Между указанными видами дисперсий существует соотношение, которое называется правилом сложения дисперсий: общая дисперсия равна сумме средней из частных дисперсий и межгрупповой:

$$\sigma^2 = \bar{\sigma}_i^2 + \delta_x^2,$$

где σ^2 – общая дисперсия;

$\bar{\sigma}_i^2$ – средняя из внутригрупповых дисперсии;

δ_x^2 – межгрупповая дисперсия.

С помощью правила сложения дисперсий можно измерить силу влияния факторного признака, который положен в основу группировки, на результативный признак, вычислив коэффициенты детерминации и эмпирическое корреляционное отношение.

Эмпирический коэффициент детерминации показывает долю общей вариации результативного признака, обусловленную вариацией группировочного признака, равен отношению межгрупповой дисперсии к общей:

$$\eta^2 = \frac{\delta_x^{\ 2}}{\sigma^2}.$$

Эмпирическое корреляционное отношение показывает тесноту связи между группировочным и результативным признаками:

$$\eta = \sqrt{\frac{\delta_x^{\ 2}}{\sigma^2}}.$$

Эмпирическое корреляционное отношение варьирует в пределах от 0 до 1. При $\eta = 0$ связи нет, т.е. группировочный признак не оказывает влияния на результативный. При $\eta = 1$ – связь полная, т.е. изменение результативного признака полностью обусловлено группировочным признаком. Чем больше корреляционное отношение приближается к единице, тем полнее корреляционная связь между признаками. (Связь при: 0-0,2 – очень слабая, 0,2-0,3 – слабая, 0,3-0,5 – умеренная, 0,5-0,7 – заметная, 0,7-0,9 – тесная, 0,9-0,99 – весьма тесная).

Правило сложения дисперсий для доли признака:

$$\sigma_{\bar{p}}^2 = \bar{\sigma}_{p_i}{}^2 + \delta_{p_i}^2,$$

где $\sigma_{\bar{p}}^2$ – общая дисперсия доли;

$\bar{\sigma}_{p_i}{}^2$ – средняя из внутригрупповых дисперсий доли;

$\delta_{p_i}^2$ – межгрупповая дисперсия доли.

Общая дисперсия доли:

$$\sigma_{\bar{p}}^2 = \bar{p}(1-\bar{p}),$$

где $\bar{p}$ – доля изучаемого признака во всей совокупности, определяемая по формуле:

$$\bar{p} = \frac{\Sigma p_i n_i}{\Sigma n_i}.$$

Средняя из групповых дисперсий доли:

$$\bar{\sigma}_{p_i}^2 = \bar{p}_i(\bar{1}-\bar{p}_i) = \frac{\Sigma p_i(1-p_i)n_i}{\Sigma n_i}.$$

Межгрупповая дисперсия доли:

$$\delta_{p_i}^2 = \frac{\Sigma(p_i-\bar{p})^2}{\Sigma n_i}.$$

Пример 2. В порядке выборки было определено количество бутонов на 10 кустах роз и установлены следующие данные:

Наименование сорта	Число проверенных кустов	Количество бутонов, шт				
		1	2	3	4	5
Восторг	2	10	12			
Вдохновение	4	12	11	12	13	
Восхищение	5	9	8	10	7	6

Определить:
1) среднее количество бутонов по каждому сорту и для всех сортов;
2) частные дисперсии;
3) общую дисперсию;
4) среднюю из частных дисперсий;
5) дисперсию групповых средних;
6) правило сложений дисперсий;
7) коэффициент детерминации;
8) эмпирическое корреляционное отношение.

Решение.

1. Среднее количество бутонов по каждому сорту:

Восторг: $$\bar{x}_1 = \frac{\sum x_i}{n} = \frac{10+12}{2} = 11$$ шт.

Вдохновение: $$\bar{x}_2 = \frac{\sum x_2}{n} = \frac{12+11+12+13}{4} = 12$$ шт.

Восхищение: $$\bar{x}_3 = \frac{\sum x_3}{n} = \frac{9+8+10+7+6}{5} = 8$$ шт.

Среднее количество бутонов для всех сортов:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{11*2+12*4+8*5}{11} = 10$$ шт.

2. Средний квадрат отклонений для каждого сорта (частные дисперсии):

$$\sigma_1^{\ 2} = \frac{\Sigma(x-\bar{x}_i)^2}{\Sigma n} = \frac{(10-11)^2+(12-11)^2}{2} = 1{,}0$$;

$$\sigma_2^{\ 2} = \frac{\Sigma(x-\bar{x}_2)^2}{\Sigma n_2} = \frac{(12-12)^2+(11-12)^2+(12-12)^2+(13-12)^2}{4} = 0{,}5\,;$$

$$\sigma_3{}^2 = \frac{\Sigma(x-\bar{x}_{3i})^2}{\Sigma n_3} = \frac{(9-8)^2+(8-8)^2+(10-8)^2+(7-8)^2+(6-8)^2}{5} = 2{,}0$$

3. Средний квадрат отклонений для всех сортов (общая дисперсия):

$$\sigma_1{}^2 = \frac{\Sigma(x_i-\bar{x})^2}{\Sigma n} = 4{,}72$$

.

4. Средняя из частных дисперсий:

$$\bar{\sigma}_i{}^2 = \frac{\Sigma\sigma_i^2 f_i}{\Sigma f_i} = \frac{1{,}0*2+0{,}5*4+2{,}0*5}{2+4+5} = 1{,}27.$$

5. Дисперсия групповых средних (межгрупповая):

$$\delta^2 = \frac{\Sigma(\bar{x}_i-\bar{x})^2 f_i}{\Sigma f_i} = \frac{(11-10)^2\cdot 2+(12-10)^2\cdot 4+(8-10)^2\cdot 5}{2+4+5} = 3{,}45.$$

6. Правило сложений дисперсий:

$$\sigma^2 = \bar{\sigma}_i{}^2 + \delta^2 = 1{,}27+3{,}45=4{,}72$$

7. Коэффициент детерминации:

$$\eta^2 = \frac{\delta_x{}^2}{\sigma^2} = \frac{3{,}45}{4{,}72} = 0{,}73$$

(73%).

Таким образом, на 73 % вариация количества бутонов обусловлена различиями между сортами, и только на 27 % – другими факторами.

8. Эмпирическое корреляционное отношение:

$$\eta = \sqrt{\frac{\delta_x{}^2}{\sigma^2}} = \sqrt{\frac{3{,}45}{4{,}72}} = 0{,}855.$$

Следовательно, можно утверждать, что связь значительная.

7 ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Выборочное наблюдение: понятие, задачи и ошибки выборки

Выборочное наблюдение или выборка – это вид несплошного наблюдения, при котором статистическому обследованию подвергаются не все единицы изучаемой совокупности, а лишь отобранные в определенном порядке.

Целью выборочного наблюдения является определение характеристик генеральной совокупности по характеристике выборочной совокупности.

Виды совокупности:

генеральная (N) – это совокупность, из которой производится отбор единиц совокупности;

выборочная совокупность (n) – это совокупность отобранных в определенном порядке единиц, по которым собирается информация.

С помощью выборки решают две основные задачи:

1) оценка средней генеральной совокупности,
2) оценка доли единиц генеральной совокупности, обладающих заданным признаком.

Сначала определяют среднюю ошибку выборки при оценке средней или доли генеральной совокупности, а затем устанавливают пределы изменения оцениваемого показателя в зависимости от избранного значения так называемой доверительной вероятности.

Основные характеристики параметров генеральной совокупности:

– генеральная средняя ($\overline{x}$) – среднее значение признака в генеральной совокупности:

$$\overline{x} = \frac{\sum x_i}{N}$$

– генеральная доля (р) – доля единиц, обладающих данным значением признака в общем числе единиц генеральной совокупности:

$$p = \frac{M}{N},$$

где М – численность единиц, обладающих определенным признаком в генеральной совокупности.

– дисперсия количественного признака в генеральной совокупности:

$$\sigma^2_{\bar{x}} = \frac{\sum (x_i - \bar{x})^2}{N}$$

– дисперсия доли признака в генеральной совокупности:

$$\sigma^2_p = p * (1 - p)$$

Основные характеристики параметров выборочной совокупности:

– выборочная средняя ($\tilde{x}$) – среднее значение признака в выборочной совокупности:

$$\tilde{x} = \frac{\sum x_i}{n}$$

– выборочная доля (ω) или частость – доля единиц, обладающих данным значением признака в выборочной совокупности:

$$\omega = \frac{m}{n},$$

где m – численность единиц, обладающих определенным признаком в выборочной совокупности.

– дисперсия количественного признака в выборочной совокупности:

$$\sigma^2_{\tilde{x}} = \frac{\sum (x_i - \tilde{x})^2}{n}$$

– дисперсия доли признака в выборочной совокупности:

$$\sigma_{\omega}^{2} = \omega * (1-\omega).$$

При выборочном наблюдении должна быть обеспечена случайность отбора единиц. Количество отобранных в выборочную совокупность единиц обычно определяется исходя из принятой доли выборки:

$$K_{в} = \frac{n}{N}.$$

Существуют различные методы, способы и виды отбора формирования выборочной совокупности.

Методы отбора:

повторный – общая численность единиц генеральной совокупности остается в процессе выборки неизменной, за счет возврата после обследования каждой единицы, отобранной в случайном порядке, в генеральную совокупность;

бесповторный – общая численность единиц генеральной совокупности остается в процессе выборки сокращается, за счет не возврата после обследования каждой единицы, отобранной в случайном порядке, в генеральную совокупность.

Виды отбора:

индивидуальный – в выборочную совокупность отбираются отдельные единицы совокупности;

групповой – в выборочную совокупность отбираются качественно однородные группы или серии изучаемых единиц;

комбинированный – предполагает сочетание первого и второго видов.

Способ отбора определяет конкретный механизм или процедуру выборки единиц из генеральной совокупности.

Собственно-случайный отбор – заключается во включении единиц в выборочную совокупность наудачу.

Механический отбор заключается в отборе единиц из генеральной совокупности, производимый в каком-либо механическом порядке (каждая десятая единица – 10% отбор).

При механическом отборе устанавливается расстояние меду отбираемыми единицами, т.е. *шаг отсчета* ($\frac{N}{n}$ – величина обратная доле выборки) и номер единицы, которая должна быть обследована первой – *начало отсчета*.

Механический отбор всегда бывает бесповторным и при его осуществлении применяют те же формулы, что и при собственно-случайном бесповторном отборе.

Серийный отбор – когда вместо случайного отбора единиц совокупности осуществляется отбор групп единиц (серий). Внутри отобранных серий обследованию подвергаются все единицы, т.е. применяется сплошное наблюдение.

Типический отбор – это отбор, когда генеральная совокупность делится на группы по какому-либо типическому признаку, а затем внутри каждой группы производится отбор единиц в выборочную совокупность.

Ошибки выборки

Т.к. наблюдению подвержена не вся совокупность, а только ее часть, поэтому при проведении выборочного наблюдения неизбежны некоторые ошибки.

Расхождения между выборочной характеристикой и характеристикой генеральной совокупности называются **ошибками репрезентативности**.

Ошибки выборки являются случайными величинами и могут принимать различные значения. Поэтому определяют среднюю из возможных ошибок – **среднюю ошибку выборки**.

