

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Брянский государственный аграрный университет»
Брасовский промышленно-экономический техникум

Л.А. Чиркова

Статистика

Учебное пособие

Рекомендовано Учебно — методическим советом филиала в качестве учебно — методического пособия по специальности 38.02.01 Экономика и бухгалтерский учет (по отраслям)

Брянская область 2015

УДК 372.834
ББК 74.57
Ш 34

Чиркова Л.А. **Статистика**: учебное пособие / Л.А. Чиркова. – Брянск: ФГБОУ ВО Брянский ГАУ, 2015. – 90 с.

В учебном пособии рассмотрены основные статистические методы, применяемые для сбора, обобщения, представления и анализа массовых данных, относящихся к социальным или экономическим явлениям и процессам.

Для студентов среднего профессионального образования, обучающихся по специальности 38.02.01 «Экономика и бухгалтерский учет (по отраслям)».

Рецензенты

Юдина Н.Г. преподаватель экономических дисциплин (Брасовский филиал ФГБОУ ВО Брянский ГАУ)

Другова Г.Е., методист (Брасовский филиал ФГБОУ ВО Брянский ГАУ)

Рекомендовано к изданию решением учебно-методическим советом филиала ФГБОУ ВО «Брянский аграрный университет» - Брасовский промышленно-экономический техникум от 25.05.2015 года, протокол № 5.

© Брянский ГАУ, 2015

© Чиркова Л.А., 2015

СОДЕРЖАНИЕ

	Стр.
1. Предмет, метод и задачи статистики. Организация государственной статистики в РФ	4
2. Статистическое наблюдение и этапы его проведения	11
3. Формы статистического наблюдения. Статистическая отчетность и ее виды	16
4. Статистическая сводка	22
5. Метод группировок в статистике	26
6. Абсолютные и относительные величины в статистике	56
7. Средние величины в статистике	62
8. Показатели вариации в статистике	70
9. Ряды динамики. Показатели изменения уровней рядов динамики	74
10. Основные компоненты динамического ряда. Методы анализа основной тенденции в рядах динамики	82
11. Понятие и классификация индексов	89
12. Выборочное наблюдение. Способы формирования выборочной совокупности. Методы оценки результатов выборочного наблюдения	100
13. Методы изучения связи между явлениями	108

«ПРЕДМЕТ, МЕТОД И ЗАДАЧИ СТАТИСТИКИ. ОРГАНИЗАЦИЯ ГОСУДАРСТВЕННОЙ СТАТИСТИКИ»

1. Происхождение термина «статистика» и его значение

В настоящее время термин «статистика» употребляется в различных значениях.

Статистика – это общественная наука, изучающая явления и процессы общественной жизни, она раскрывает законы возникновения и развития этих явлений и их взаимосвязи. Для того чтобы изучить статистическую науку, необходимо иметь представление о предмете её исследования и знать, какие научные принципы положены в основу этих явлений.

Статистика – это особо важная наука, т. е. отрасль знаний, изучающая с количественной стороны все явления из жизни общества.

В переводе с латинского слово «status» означает определённое положение вещей. Термин «статистика» впервые был употреблён немецким учёным Г. Ахенвалем в 1749 г., в его книге о государственном управлении.

Статистика понимается в настоящее время в трёх значениях:

1) **статистическая наука** – вся практическая деятельность человека по сбору, обработке, накоплению и анализу цифровых данных, которые характеризуют образование, экономику страны, её культуру и другие жизненно важные явления в жизни общества;

2) **статистика** – наука, которая занимается разработкой технических положений и методов, используемых статистической практикой.

Существует тесная связь между статистической наукой и статистической практикой. Статистическая практика применяет правила, которые разработала статистическая наука, но в то же время статистическая наука опирается на те материалы, которые были получены статистической практикой, обобщает её опыт и разрабатывает на основе всего этого свои новые положения;

3) представленные предприятием, организацией **статистические данные** в виде финансовой отчётности называют статистикой, а также могут быть использованы данные, которые публикуются в справочниках, в периодических изданиях, в сборниках, они и представляют собой результат статистической работы.