Средняя ошибка выборки зависит от:

объема выборки (чем больше численность при прочих равных условиях, тем меньше величина средней ошибки выборки);

степени варьирования признака (чем меньше вариация признака, а следовательно, и дисперсия, тем меньше ошибка выборки и наоборот).

Средняя ошибка выборки для:

– количественного признака:

$$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2_{\tilde{x}}}{n}}$$

– альтернативного признака:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}}.$$

Определение средней ошибки выборки при разных способах отбора

Показатель	Повторный выбор	Бесповторный отбор
Собственно-случайный отбор		
Средняя ошибка выборки для средней	$\mu_{\tilde{x}} = \sqrt{\frac{\sigma^2_{\tilde{x}}}{n}}$	$\mu = \sqrt{\frac{\sigma^{2\,2}_{\tilde{x}}}{n}(1-\frac{n}{N})}$
Средняя ошибка выборки для доли	$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}}$	$\mu = \sqrt{\frac{\omega(1-\omega)}{n}(1-\frac{n}{N})}$
Серийный отбор		
Средняя ошибка выборки для средней	$\mu = \sqrt{\frac{\delta^2_{\tilde{ч}}}{r}}$	$\mu = \sqrt{\frac{\delta^2_{\tilde{ч}}}{r}(1-\frac{r}{R})}$
Типический отбор		
Средняя ошибка для средней	$\mu = \sqrt{\frac{\bar{\sigma}_i^2}{n}}$	$\mu = \sqrt{\frac{\bar{\sigma}_i^2}{n}(1-\frac{n}{N})}$

где δ^2 – межгрупповая (межсерийная) дисперсия;
r – число отобранных серий;
R – число серий в генеральной совокупности.
$\bar{\sigma}_i^2$ – средняя из внутригрупповых дисперсий;
σ_i^2 – дисперсия признака в *i*-й группе;
N_i – численность генеральной совокупности в *i*-й группе;
n_i – численность выборки в *i*-й группе.

Предельная ошибка выборки (Δ) – это максимально возможное расхождение выборочной средней и генеральной $(\tilde{x} - \bar{x})$, т.е. максимум ошибки при заданной вероятности (P).

Предельная ошибка выборки отвечает на вопрос о точности выборки с определенной вероятностью (P), значение которой определяется коэффициентом доверия, пропорциональности, кратностью ошибки (t).

Значения доверительной вероятности	Кратность ошибки
P = 0,683	t = 1
P = 0,954	t = 2
P = 0,997	t = 3

Предельная ошибка выборки для:
средней:

$$\Delta_{\tilde{x}} = t * \mu_{\tilde{x}}$$

доли:

$$\Delta_{\omega} = t * \mu_{\omega}.$$

Предельная ошибка выборки позволяет определить **предельные значения характеристик** генеральной совокупности при заданной вероятности и их доверительные интервалы.

Интервальная оценка для:
генеральной средней:

$$\bar{x} = \tilde{x} \pm \Delta_{\tilde{x}}, \quad \tilde{x} - \Delta_{\tilde{x}} \le \bar{x} \le \tilde{x} + \Delta_{\tilde{x}}$$

генеральной доли:

$$p = \omega \pm \Delta_{\omega}, \quad \omega - \Delta_{\omega} \le p \le \omega + \Delta_{\omega}.$$

Это означает, что с заданной вероятностью можно утверждать, что значение генеральной средней следует ожидать в пределах от $\tilde{x} - \Delta_{\tilde{x}}$ до $\tilde{x} + \Delta_{\tilde{x}}$ и от $\omega - \Delta_{\omega}$ до $\omega + \Delta_{\omega}$.

При проведении наблюдения допускаются ошибки: регистрации и репрезентативности.

При сплошном статистическом наблюдении могут возникать следующие виды ошибок: систематические и случайные ошибки регистрации.

При проектировании выборочного наблюдения с заранее заданным значением допустимой ошибки выборки очень важно правильно определить численность (объем) выборочной совокупности, которая с определенной вероятностью обеспечит заданную точность результатов наблюдения.

Определение численности выборки при разных способах отбора

Показатель	Повторный выбор	Бесповторный отбор
Собственно-случайный отбор		
Численность выборки при определении среднего размера признака	$n = \frac{t^2 \sigma_{\tilde{x}}^2}{\Delta_{\tilde{x}}^2}$	$n = \frac{t^2 \sigma_{\tilde{x}}^2 N}{\Delta_{\tilde{x}}^2 N + t^2 \sigma^2}$
Численность выборки при определении доли признака	$n = \frac{t^2 \omega(1-\omega)}{\Delta_{\omega}^2}$	$n = \frac{t^2 \omega(1-\omega) N}{\Delta_{\omega}^2 N + t^2 \omega(1-\omega)}$
Серийный отбор		
Численность выборки при определении среднего размера признака	$r = \frac{t^2 \delta_{\tilde{x}}{}^2}{\Delta_{\tilde{x}}{}^2}$	$r = \frac{t^2 \delta_{\tilde{x}}{}^2 R}{\Delta_{\tilde{x}}^2 R + t^2 \delta_{\tilde{x}}^2}$
Типический отбор		
Численность выборки при определении среднего размера признака	$n = \frac{t^2 \overline{\sigma}_i{}^2}{\Delta^2_{\tilde{x}}}$	$n = \frac{t^2 \overline{\sigma}_i{}^2 N}{\Delta^2_{\tilde{x}} N + t^2 \overline{\sigma}_i{}^2}$

Пример 1. На предприятии в порядке случайной бесповторной выборки было опрошено 50 рабочих из 500 и получены следующие данные об их доходах:

Месячный доход, р.	8000 – 8700	8700-9400	9400 – 10100	10100 – 10800	10800 – 11500
Число рабочих, чел.	4	6	11	22	7

Определить:

1) среднемесячный размер дохода рабочего предприятия, гарантируя результат с вероятностью 0,997;

2) долю рабочих предприятия, имеющих месячный доход 10100 руб. и выше, гарантируя результат с вероятностью 0,954;

3) необходимую численность выборки при определении среднего месячного дохода рабочих предприятия, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 100 р.;

4) необходимую числен- ность выборки при определе-

нии доли рабочих с размером месячного дохода 10100 р., и выше, чтобы с вероятностью 0,954 предельная ошибка не превышала 5%.

Решение

1. Доверительный интервал среднего размера месячного дохода работников предприятия:

$$\tilde{x} - \Delta\bar{x} \le \bar{x} \le \tilde{x} + \Delta\bar{x}.$$

Средний месячный доход по выборке:

$$\tilde{x} = \frac{\sum x_i f_i}{\sum f_i} =$$

$$= \frac{8350 \cdot 4 + 9050 \cdot 6 + 9750 \cdot 11 + 10450 \cdot 22 + 11150 \cdot 7}{50} = 10058{,}0 \text{ р.}$$

Предельная ошибка выборки $\Delta x = t\mu$. При вероятности P=0,997 t=3,0.

$$\Delta\bar{x} = t\mu = t\sqrt{\frac{\sigma^2_{\tilde{x}}}{n}(1 - \frac{n}{N})};$$

$$\sigma^2 = \frac{\sum (x_i - \tilde{x})^2 f}{\sum f} =$$

$$= \frac{(8350 - 10058)^2 \cdot 4 + (9050 - 10058)^2 \cdot 6 + (9750 - 10058)^2 \cdot 11 + (10450 - 10058)^2 \cdot 22 + (11150 - 10058)^2 \cdot 7}{4 + 6 + 11 + 22 + 7} =$$

$$= 610736$$

$$\Delta\bar{x} = 3 * \sqrt{\frac{610736}{50}(1 - \frac{50}{500})} = 314.5 \text{ р.}$$

9743,5 р. < $\bar{x}$ < 10372,5 р.

2. ω – доля рабочих, имеющих размер месячного дохода 10100 р. и выше

$$\omega = \frac{22+7}{50} = 0{,}58.$$

Предельная ошибка доли

$$\Delta p = t\mu_p = t\sqrt{\frac{\omega(1-\omega)}{n}(1-\frac{n}{N})}.$$

При вероятности p = 0,954, t = 2

$$\Delta p = 2\sqrt{\frac{0{,}58(1-0{,}58)}{50}\cdot(1-\frac{50}{500})} = 0{,}132$$ или 13,2%.

Доверительные интервалы для генеральной доли:

$$\omega - \Delta p \le p \le \omega + \Delta p;$$
$$0{,}58 - 0{,}132 \le p \le 0{,}58 + 0{,}132;$$
$$0{,}448 \le p \le 0{,}712.$$

3. Необходимую численность выборки для определения среднего месячного дохода определяется по формуле:

$$n = \frac{t^2 N\sigma^2}{\Delta^2_{\bar{x}} N + t^2\sigma^2}.$$

При вероятности p=0,954, t=2, $\Delta x = 100$ р.; $\sigma^2 = 610736$.

$$n = \frac{2^2\cdot 500\cdot 610736}{100^2\cdot 500 + 2^2\cdot 610736} = 164$$ чел.

4. Необходимая численность выборки для определения доли рабочих, имеющих доход 10100 р. и выше, определяется по формуле:

$$n = \frac{t^2 N\omega(1-\omega)}{\Delta_p^2 N + t^2\omega(1-\omega)}.$$

$\Delta p = 5$ %, или 0,05; при вероятности p = 0,954, t = 2; $\omega = 0{,}58$

$$n = \frac{2^2 \cdot 500 \cdot 0{,}58 \cdot (1-0{,}58)}{0{,}05^2 \cdot 500 + 2^2 \cdot 0{,}58 \cdot (1-0{,}58)} = 219 \text{ чел.}$$

Пример 2. При проверке качества деталей проведена 10 %- я серийная выборка из партии, содержащей 40 контейнеров деталей, методом механического отбора взято 4 контейнера. В результате сплошного обследования находящихся в контейнере деталей получили данные об удельном весе бракованных деталей:

Номер контейнера	1	2	3	4
Удельный вес деталей, %	0,8	1,0	0,9	1,1

Установить с вероятностью 0,995 доверительные интервалы удельного веса бракованных деталей для всей партии.

Решение

Для установления доверительного интервала, в котором для всей партии находится доля бракованной продукции, используется формула:

$$p = \omega \pm \Delta p;$$

$$\Delta p = t\mu_p = t\sqrt{\frac{\sigma_\omega^2}{r}\left(1-\frac{r}{R}\right)},$$

где σ_ω^2 – межсерийная выборочная дисперсия доли;
r – число контейнеров, попавших в выборку;
R – общее число контейнеров.

При вероятности p = 0,0995, t = 3,2

$$\omega = \frac{0{,}8+1{,}0+0{,}9+1{,}1}{4} = 0{,}95$$ %, или 0,0095;

$$\sigma_{\omega}^{2} = \frac{\sum(\omega_i - \omega)^2}{r} =$$

$$= \frac{(0{,}008-0{,}0095)^2+(0{,}01-0{,}0095)^2+(0{,}009-0{,}0095)^2+(0{,}011-0{,}0095)^2}{4} = 0{,}00000125$$

$$\Delta p = 3{,}2\sqrt{\frac{0{,}00000125}{4} * (1-\frac{4}{40})} = 0{,}0016$$, или 0,16 %

$$p = 0{,}95 \pm 0{,}16$$

$$0{,}79 \le p \le 1{,}1.$$

8 РЯДЫ ДИНАМИКИ И ИХ ПРИМЕНЕНИЕ В АНАЛИЗЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Ряды динамики: понятие, правила построения, показатели анализа и методы анализа основной тенденции развития в рядах динамики

Процесс развития социально-экономических явлений во времени в статистике называют динамикой. Для отображения динамики строят ряды динамики.

Ряд динамики – числовое значение статистического показателя, представленное во временной последовательности (т.е. расположенные в хронологическом порядке).

Каждый ряд динами имеет два основных элемента:

1. период времени (годы, кварталы, месяцы, сутки) или моменты (даты) времени;
2. уровень ряда (у).

Основная задача изучения рядов динамики – выявление тенденции (закономерности) в изменении уровней ряда, которая именуется трендом.

Ряды динамики различаются по признакам.

По форме представления уровни в динамическом ряду могут характеризоваться абсолютными, средними и относительными величинами.

По времени:

- моментные ряды, уровни которого характеризуют значение показателя по состоянию на определенные моменты времени (например, величина запасов на начало периода и т.д.);
- интервальные ряды, уровни которого характеризуют зна-

чение показателя, достигнутое за определенный период (интервал) (например, ряды показателей объема продаж по месяцам и т.д.).

Сумма уровней интервального ряда дает вполне реальный показатель- общий объем продаж за год и т.д.). Сумма уровней моментного ряда, хотя иногда подсчитывается, но реального содержания, как правило, не имеет.

По расстоянию между датами или интервалами времени:

►с равностоящими (по времени) уровнями;

►с неравностоящими (по времени) уровнями.

Ряды динамики, как правило, представляются в виде таблицы или графически. При графическом изображении ряда динамики на оси абсцисс строится шкала времени (t), а на оси ординат – шкала уровней ряда (y).

Показатели анализа рядов динамики

Показатели изменения уровней ряда

Для характеристики интенсивности изменения во времени используются показатели:

- абсолютный прирост;
- темпы роста;
- темпы прироста;
- абсолютное значение одного процента прироста.

Абсолютный прирост исчисляется для выражения абсолютной скорости роста (снижения) уровня ряда динамики, и рассчитывается как разность двух сравниваемых уровней и показывает, на сколько (в единицах измерения уровней ряда) уровень одного периода больше или меньше уровня какого либо предшествующего периода. Поэтому он может иметь знак «+» (при увеличении уровней) или «-» (при уменьшении уровней).

В зависимости от задачи исследования абсолютные приросты (снижения), темпы роста (снижения) и темпы прироста (снижения) могут быть рассчитаны с переменной базой сравнения (цепные) и постоянной базой сравнения (базисные).

Абсолютные приросты:

базисные:

$$\Delta y^{б} = y_i - y_0$$

цепные:

$$\Delta y^{ц} = y_i - y_{i-1},$$

где y_i – порядковый уровень ряда динамики;

y_0 – базисный уровень ряда динамики.

Относительные показатели изменения уровней ряда

Относительная величина динамики характеризует:_изменение уровня развития какого-либо явления во времени.

Интенсивность изменения уровней ряда динамики оценивается отношением последующего уровня к предыдущему или какому-либо другому, принятому за базу сравнения. Этот показатель называется **темп роста (Тр):**

базисный:

$$T_p = \frac{y_i}{y_0} \cdot 100 \quad \%;$$

цепной:

$$T_p = \frac{y_i}{y_{i-1}} \cdot 100 \quad \%,$$

Формы выражения темпов роста:

в коэффициентах (если база сравнения принимается за единицу) показывая, во сколько раз уровень данного периода больше базы сравнения или какую часть его составляет;

в процентах (если база сравнения принимается за 100), показывая сколько процентов составляет уровень данного периода по сравнению с уровнем базы сравнения.

Между цепными и базисными темпами роста существует взаимосвязь:

1) произведение цепных темпов роста равно базисному;

2) результат деления двух базисных коэффициентов равен цепному (промежуточному).

Для выражения изменения величины абсолютного прироста уровней ряда динамики в относительных величинах определяется **темп прироста (снижения) (Тпр),** который показывает, на сколько процентов данный уровень больше (или меньше) другого, принимаемого за базу сравнения.

Способы расчета показателя:

1 способ, как отношение абсолютного прироста к уровню, принятому за базу сравнения:

базисный:

$$T_{np} = \frac{\Delta y_i}{y_0}$$

цепной:

$$T_{пр} = \frac{\Delta y_i}{y_{i-1}}$$

2 способ, как разность между темпами роста и единицей, если темпы роста выражены в коэффициентах: $T_{пр} = T_р - 1$; или как разность между темпами роста и 100%, если темпы роста выражены в процентах: $T_{пр} = T_р - 100$ %.

Абсолютное значение одного процента прироста (снижения) – это отношение абсолютного цепного прироста к соответствующему цепному темпу прироста, выраженному в процентах:

$$A1\% = \frac{\Delta y}{T_{пр}}.$$

Этот показатель может быть рассчитан как одна сотая часть предыдущего уровня:

$$A1\% = 0{,}01 * Y_{i-1}.$$

Расчет этого показателя имеет смысл только на цепной основе.

Коэффициент опережения – это отношение темпов роста (или прироста) по двум динамическим рядам (в одинаковые отрезки времени).

Средние показатели динамических рядов

Обобщенной характеристикой динамического ряда служит средний уровень ряда. Методы расчета среднего уровня ряда зависят от его вида и способа получения статистических данных.

Для моментного ряда динамики расчет производится по средней хронологической:

для равноотстоящих уровней:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2}y_n}{n-1} = \frac{\frac{y_1 + y_n}{2} + \sum_{i=2}^{n-1} y_i}{n-1};$$

для неравноотстоящих уровней:

$$\bar{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \ldots + (y_{n-1} + y_n)t_{n-1}}{2(t_1 + t_2 + t_3 + \ldots + t_{n-1})} = \frac{\sum (y_i + y_{i+1})t}{2\sum_{t=1}^{n-1} t}.$$

где y_i, y_n – уровни ряда динамики;

t_i – длительность интервала времени между уровнями.

В интервальном ряду динамики:

с равноотстоящими уровнями, средний уровень рассчитывается как средняя арифметическая простая из уровней ряда:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n};$$

с неравноотстоящими уровнями:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i t_i}{\sum t_1}.$$

Средний абсолютный прирост производится по цепным абсолютным приростам:

$$\bar{\Delta}_y = \frac{\sum_{i=1}^{n-1} \Delta_{i/i-1}}{n-1}$$

или

$$\bar{\Delta}_y = \frac{\Delta_{i/1}}{n-1} = \frac{y_n - y_0}{n-1},$$

где y_n – конечный уровень ряда динамики;

n – число уровней ряда динамики.

Среднегодовой темп роста ($\bar{T}_p$ – средняя геометрическая в рядах динамики) вычисляется по:

цепным коэффициентам (темпам) роста:

$$\bar{T}_p = \sqrt[m]{k_{p2/1} * k_{p3/2} * \ldots * k_{n/n-1}} = \sqrt[m]{\prod k_{pi/i-1}},$$

где m – число темпов роста.

по абсолютным уровням ряда динамики:

$$\overline{T}_p = \sqrt[n-1]{\frac{y_n}{y_0}} .$$

Среднегодовой темп прироста получается вычитанием из среднего темпа роста 100%.

$$\overline{T}_{np} = \overline{T}_p - 100 .$$

Пример 1. Имеются данные о количестве денежных переводов отделения почтовой связи за 2006-2010 г.г., тыс. ед. (табл.1).

Таблица 1

2006 г.	2007 г.	2008 г.	2009 г.	2010 г.
632	640	660	690	705

Задание.

Определить:

1) цепные и базисные абсолютные приросты;
2) цепные и базисные темпы роста;
3) абсолютное значение одного процента прироста;
4) средний абсолютный прирост;
5) среднее значение одного процента прироста;
6) среднегодовые темпы роста и прироста.
7) нанести на график динамику ряда.

Полученные данные представить в таблице. Сделать вывод о характере изменения денежных переводов по годам.

Решение

Расчет показателей представим в таблице:

Год	Объем т.ед.	Абсолютные приросты, ед.		Тепы роста, %		Темпы прироста, %		А 1% прироста, т.ед.
		цепные	базисные	цепные	базисные	цепные	базисные	
2006	632	-	-	-	100,00	-	-	-
2007	640	8	8	101,27	101,27	1,27	1,27	0,632
2008	660	20	28	103,13	104,43	3,13	4,43	0,640
2009	690	30	58	104,55	109,18	4,55	9,18	0,660
2010	705	15	73	102,17	111,55	2,17	11,55	0,690

Средний абсолютный прирост:

а) как средняя арифметическая простая годовых (цепных) приростов:

$$\overline{\Delta}_y = \frac{\sum \Delta y}{n-1} = \frac{8+20+30+15}{4} = 18{,}25$$ тыс.ед.

б) $$\overline{\Delta}_y = \frac{y_n - y_0}{n-1} = \frac{705-632}{4} = 18{,}25$$ тыс.ед.

Среднее значение одного процента прироста:

$$\overline{A1\%} = \frac{\sum A1\%}{n} = \frac{0{,}632+0{,}640+0{,}660+0{,}690}{4} = 0{,}656$$ тыс.ед.

Среднегодовые темпы роста:
а) по цепным коэффициентам (темпам) роста:

$$\overline{T}_p = \sqrt[m]{k_{p2/1} \cdot k_{p3/2} \cdot \ldots \cdot k_{n/n-1}} = \sqrt[4]{1{,}0127 * 1{,}0313 * 1{,}0455 * 1{,}0217} = \sqrt[4]{1{,}1156} =$$
$$= 1{,}0277$$

б) по абсолютным уровням ряда динамики:

$$\overline{T}_p = \sqrt[n-1]{\frac{y_n}{y_0}} = \sqrt[4]{\frac{705}{632}} = \sqrt{1{,}1156} = 1{,}0277 .$$

Среднегодовые темпы прироста:

$$\overline{T}_{np} = \overline{T}_p - 100 = 102{,}77 - 100{,}00 = 2{,}77$$ %.

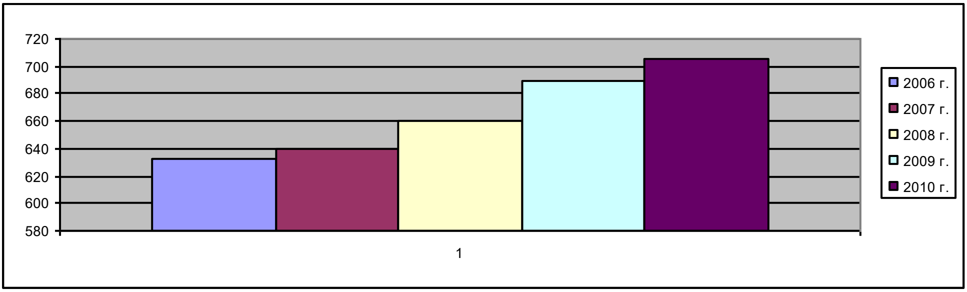


Рисунок 1 – Динамика количества денежных переводов

Приемы обработки и анализа рядов динамики

1. Приведение рядов к сопоставимому виду

Если уровни ряда, исчислены по разным методикам или находятся в неодинаковых границах, то такой ряд приводят к сопоставимому, используя метод смыкания рядов, т.е. осуществляется объединение двух и более рядов, характеризующих изменение явления в один ряд.

Существует несколько способов приведения рядов динамики к сопоставимому виду.

Первый. Абсолютный способ, предполагает расчет коэффициента перехода из старых границ в новые (или наоборот) или старой методики в новую, на который умножают прежние уровни.

Второй. Относительный способ, предполагает проведение сравнительной оценки по относительным уровням ряда. Показатели года, в котором происходили изменения (как до, так и после изменений) принимают за 100 %, а остальные пересчитываются относительно этих уровней соответственно. В результате получают сомкнутый ряд.

Пример 2. Имеются данные о реализации продукции в районе (табл. 2).

Таблица 2

Объем продажи, млн руб.	Год					
	2007	2008	2009	2010	2011	2012
В прежних границах	6,1	6,3	**6,5**	-	-	-
В новых границах	-	-	**7,8**	8,0	8,4	8,9

Привести ряды динамики к сопоставимому виду, применив прием смыкания рядов динамики.

Решение

Для приведения ряда динамики к сопоставимому виду используем два способа.

1. Определим коэффициент пересчета уровней 2009 года, в котором произошло изменение границ района:

$$K = \frac{7{,}8}{6{,}5} = 1{,}2$$

Умножим уровни ряда динамики в прежних границах на коэффициент пересчета, приведем их сопоставимому виду.

В 2007 г. :$y_{н.г.}$= 6,1*1,2=7,32 млн руб.

Получен сопоставимый ряд динамики реализации продукции в районе в новых границах, млн руб.:

Объем продажи, млн руб.	Год					
	2007	2008	2009	2010	2011	2012
Сомкнутый (сопоставимый) ряд абсолютных величин	7,32	7,56	7,8	8,0	8,4	8,9

2 способ. Уровни года, в котором произошли изменения – это объем продаж 2009 г. примем за 100,00%, а остальные (как до изменения, так и после изменения) пересчитаем относительно этих уровней соответственно: в старых границах по отношению к 6,5 млн руб., в новых границах по отношению к 7,8 млн руб. В результате получим сомкнутый ряд:

Объем продажи, млн руб.	Год					
	2007	2008	2009	2010	2011	2012
Сомкнутый (сопоставимый) ряд относительных величин, %	93,85	96,92	100,00	102,56	107,69	114,10

2. Выявление основной тенденции ряда динамики

Любой ряд динамики теоретически может быть представлен в виде следующих составляющих:

1. тренда – основной тенденции развития ряда, обуславливающей либо увеличение либо снижение его уровней;
2. периодических колебаний, в том числе сезонных;
3. случайных колебаний.

Изучение тренда включает два этапа:

1. проверку ряда на наличие тренда;
2. выравнивание ряда и непосредственное выделение тренда с экстраполяцией полученных результатов.

Важной задачей статистики при анализе рядов динамики является определение основной тенденции развития с применением различных способов.

Непосредственное выделение тренда можно осуществлять тремя методами:

1) укрупнения интервалов;
2) скользящей средней;
3) аналитического выравнивания.

1. Метод укрупнения интервалов.

Сущность метода заключается в укрупнении периодов времени, к которому относятся уровни ряда. Например, ряд ежедневной величины товарооборота заменяется рядом месячной величины товарооборота и т.д.

2. Метод скользящей средней.

Сущность метода заключается в расчет средних уровней динамического ряда по укрупненным интервалам путем последовательного смещения начала отсчета на один временной период, т.е. исключают из укрупненного интервала первые уровни и включают последующие.

Выбор интервала сглаживания (порядковой скользящей средней) обосновывается четностью/нечетностью ряда и результатом сглаживания (должна достигаться общая тенденция). $\frac{y_1 + y_2 + \ldots + y_n}{n}$; $\frac{y_2 + y_3 + \ldots + y_{n+1}}{n}$ и т.д. Например, для нечетного ряда, состоящего из 15 уровней, применяют в трех-, пяти- и семичленную скользящую среднюю (но не более половины ряда).

Для трехчленной скользящей средней:

$\frac{y_1 + y_2 + y_3}{3}$; $\frac{y_2 + y_3 + y_4}{3}$ и т.д.

Интервал сглаживания можно брать четный (четыре, шесть и т.д.). В этом случае, сложность заключается в том, что средняя может быть отнесена только к середине между двумя датами. Необходимо применить центрирование, т.е. нахождение средней для отнесения полученного уровня к определенной дате.

Пример 4. Имеются данные об объеме оказанных услуг, млн р.

2007 г.	2008 г.	2009 г.	2010 г.	2011 г.	2012 г.	2013г.
6,9	7,3	7,1	7,5	7,4	7,3	7,8

Выявить основную тенденцию по объему оказанных услуг за 2005-2011 г.г. методом скользящей средней.

Решение

Исчислим трехлетние скользящие средние уровни ряда за 2007-2009 г.г.:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3} = \frac{6{,}9 + 7{,}3 + 7{,}1}{3} = 7{,}1 \text{млнр}.$$

за 2008-2010 г.г.:

$$\bar{y}_2 = \frac{y_2 + y_3 + y_4}{3} = \frac{7{,}3 + 7{,}1 + 7{,}5}{3} = 7{,}3 \text{млнр}.$$

и т.д.

Результаты расчета представим в таблице 3.

Таблица 3

Годы	Объем услуг, млн р. (y_i)	Скользящие трехлетние суммы ($\Sigma\bar{y}_1$)	Трехлетние скользящие средние ($\frac{\Sigma y_i}{n}$)
2007	6,9(y_1)	-	-
2008	7,3 (y_2)	21,3 ($y_1+y_2+y_3$)	7,1 ($\bar{y}_1$)
2009	7,1 (y_3)	21,9 ($y_2+y_3+y_4$)	7,3 ($\bar{y}_2$)
2010	7,5 (y_4)	22 ($y_3+y_4+y_5$)	7,3 ($\bar{y}_3$)
2011	7,4 (y_5)	22,2 ($y_4+y_5+y_6$)	7,4 ($\bar{y}_4$)
2012	7,3 (y_6)	22,5 ($y_5+y_6+y_7$)	7,5 ($\bar{y}_5$)
2013	7,8 (y_7)	-	-

В результате проведенных расчетов методом скользящей средней проявилась тенденция к росту объема оказываемых услуг.

3. Метод аналитического выравнивания

Сущность метода заключается в отыскании такой прямой или кривой, ординаты точек которой были бы наиболее близки к значениям исходного динамического ряда.

Подбор адекватной математической функции, осуществляется таким образом, чтобы она давала содержательное объяснение изучаемого признака. Чаще всего при выравнивании используются зависимости:

линейная – выбирается в тех случаях, когда в исходном временном ряду наблюдаются боле или менее постоянные абсолютные цепные приросты, не проявляющие тенденции ни к увеличению, ни к снижению;

параболическая – используется, если абсолютные цепные приросты само по себе обнаруживают некоторую тенденцию развития, но абсолютные цепные приросты абсолютных цепных приростов (разности второго порядка) никакой тенденции развития не проявляют;

экспоненциальные – применяются, если в исходном временном ряду наблюдается либо более или менее постоянный относительный рост (устойчивость цепных темпов роста, темпов прироста, коэффициентов роста), либо при отсутствии такого постоянства.

В большинстве расчетов выравнивание осуществляется с помощью способа наименьших квадратов, который обеспечивает наименьшую сумму квадратов отклонений фактических уровней от выравненных.

При выравнивании ряда динамики по прямой используют уравнение:

$$\overline{y}_t = a_0 + a_1 t \text{.}$$

где $\overline{y}_t$ – теоретический уровень ряда;

t – показатель времени (дни, месяцы, годы и т.д.);

$a_{0,} a_1$ – параметры прямой.

Для линейной зависимости $\overline{y}_t = a_0 + a_1 t$ параметр a_0 обычно интерпретации не имеет, но иногда его рассматривают как обобщенный начальный уровень ряда.

Параметр a_1 – сила связи, т.е. параметр, показывающий насколько изменится результат при изменении времени на единицу. Т.о. a_1 можно представить как постоянный теоретический абсолютный прирост.

Для нахождения a_0 и a_1 используется система нормальных уравнений:

$$\begin{cases} \sum y = na_0 + a_1 \sum t. \\ \sum yt = a_0 \sum t + a_1 \sum t^2, \end{cases}$$

где y – фактические уровни ряда;

n – число уровней;

t – показатель времени, который обозначается порядковыми номерами, начиная с низшего.

Для упрощения системы уравнений показатели времени t обозначают так, чтобы $\Sigma t = 0$, т.е. обозначим время так, чтобы начало его отсчета приходилось на середину рассматриваемого периода.

Так как $\Sigma t = 0$, то система нормальных уравнений будет иметь вид:

$$\begin{cases} \sum y = na_0; \\ \sum yt = a_1 \sum t^2 \end{cases}$$

Отсюда:

$$a_0 = \frac{\sum y}{n},$$

$$a_1 = \frac{\sum yt}{\sum t^2}.$$

Рассмотрим применение метода аналитического выравнивания по прямой для выражения основной тенденции на следующем примере.

Пример 3. Реализация товара в торговой сети характеризуется следующими данными:

Год	2006	2007	2008	2009	2010	2011
Объем продажи, млн руб.	6,1	6,3	7,1	7,6	8,0	8,4

Задание. Для изучения общей тенденции объема продажи продукции произведите **аналитическое выравнивание** по прямой и экстраполируя тенденцию, определитель показатель в 2012 году.

Решение

В данном примере число исходных уровней четное, условный отчет (t) в решении произведем следующим образом:

Год	2006	2007	2008	2009	2010	2011
t	-5	-3	-1	+1	+3	+5

Далее произведем расчет необходимых значений, построив вспомогательную таблицу.

Таблица 4

Год	Эмпирические уровни ряда (y_i)	Условное обозначение времени (t)	t^2	$y_i{*}t$	Теоретический уровень ряда ($\overline{y}_t$)	$(y-\overline{y}_t)^2$
2006	6,1	-5	25	-30,5	6,03	
2007	6,3	-3	9	-18,5	6,52	0,048
2008	7,1	-1	1	-7,1	7,01	0,008
2009	7,6	+1	1	7,6	7,49	0,012
2010	8,0	+3	9	24	7,98	0,001
2011	8,4	+5	25	42	8,47	0,005
Итого	$\sum y = 43{,}5$	$\sum t = 0$	$\sum t^2 = 70$	$\sum yt = 17.1$	$\Sigma \overline{y}_t = 43{,}5$	0,074

Определим параметры:

$$a_0 = \frac{43{,}5}{6} = 7{,}25$$ млн руб.

$$a_1 = \frac{17{,}1}{70} = 0{,}244$$ млн руб.

Уравнение прямой, представляющие трендовую модель искомой функции, будет иметь вид:

$$\overline{y}_t = a_0 + a_1 t;$$

$$\overline{y}_t = 7{,}25 + 0{,}244t.$$

Подставляя в уравнение принятые обозначения (t), вычисляем выровненные теоретические значения ($\overline{y}_t$):

2006г. $\overline{y}_t = 7{,}25 + 0{,}244 \cdot (-5) = 6{,}03$

2007г. $\overline{y}_t = 7{,}25 + 0{,}244 \cdot (-3) = 6{,}52$ и т.д.

Правильность расчета уровней выравниваемого ряда динамики может быть проверена следующим образом: сумма значений эмпирического ряда должна быть близка или совпадать с сумой вычисленных уровней выравниваемого ряда, т.е. $\Sigma y_i = \Sigma \overline{y_t}$

После решения уравнения наносим на график фактические уровни и исчисленную прямую линию, характеризующую тенденцию динамического ряда.

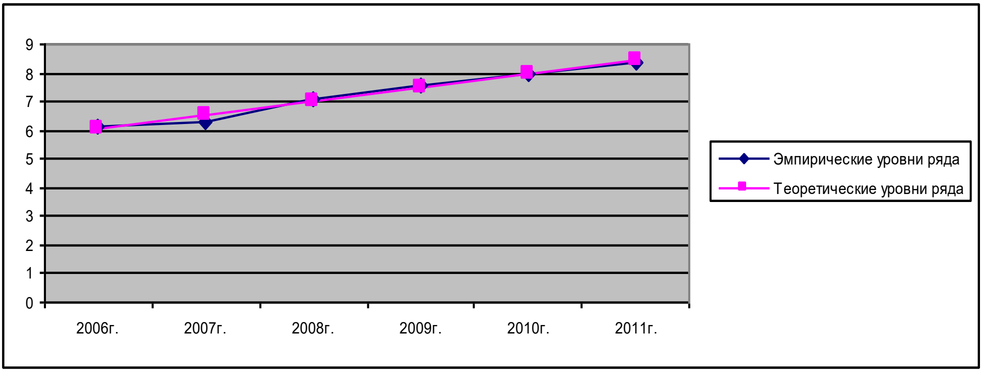


Рисунок 2 – Тенденции динамического ряда

Если число уровней ряда нечетное, то условное обозначение показателя времени примет следующий вид:

Годы	2007	2008	2009	2010	2011
t	-2	-1	0	1	2

Продление в будущее тенденции, наблюдавшейся в прошлом, носит название **экстраполяции.**

Экстраполируя при t =7, находим уровень 2012 года.

Точечный прогноз на 2012г.:

$\overline{y_t} = 7{,}25 + 0{,}244 \cdot 7 = 8{,}96$ млн руб.

Рассчитаем интервальный прогноз объема продажи на 2012г. с вероятностью 0,99.

Интервальный прогноз объёма продажи на 2012г.:

$$S\overline{y_t} = \sqrt{\frac{\Sigma(y - \overline{y_t})^2}{n - l}}$$

где *l* – число параметров в уравнении тренда;

n – число уровней ряда;

t – коэффициент доверия по распределению Стьюдента при уровне значимости α;

$\overline{y}_t$ – точечный прогноз, рассчитанный по модели.

$$S\overline{y}_t = \sqrt{\frac{0{,}074}{8-2}} = 0{,}111$$ млн руб.

Интервальный прогноз объёма продажи на 2012 г.:

$$\overline{y}_{прог} = 8{,}96 \pm 3{,}4 \cdot 0{,}111$$

При вероятности р = 0,99 t_α = 3,4

8,58 млн руб. < $\overline{y}_{прог}$ < 9,33 млн руб.

3. Анализ сезонных колебаний

Под влиянием природно-климатических условий и различных факторов возникают периодические колебания, которые имеют определенный постоянный и определенный период, носят название «сезонные колебания» или «сезонные волны».

Сезонные колебания измеряются при помощи **индексов сезонности**.

Индеек сезонности показывает, во сколько раз фактический уровень ряда в момент или интервал времени (t) больше или меньше уровня среднего или вычисленного по уровню тренда f (t).

Для исчисления индексов сезонности применяют различные методы, выбор которых зависит от характера основной сезонности тенденции ряда динамики.

Если ряд динамики не содержит ярко выраженной тенденции развития, то индексы сезонности исчисляют непосредственно по эмпирическим данным без их предварительного выравнивания.

Так как годовой уровень явления остается относительно неизменным, то индекс сезонности рассчитывается как процентное отношение средних для каждого месяца к общему среднемесячному уровню ряда:

$$i_c = \frac{\overline{y}_i}{\overline{y}_0} \cdot 100,$$

где $\overline{y}_i$ – средняя из фактиче- ских уровней одноименных месяцев;

$\bar{y}_0$ – общая средняя за исследуемый период.

Определяются среднесуточные уровни для каждого месяца:

$$\bar{y}_i = \frac{\sum y_i}{n}.$$

Определяется общая средняя за весь период:

$$\bar{y}_0 = \frac{\sum \bar{y}_i t_k}{365},$$

где t_k – число календарных дней каждого месяца.

Пример 5. Имеются следующие данные об объёме продаж, тыс.р.:

Месяц	2011г.	2012г.	2013г.	Месяц	2011г.	2012г.	2013г.
Январь	62	70	92	Июль	80	92	113
Февраль	64	73	94	Август	83	97	117
Март	67	78	98	Сентябрь	82	90	109
Апрель	72	82	100	Октябрь	79	84	105
Май	75	84	103	Ноябрь	76	81	98
Июнь	79	89	108	Декабрь	71	79	95

Вычислить индексы сезонности.

Решение

Применяя формулу средней арифметической простой ($\bar{y}_i = \frac{\sum y_i}{n}$) определим среднемесячные уровни за три года:

Январь = (62+70+92)/3 = 74 тыс.р.

Февраль= (64+73+94)/3 =77 тыс.р. и т.д.

Исчислим общую (постоянную) среднюю:

$$\bar{y}_0 = \frac{\sum \bar{y}_i t_k}{365}$$

$$\bar{y}_0 = (74\cdot 31 + 77\cdot 28 + 81\cdot 31 + 85\cdot 30 + 87\cdot 31 + 92\cdot 30 + 95\cdot 31 + 99\cdot 31 + \\ + 94\cdot 30 + 89\cdot 31 + 85\cdot 30 + 82\cdot 31):365 = 87\,\text{тыс.р.}$$

Индексы сезонности показывают, что наименьший спрос приходится на январь-февраль, а наибольший – на июль-август.

Расчёт индексов сезонности проведен в табл. 5.

Таблица 5

Месяц	Среднемесячные уровни ряда	Индекс сезонности, %
Январь	74	(74/87*100) 85,05
Февраль	77	88,51
Март	81	93,10
Апрель	85	97,70
Май	87	100,00
Июнь	92	105,74
Июль	95	109,2
Август	99	113,79
Сентябрь	94	108,01
Октябрь	99	102,3
Ноябрь	85	97,7
Декабрь	82	94,25

Для наглядности можно построить график сезонной волны объема продаж (по оси абсцисс располагают месяцы, а по оси ординат – индексы сезонности в процентах). Общая средняя за месяц величина объема продаж располагается на уровне 100,00%, а средние месячные индексы сезонности в виде точек нанесем на поле графика в соответствии с принятым масштабом по оси ординат.

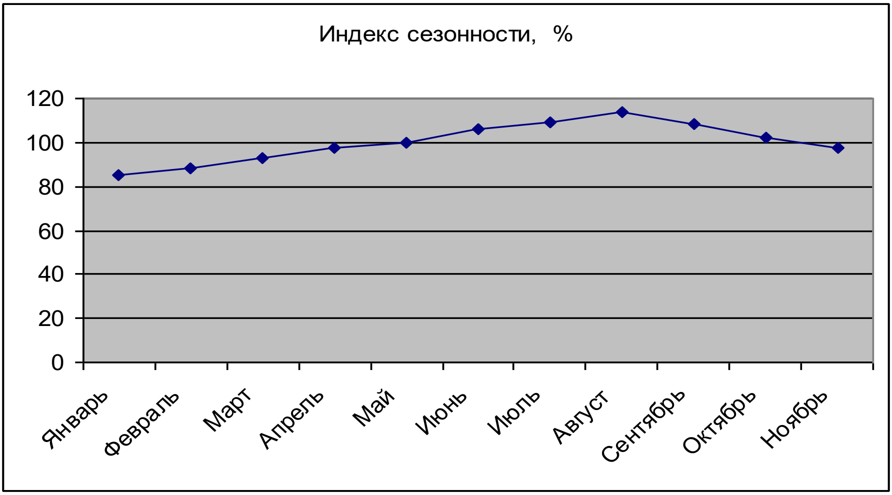


Рисунок 3 – Индекс сезонности

9 ИНДЕКСНЫЙ МЕТОД АНАЛИЗА

Индексы: понятие и их классификация

Понятие и классификация индексов

Индекс (от латинского index) – относительный показатель, характеризующий изменение какого-либо явления во времени, пространстве или по сравнению с любым эталоном (план, прогноз, норматив и т.д.).

Индексы позволяют: измерять изменения сложных явлений; определять влияние отдельных факторов и изменение динамики сложного явления; они являются показателями сравнения во времени, пространстве, с нормативами, планами и прогнозами.

При построении индексов рекомендуется придерживаться следующей символики:

– количество единиц данного вида произведенной или реализованной продукции (q);
– цена единицы изделия (p);
– себестоимость единицы изделия (z);
– трудоемкость единицы изделия (t);
– выработка продукции на одного работающего (w) и т. д.

Подстрочный значок «0» означает базисный, а «1» – отчетный периоды. Индивидуальный индекс обозначается латинской буквой i, а общий – I .

В зависимости **от степени охвата единиц** совокупности индексы делятся на индивидуальные, групповые и общие.

Индивидуальные индексы характеризуют изменение одного элемента совокупности. Они показывают, каково соотношение между отчетным (со знаком «1») и базисным (со знаком «0») показателями или во сколько раз увеличилась (уменьшилась) индексируемая величина.

$$iq = \frac{f_1}{f_0}, \quad ip = \frac{x_1}{x_0}, \quad ipq = \frac{x_1 f_1}{x_0 f_0}.$$

Общие индексы отражают изменение всех элементов статистической совокупности, которые непосредственно не подлежат суммированию.

Агрегатные общие индексы

Индекс стоимости товаров:

$$Iqp = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_0},$$

Индекс показывает во сколько раз возросла (уменьшилась) стоимость продукции (товарооборота) отчетного периода по сравнению с базисным за счет изменения цен на товары и объемов их реализации или сколько процентов составляет рост (снижение) стоимости товаров.

Разница между числителем и знаменателем индекса составляет абсолютное изменение стоимости (прирост или снижение) за счет совместного действия двух факторов: цен и количества.

$$\Delta qp = \sum q_1 p_1 - \sum q_0 p_0.$$

Абсолютное изменение произошло под влиянием факторов:

где $\sum \Delta^p pq$ – абсолютное изменение стоимости, обусловленное изменением уровня цен;

$\sum \Delta^й pq$ – абсолютное изменение стоимости, обусловленное изменением физического объема продаж.

Изменение индекса стоимости зависит от двух факторов: цены и физического объема.

$$\sum \Delta pq = \sum \Delta^p pq + \sum \Delta^q pq$$

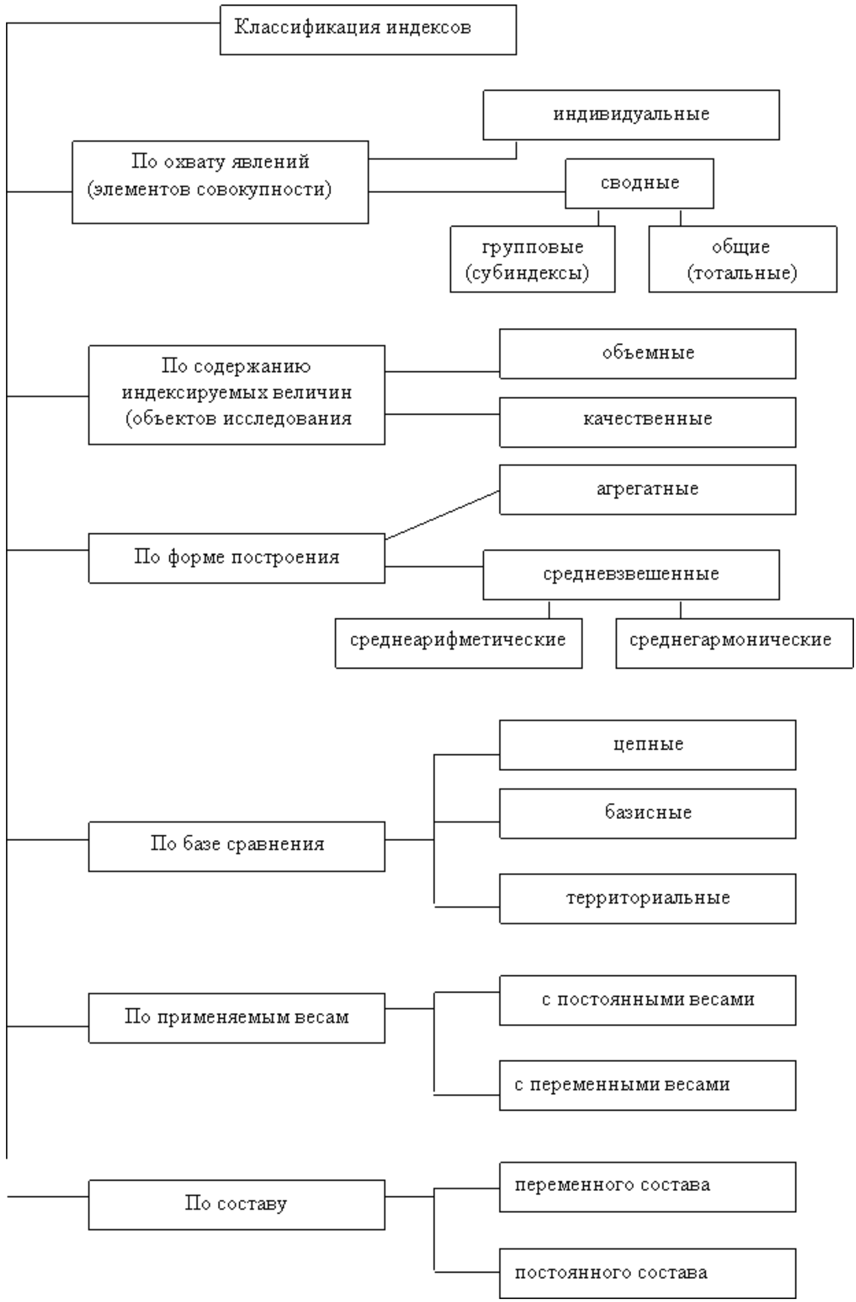


Рисунок 1 – Классификация индексов

Агрегатный общий индекс цен Пааше (по отчетным весам):

$$I_p = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1},$$

где $\sum p_1 q_1$ – фактическая стоимость (товарооборот) отчетного периода;

$\sum q_1 p_0$ – условная стоимость текущем периоде продукции в ценах базисного периода.

Общий индекс цен показывает, во сколько раз возросла (уменьшилась) стоимость в отчетном периоде по сравнению с базисным периодом в результате изменения цен.

$$\Delta qp^{\Delta з} = \sum q_1 p_1 - \sum q_1 p_0 .$$

Разница между числителем и знаменателем индекса показывает, на сколько денежных единиц изменилась стоимость в результате изменения цен или на сколько товары в отчетном периоде стали дороже (дешевле), чем в базисном.

Агрегатный общий индекс цен Ласпейреса:

$$I_p = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}.$$

Индекс цен Ласпейреса показывает, на сколько изменились цены в отчетном периоде по сравнению с базисным, но по той продукции, которая была реализована в базисном периоде, и экономию (перерасход), которую можно было бы получить от изменения цен, т.е. условную экономию (перерасход). Т.е. он показывает, во сколько раз товары базисного периода подорожали (подешевели) из-за изменения цен на них в отчетном периоде.

$$\Delta qp^{\Delta з} = \sum q_0 p_1 - \sum q_0 p_0 .$$

Разница между числителем и знаменателем индекса показывает, на сколько денежных единиц товары в базисном периоде стали дороже дешевле) из-за изменения цен на них в отчетном периоде.

Индекс физического объема (и любых количественных показателей):

$$I_q = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0},$$

где $\sum q_0 p_0$ – стоимость продукции, произведенной в базисном периоде.

Индекс физического объема продукции показывает, во сколько раз изменилась стоимость в отчетном периоде по сравнению с базисным периодом в результате изменения физического объема ее реализации.

$$\Delta qp^{\Delta q} = \sum q_1 p_0 - \sum q_0 p_0 .$$

Разница между числителем и знаменателем индекса означает абсолютное изменение стоимости (прирост или снижение) за счет изменения только физического объема.

Пример 1. Реализация продукции характеризуется данными:

Вид продукции	Объем реализации, штук		Цена реализации, р.	
	Июль (q_0)	Август (q_1)	Июль (p_0)	Август (p_1)
A	50	65	200	250
B	20	25	1300	1200
C	100	120	900	800

Определить:

1) среднее изменение цен на реализованную продукцию и абсолютное изменение стоимости реализованной продукции за счет изменения цен;

2) общее изменение физического объема реализованной продукции предприятия и абсолютное изменение стоимости реализованной продукции за счет изменения ее физического объема.

Решение

1. Среднее изменение цен на реализованную продукцию определяется по формуле агрегатного индекса цен:

$$I_p = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = \frac{250 \cdot 65 + 1200 \cdot 25 + 800 \cdot 120}{200 \cdot 65 + 1300 \cdot 25 + 900 \cdot 120} = 0{,}927.$$

Изменение стоимости реализованной продукции за счет изменения цен:

$$\Delta \Sigma pq_{\Delta p} = \Sigma p_1 q_1 - \Sigma p_0 q_1 = 142250 - 153500 = -11250 \text{ р.}$$

Стоимость реализованной продукции снизилась на 11250 р. или на 7,33% за счет изменения цен.

2. Изменение физического объема реализованной продукции предприятия определяется по формуле агрегатного индекса физического объема продукции:

$$I_q = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} = \frac{65 \cdot 200 + 25 \cdot 1300 + 120 \cdot 900}{50 \cdot 200 + 20 \cdot 1300 + 100 \cdot 900} = 1{,}218.$$

Изменение стоимости реализованной продукции за счет изменения объема реализованной продукции:

$$\Delta \Sigma qp_{\Delta q} = \Sigma q_1 p_0 - \Sigma q_0 p_0 = 153500 - 126000 = +27500 \text{ р.}$$

Стоимость реализованной продукции выросла на 27500 р. или на 21,8 % за счет изменения объема реализованной продукции.

Проверка.

Общий прирост стоимости реализованной продукции составил:

$$\Delta \Sigma qp = \Sigma q_1 p_1 - \Sigma q_0 p_0 = 142250 - 126000 = +16250 \text{ р.}$$

$$\Delta \Sigma qp = \Delta \Sigma qp_{\Delta q} + \Delta \Sigma qp_{\Delta p} = 27500 + (-11250) = +16250 \text{ р.}$$

Для определения степени влияния факторов на общий прирост используют метод цепных подстановок, суть которого состоит в многошаговом переходе от базисного значения результативного показателя к отчетному, путем последовательной (цепной) замены базисных значений факторов на отчетные.

Метод цепной подстановки требует соблюдения следующего правила: подстановка отчетных данных в место базисных начинается с количественного (экстенсивного) фактора и заканчивается качественным (интенсивным) фактором.

Пример 2. Имеются данные о реализации товара «А» на рынке города:

Показатель	Январь	Март
Цена 1 кг., руб. (p)	25	32
Объем продаж, кг (q)	1300	1100

Определить изменение выручки за счет влияния факторов.

Решение

Метод цепных подстановок:

$$\Delta B = B_1 - B_0 = 35200 - 32500 = 2700 руб.$$

$$B_0 = p_0 * q_0 = 25 * 1300 = 32500 руб.$$

$$B' = p_0 * q_1 = 25 * 1100 = 27500 руб.$$

$$B'' = p_1 * q_1 = 32 * 1100 = 35200 руб.$$

Изменение выручки за счет изменения физического объема продаж:

$$\Delta B_{\Delta q} = B' - B_0 = 27500 - 32500 = -5000 руб.$$

Изменение выручки за счет изменения физического цены:

$$\Delta B_{\Delta p} = B'' - B' = 35200 - 27500 = 7700 руб.$$

Проверка: Суммарное влияние факторов:

$\Delta B_1 = \Delta B_{1\Delta p} + \Delta B_{1\Delta q}$ = 7700+ (-5000)=2700 руб.

Метод абсолютных разниц:

$$B_0 = p_0 * q_0 = 25*1300 = 32500 руб.$$
$$B_1 = p_1 * q_1 = 32*1100 = 35200 руб.$$
$$\Delta B = B_1 - B_0 = 35200 - 32500 = 2700 руб.$$

Выручка в марте по сравнению с январем вырос на 2700 руб. за счет одновременного изменения и цены, и количества проданного товара.

Изменение выручки за счет изменения цены:

$$\Delta B_{1\Delta p} = (p_1 - p_0) * q_1 = (32 - 25) * 1100 = 7700 руб.$$

Изменение выручки за счет изменения физического объема продаж:

$$\Delta B_{1\Delta q} = (q_1 - q_0) * p_0 = (1100 - 1300) * 25 = -5000 руб.$$

Проверка: $\Delta B_1 = \Delta B_{1\Delta p} + \Delta B_{1\Delta q}$ = 7700+ (-5000)=2700 руб.

Т.о. рост выручки был обусловлен ростом цены, при этом сокращение физического объема продаж привел к снижению выручки на 5000 руб.

Средние индексы

При анализе экономической деятельности пользуются не только объемными, но и средними показателями.

Средняя величина, является характеристикой качественного показателя и складывается под влиянием значений показателя у отдельных единиц, из которых состоит объект, но и под влиянием соотношения их весов (структуры).

Динамику среднего показателя можно отразить как за счет изменения обоих факторов, так и за счет каждого фактора отдельно. Для этого используются индексы переменного состава, постоянного состава и структурных сдвигов.

Индекс переменного состава отражает динамику среднего показателя за счет изменения индексируемой величины (x) и за счет изменения весов (f), по которым взвешиваются отдельные значения.

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\Sigma x_1 f_1}{\Sigma f_1} \div \frac{\Sigma x_0 f_0}{\Sigma f_0}$$

Разница между дробями индекса переменного состава показывает абсолютное изменение среднего уровня признака как за счет изменения значений самого признака у отдельных единиц совокупности, так и за счет структурных изменений:

$$\Delta\bar{x} = \bar{x}_1 - \bar{x}_0 = \frac{\Sigma x_1 f_1}{\Sigma f_1} - \frac{\Sigma x_0 f_0}{\Sigma f_0} \text{ или } \Delta\bar{x} = \sum x_1 d_1 - \sum x_0 d_0.$$

Индекс постоянного состава отражает динамику среднего показателя лишь за счет изменения индексируемой величины при фиксированном весе.

$$I_x = \frac{\Sigma x_1 f_1}{\Sigma f_1} \div \frac{\Sigma x_0 f_1}{\Sigma f_1} \text{ или } I_x = \frac{\Sigma x_1 f_1}{\Sigma x_0 f_1}.$$

Разница между дробями индекса постоянного состава показывает абсолютное изменение среднего уровня признака за счет изменения значений самого признака у отдельных единиц совокупности:

$$\Delta\bar{x}_{\Delta x} = \frac{\Sigma x_1 f_1}{\Sigma f_1} - \frac{\Sigma x_0 f_1}{\Sigma f_1} \text{ или } \Delta\bar{x}_{\Delta x} = \sum x_1 d_1 - \sum x_0 d_1.$$

Индекс структурных сдвигов отражает динамику среднего показателя за счет изменения весов.

$$I_{стр} = \frac{\Sigma x_0 f_1}{\Sigma f_1} \div \frac{\Sigma x_0 f_0}{\Sigma f_0}.$$

Под структурными изменениями понимается изменение доли отдельных групп единиц совокупности в общей их численности.

Разница между дробями индекса структурных сдвигов показывает абсолютное изменение среднего уровня признака за счет структурных изменений:

$$\Delta \bar{x}_{\Delta f} = \frac{\Sigma x_0 f_1}{\Sigma f_1} - \frac{\Sigma x_0 f_0}{\Sigma f_0} \quad \text{или} \quad \Delta \bar{x}_{\Delta d} = \sum x_0 d_1 - \sum x_0 d_0.$$

Между индексами существует взаимосвязь:

$$I_{пер.} = I_{пост.} * I_{стр.}$$

Совместное влияние факторов:

$$\Delta \bar{x} = \Delta \bar{x}_{\Delta x} + \Delta \bar{x}_{\Delta d}.$$

Пример 3. Имеются данные о выпуске однородной продукции по предприятиям акционерного общества:

№ предприятия	Выпуск, тыс. шт		Себестоимость, р.	
	1 кв. (q_0)	2 кв. (q_1)	1 кв. (z_0)	2 кв. (z_1)
1	40	36	7	8
2	60	84	6	6,5
Итого	100	120	6,4	6,9

Определить: изменение себестоимости по акционерному обществу, за счет влияния факторов.

Решение

1. Индекс себестоимости продукции переменного состава:

$$I_{пер} = \frac{\Sigma z_1 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_0}{\Sigma q_0} = \frac{6{,}95}{6{,}4} = 1{,}086$$

, или 108,6 %

Себестоимость единицы продукции по двум предприятиям выросла на 8,6 %, за счет изменения себестоимости единицы продукции и объема по каждому предприятию.

2. Индекс себестоимости постоянного состава:

$$I_{пост} = \frac{\Sigma z_1 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_1}{\Sigma q_1} = \frac{6{,}95}{6{,}3} = 1{,}103$$

, или 110,3 %

Себестоимость единицы продукции по двум предприятиям выросла на 10,3%, за счет изменения себестоимости единицы продукции по каждому предприятию.

3. Индекс структурных сдвигов:

$$I_{стр} = \frac{\Sigma z_0 q_1}{\Sigma q_1} : \frac{\Sigma z_0 q_0}{\Sigma q_0} = 0{,}985$$

, или 98,5 %

Себестоимость единицы продукции по двум предприятиям снизилась на 1,5 %, за счет изменения объема продаж по каждому предприятию.

Между индексами существует взаимосвязь:

$$I_{пер} = I_{пост} \cdot I_{стр} = 1{,}103 \cdot 0{,}985 = 1{,}086.$$

Средние индексы из индивидуальных

Средняя арифметическая форма агрегатного индекса физического объема стоимости продукции используется в том случае, когда известны объемы стоимости продукции базисного периода $(pq)_0$, но неизвестны по отдельности цены (p) и количества товара (q).

$$i_q = \frac{q_1}{q_0} \rightarrow I_q = \frac{\Sigma p_0 q_1}{\Sigma p_0 q_0} \rightarrow q_1 = i_q * q_0 \rightarrow I_q = \frac{\Sigma i_q p_0 q_0}{\Sigma p_0 q_0}$$

Здесь агрегатный индекс физического объема представлен как средневзвешенная арифметическая из индивидуальных индексов физических объемов, и весами в ней являются объемы стоимости продукции базисного периода в текущих ценах.

Средняя гармоническая форма индекса цен используется в тех случаях, когда известны объемы стоимости продукции в текущих ценах $(pq)_1$, но неизвестны по отдельности цены (p) и количества товара (q).

$$i_p = \frac{p_1}{p_0} \rightarrow Ip = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \rightarrow p_0 = \frac{p_1}{i_p} \rightarrow I_p = \frac{\Sigma p_1 q_1}{\Sigma \frac{p_1 q_1}{i_p}}$$

Здесь агрегатный индекс цен представлен как средневзвешенная гармоническая из индивидуальных индексов цен, где весами являются объемы стоимости продукции отчетного периода в текущих ценах.

Пример 4. По заводу имеются следующие данные по выпуску продукции.

Вид продукции	Выпуск продукции в 1-м квартале, млн р.	Изменение выпуска во 2-м квартале, по сравнению с 1-м кварталом, %
А	3,0	+ 1,5
Б	3,8	- 1,8

Определить, на сколько процентов изменился выпуск продукции по заводу.

Решение

Для определения изменения физического общего объема продукции, используется формула среднего взвешенного арифметического индекса, так как по условиям задачи известны индивидуальные индексы физического объема.

Индивидуальные индексы по двум видам продукции:

- продукция А- $i_q = \frac{100+1{,}5}{100} = 1{,}015;$

- продукция Б- $i_q = \frac{100-1{,}8}{100} = 0{,}982;$

$$I_q = \frac{\Sigma i_q q_0 p_0}{\Sigma q_0 p_0} = \frac{1{,}015 \cdot 3 + 0{,}982 \cdot 3{,}8}{3+3{,}8} = 0{,}9966,$$ или 99,66 %

Следовательно, физический объем продукции в целом по предприятию снизился на 0,34 %.

Пример 5. Средняя выработка на одного рабочего увеличилась на 2,5 %. Объем произведенной продукции вырос на 4 %. Определить, как изменилась в среднем численность рабочих.

Решение

Для определения индекса численности используется взаимосвязь между тремя индексами:

$$I_{cw} = I_c \times I_w$$

$$I_c = I_{cw} : I_w = 1{,}04 : 1{,}025 = 1{,}0146,$$ или 101,46 %

Следовательно, численность рабочих выросла на 1,46 %.

Пример 6. В отчетном периоде было реализовано товара А на 2,8 млн р., товара Б – на 3,2 млн р., товара С – на 1,8 млн р.

Исчислить общий индекс цен на все товары, если известно, что цены на товар А были снижены на 2,8 %, на товар Б повысились на 1,3 %, а на товар С остались без изменения.

Решение

Для характеристики изменения цен на все виды товара вместе исчисляется общий индекс цен по формуле средневзвешенного гармонического индекса, поскольку известны индивидуальные индексы цен по каждому виду товара:

$$I_p = \frac{\Sigma p_1 q_1}{\Sigma \frac{p_1 q_1}{i_p}} = \frac{2{,}8 + 3{,}2 + 1{,}8}{\frac{2{,}8}{0{,}972} + \frac{3{,}2}{1{,}013} + \frac{1{,}8}{1{,}0}} = 0{,}9949 \text{ или } 99{,}49\%$$

По всему ассортименту товаров цены в отчетном периоде снизились в среднем на 0,51 %.

Пример 7. Затраты на производство продукции по промышленному предприятию за отчетный период выросли на 12 %, себестоимость единицы продукции при неизменной структуре производства увеличилась на 3 %, количество произведенных изделий возросло на 5 %.

Определите, как повлияли на изменение общей суммы затрат структурные изменения в производстве изделий.

Решение

$$I_{zq} = I_{стр.} I_q I_z,$$

где I_q – индекс количества произведенных изделий;
$I_{стр.}$ – индекс влияния структурных сдвигов в производстве продукции;
I_z – индекс затрат.

$$I_{стр.сдв} = \frac{Izq_1}{Iq \cdot Iz_1} = \frac{1{,}12}{1{,}03 \cdot 1{,}05} = 1{,}036, \text{ или } 103{,}6\ \%.$$

Следовательно, в результате увеличения доли изделий с наиболее высокими затратами на их производство, общая сумма затрат увеличилась на 3,6 %.

Базисная и цепная система индексов

Для изучения динамики экономических явлений за несколько периодов используется система базисных и цепных индексов.

Если индексируемые величины сравниваются с уровнем одного и того же периода, то говорят о системе *базисных индексов*. При оценке относительного изменения уровня изучаемого явления по сравнению с предшествующим периодом получают систему *цепных индексов*.

Индексы стоимости продукции:

а) цепные индексы:

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0};\frac{\Sigma p_2 q_2}{\Sigma p_1 q_1};...;\frac{\Sigma p_n q_n}{\Sigma p_{n-1} q_{n-1}}$$

б) базисные индексы:

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0};\frac{\Sigma p_2 q_2}{\Sigma p_0 q_0};...;\frac{\Sigma p_n q_n}{\Sigma p_0 q_0}$$

В системе базисных индексов сравнения уровней индексируемого показателя в каждом индексе производят с уровнем базисного периода, а в системе цепных индексов уровни индексируемого показателя сопоставляются с уровнем предыдущего периода.

Базисные индексы физического объема продукции с постоянными весами (p_0):

$$\frac{\Sigma q_1 p_0}{\Sigma p_0 q_0};\frac{\Sigma q_2 p_0}{\Sigma p_0 q_0};...;\frac{\Sigma q_n p_0}{\Sigma p_0 q_0}.$$

Цепные индексы физического объема продукции с постоянными весами (p_0):

$$\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0};\frac{\Sigma q_2 p_0}{\Sigma q_1 p_0};...;\frac{\Sigma q_n p_0}{\Sigma q_{n-1} p_0}.$$

Базисные индексы цен с переменными весами:

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1};\frac{\Sigma p_2 q_2}{\Sigma p_0 q_2};...;\frac{\Sigma p_n q_n}{\Sigma p_0 q_n}.$$

Между цепными и базисными индексами существует взаимосвязь:

- произведение цепных индексов равно конечному базисному;
- частное от деления двух смежных базисных индексов равно промежуточному цепному.

Индексы, построенные на основе переменных весов, непосредственно перемножать и делить нельзя.

10. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Статистические методы изучения взаимосвязи социально-экономических явлений

Между явлениями и их признаками различают прежде всего два вида связей: функциональные и стохастические, каждая из которых имеет свои особенности. Частным случаем стохастических связей являются корреляционные.

При **функциональной (полной) связи** изменение результативного признака (y) всецело зависит от изменения факторного признака (x): y = f (x). Каждому значению величины факторного признака соответствует только одно или несколько точно определенных значений.

Стохастическая связь – это связь между величинами, при которой одна из них, случайная величина y, реагирует на изменение другой величины (x) или других величин x_1,x_2 …x_n (случайных или неслучайных) изменением закона распределения. Это обуславливается тем, что зависимая переменная (результативный признак), кроме рассматриваемых независимых, подвержена влиянию ряда неучтенных или неконтролируемых (случайных) факторов, а также некоторых неизбежных ошибок измерения переменных.

При **корреляционной (неполной) связи** средняя величина результативного признака (y)меняется под влиянием изменения многих факторных признаков (x). Корреляционные связи обнаруживаются не в единичных случаях, а в массе. При наличии корреляционной зависимости устанавливается лишь тенденция изменения результативного признака при изменении величины факторного признака.

Термин «корреляция» был введен в науку выдающимся английским естествоиспытателем Френсисом Гальтоном в 1886 г. Однако точную формулу для подсчета коэффициента корреляции разработал его ученик Карл Пирсон.

Задача корреляционного анализа сводится к установлению направления (положительное или отрицательное) и формы (линейная, нелинейная) связи между варьирующими признаками, измерению ее тесноты, и, наконец, к проверке уровня значимости полученных коэффициентов корреляции.

Корреляционные связи различаются по:

1) аналитической форме:

– линейные;
– нелинейные.

– *прямолинейные (линейные*) – когда с возрастанием значения факторного признака происходит непременное возрастание (или убывание) значений результативного признака (например, связь между количеством тренировок на тренажере и количеством правильно решаемых задач в контрольной сессии). Математически эта связь представляется уравнением прямой, а графически – прямой линией;

– *криволинейые (нелинейные*) – когда с возрастанием значения факторного признака возрастание (или убывание) результативного признака происходит неравномерно или же направление его изменения меняется на обратное (например, связь между уровнем мотивации и эффективностью выполнения задачи). Геометрически такие связи представляются кривыми линиями (гиперболой, параболой и т.д.).

2) направлению действия:

– *прямые (положительные)* – более высоким значениям одного признака соответствуют более высокие значения другого, а более низким значениям одного признака – низкие значения другого. При этом коэффициент корреляции имеет положительный знак.

– *обратные (отрицательные)* – более высоким значениям одного признака соответствуют более низкие значения другого, а более низким значениям одного признака – высокие значения другого. При этом коэффициент корреляции имеет отрицательный знак.

3) количеству взаимодействующих факторов:

– *однофакторные (парные)* – если характеризуется связь двух признаков;

– *многофакторные (множественные)* – если изучаются более чем две переменные.

4) силе связи:

– сильные;

– слабые.

Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции.

Силу связи между признаками можно оценить по шкале Чеддока:

$0,1 < r < 0,3$ – слабая;

$0,3 < r < 0,5$- умеренная;

$0,5 < r < 0,7$ – заметная;

$0,7 < r < 0,9$ – высокая;

$0,9 < r < 1,0$ – весьма высокая.

Кроме этого выделяют связи:

непосредственные – когда два фактора взаимодействуют между собой непосредственно;

косвенные – кода участвует какая-то третья переменная, определяющая связь между изучаемыми факторами;

ложные – не имеет в себе смысловой основы.

Для изучения статистических взаимосвязей применяют два **метода анализа:**

1) *корреляционный*, задачи которого, сводятся к измерению тесноты связи между факторами, выявлению неизвестных причин связей и оценке факторов, оказывающих максимальное влияние на результат;

2) *регрессионный анализ* – задачи которого, лежат в сфере установления формы зависимости, определения уравнения регрессии и его использования для оценки неизвестных зависимостей переменной.

Прямолинейная связь между факторами исследуется с помощью уравнения регрессии:

$$\overline{y}_x = a_0 + a_1 x,$$

где: a_0, a_1 – коэффициенты (параметры) уравнения регрессии.

Теснота корреляционной связи между переменными x и y может быть измерена **эмпирическим корреляционным отношением,** которое представляет собой корень квадратный из отношения межгрупповой дисперсии (δ^2)к общей дисперсии (σ^2):

$$\eta_э = \sqrt{\frac{\delta^2}{\sigma^2}}$$

Теоретическое корреляционное отношение характеризует долю общей вариации результативного признака, объясняемую на основе выбранного уравнения связи результативного и факторного признаков.

$$\eta_T = \sqrt{\frac{\sigma^2_{(y/x)}}{\sigma^2_{(y)}}}$$

Коэффициент детерминации (мера определенности, причинности) равен квадрату множественного коэффициента корреляции.

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

Он характеризует долю дисперсии результативной переменной, обусловленной влиянием независимых переменных, входящих в модель, т.е. долю вариации результативного признака под влиянием вариации признака-фактора.

Может принимать значения от 0 до 1.

При непрямолинейной форме для измерения тесноты связи определяется **индекс корреляции (R):**

$$R_{xy} = \frac{r_{xy}}{\sigma_x \sigma_y}.$$

При линейной зависимости между признаками для определения тесноты корреляционной связи применяются **линейный коэффициент корреляции**:

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

где x, y – значения факторного и результативного показателей соответственно;

$\bar{x}, \bar{y}$ – средние значения соответствующих показателей;

Линейный коэффициент корреляции не может превышать +1 и быть меньше чем -1. Эти два числа +1 и -1 — являются границами для коэффициента корреляции. Когда при расчете получается величина большая +1 или меньшая -1 — следовательно произошла ошибка в вычислениях.

Если знак коэффициента линейной корреляции — плюс, то связь между коррелирующими признаками такова, что большей величине одного признака (переменной) соответствует большая величина другого признака (другой переменной). Иными словами, если один показатель (переменная) увеличивается, то соответственно увеличивается и другой показатель (переменная). Такая зависимость носит название **прямо пропорциональной зави-**

симости.

Если же получен знак минус, то большей величине одного признака соответствует меньшая величина другого. Иначе говоря, при наличии знака минус, увеличению одной переменной (признака, значения) соответствует уменьшение другой переменной. Такая зависимость носит название **обратно пропорциональной зависимости**.

Парный коэффициент корреляции показывает тесноту линейной зависимости между двумя признаками на фоне действия остальных, входящих в модель:

$$r_{yx1} = \frac{\overline{x_1 y} - \overline{x_1}\,\overline{y}}{\sigma_{x1}\sigma_y}$$

Он может принимать значения от -1 до 1.

Частный коэффициент корреляции показывает тесноту линейной зависимости между двумя признаками при исключении влияния остальных, входящих в модель/

Множественный коэффициент корреляции отражает связь между результативным и несколькими факторными признаками.

Может принимать значения от 0 до 1.

Корреляционное отношение:

$$\eta = \sqrt{\frac{\sum(\tilde{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}} \Rightarrow \quad \eta = \sqrt{1 - \frac{\sigma^2_{ост}}{\sigma^2_{общ}}} = \sqrt{1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y})^2}}$$

В зависимости от количества признаков, включенных в модель, корреляционные связи делят на парные и множественные.

Тесноту связи между двумя *альтернативными признаками* (представленными только группами с противоположными, взаимоисключающими характеристиками) можно измерить с помощью коэффициента:

Ассоциации:

$$K_a = \frac{ad - bc}{ad + bc}$$

Изменяется от -1 до +1; чем ближе к +1 или -1, тем сильнее связаны между собой изучаемые признаки. Если коэффициент не менее 0,3, то это свидетельствует о наличии связи между качественными признаками.

Контингенции:

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

Коэффициент контингенции всегда меньше коэффициента ассоциации.

Если коэффициент ассоциации 0,5, а коэффициент контингенции 0,3, то можно сделать вывод о наличии существенной зависимости между изучаемыми признаками.

Если необходимо оценить тесноту связи между альтернативными признаками, которые могут принимать любое число вариантов значений, применяется коэффициент взаимной сопряженности Пирсона:

$$K_n = \sqrt{\frac{\varphi^2}{1+\varphi^2}}$$

где φ^2 – показатель средней квадратической сопряженности.

$$\varphi^2 = \sum \frac{{m_{xy}}^2}{m_x m_y} - 1.$$

Коэффициент взаимной сопряженности Чупрова:

$$K_Ч = \sqrt{\frac{\varphi^2}{(К1-1)(К2-1)}}.$$

где К$_1$ – число возможных значений первой статистической величины (число групп по столбцам);

К$_2$ – число возможных значений второй статистической величины (число групп по строкам);

φ^2 – показатель взаимной сопряженности (определяется как сумма отношений квадратов частот клетки таблицы распределения к произведению итоговых частот соответствующего столбца и строки).

Коэффициент изменяется от 0 до 1, но уже при значении 0,3 можно говорить о тесноте связи между вариацией изучаемых признаков.

Различают два основных вида корреляции (от лат.: «соотношение») признаков: линейная и нелинейная. Отсюда – многомерный и двумерный, регрессионный и корреляционный анализы применяются при исследовании линейных и нелинейных корреляционных связей признаков.

Пример 1. При контрольной проверке качества поступившей продукции в торговлю партии товара получены следующие данные об удельных весах стандартной продукции по категориям:

Категория продукции	Стандартная продукция	Нестандартная продукция	Итого
Высшая	90	10	100
Первая	70	30	100
Вторая	50	50	100
Итого	120	90	300

Для определения тесноты связи между категорией продукции и ее качеством найдите коэффициенты взаимной сопряженности Пирсона и А.А.Чупрова.

Решение

Коэффициент взаимной сопряженности Пирсона:

$$K_n = \sqrt{\frac{\varphi^2}{1+\varphi^2}},$$

$$\varphi^2 = \sum \frac{m_{xy}{}^2}{m_x m_y} - 1.$$

$$\varphi^2 = \sum \frac{\sum \frac{m_{xy}{}^2}{m_x}}{m_x} - 1 =$$

$$\frac{\frac{90^2}{210}+\frac{10^2}{90}}{100} + \frac{\frac{70^2}{210}+\frac{30^2}{90}}{100} + \frac{\frac{50^2}{210}+\frac{50^2}{90}}{100} - 1 = \frac{39{,}68}{100} + \frac{33{,}33}{100} + \frac{39{,}68}{100} - 1 = 1{,}1269 - 1 = 0{,}127$$

$$K_n = \sqrt{\frac{\varphi^2}{1+\varphi^2}} = \sqrt{\frac{0{,}127}{1+0{,}127}} = 0{,}34$$

Коэффициент взаимной сопряженности: А.А.Чупрова

$$K_ч = \sqrt{\frac{\varphi^2}{(K1-1)(K2-1)}} = \sqrt{\frac{0{,}127}{2*1}} = 0{,}25.$$

По проведенным расчетам можно отметить, что связь между качественными признаками слабая.