Статистика – это инструмент познания. **Особенности статистики:**

1) в количественном выражении сообщаются статистические данные;

2) статистическую науку интересуют выводы, сделанные в результате анализа собранных и обработанных числовых данных;

3) состояние изучаемого явления на определённой ступени его развития в конкретных условиях места и времени отражают статистические данные.

Рост производительных сил и научной деятельности в России вызвал развитие статистики и применение её в практической деятельности.

Основные задачи статистической науки:

1) исследовать происходящие в обществе преобразования социально-экономических процессов;

2) выявить резервы эффективности общественного производства;

3) своевременно обеспечить органы законодательной власти надёжной информацией

2. Предмет изучения статистики

Любая наука обладает специфическими особенностями, которые отличают её от других наук и дают ей право на самостоятельное существование. В предмете познания, в принципах и методах изучения науки, которые в совокупности образуют её методологию, заключается главная особенность каждой науки.

Предметом статистической науки являются:

- 1) массовые социально–экономические явления жизни;
- 2) количественная сторона этих явлений в конкретных условиях места и времени.

Посредством статистических показателей статистика изучает все явления и процессы, протекающие в жизни общества.

Количественная оценка свойства изучаемого объекта – это статистический показатель. В зависимости от функции статистические показатели можно разделить на: аналитические показатели, учётно–оценочные показатели.

Аналитические показатели применяются для того, чтобы проанализировать статистическую информацию и охарактеризовать особенность развития изучаемых явлений: скорость развития во времени, типичность признака, соотношение его отдельных частей, меру распределения в пространстве и т. д.

Относительные и средние величины, показатели вариации и динамики, тесноты связи и многие другие применяются в статистической науке в качестве аналитических показателей.

Учётно–оценочные показатели – это статистическая характеристика размера качественно определённых социально–экономических явлений в конкретных условиях места и времени.

Учётно–оценочные показатели могут отображать объёмы распространения их в пространстве или достигнутые на определённые моменты уровни развития.

В статистике признаки могут выражаться смысловыми понятиями и числовыми значениями.

Атрибутивными принято называть признаки, которые выражаются смысловыми понятиями, например, к атрибутивному признаку можно отнести пол человека – мужчина и женщина. Если атрибутивные признаки принимают одно из двух противоположных значений, их называют альтернативными.

Количественными принято называть признаки, которые выражены числовыми значениями, например, получаемая заработная плата, возраст, получаемый доход, стаж работы.

Варьирующими называют признаки, принимающие различные значения у отдельных единиц изучаемого явления. Значение варьирующего признака у отдельных единиц изучаемого явления называется **вариантой**.

В статистическом исследовании признаки подразделяются на:

1) основные – определяют основное (главное) содержание изучаемого объекта;

2) второстепенные – это признаки, которые непосредственно не связаны с основным их содержанием.

Изучая свой предмет, статистическая наука образует статистические совокупности.

Множество единиц, объединённых в соответствии с задачей исследования качественной единой основой, называют **статистической совокупностью**.

Единицей совокупности называют первичный элемент статистической совокупности. Единица совокупности является носителем признаков, подлежащих регистрации, и основой ведущегося при обследовании счёта.

Статистический показатель – та категория, которая отображает количественные характеристики соотношения признаков общественных явлений. Статистические показатели бывают:

- 1) объёмными (численность населения);
- 2) расчётными (средние величины);
- 3) плановыми;
- 4) отчётными;
- 5) прогностическими.

Статистические показатели нужно отличать от понятия «статистические данные», так как статистические данные – это конкретные численные значения статистических показателей. Статистические данные зависят от конкретных условий места и времени и определены как количественные и качественные.

Опираясь на основные положения экономической теории статистика, обогащает экономические науки фактами, которые она получает в результате статистического исследования.

3. Метод статистики

В исторической обусловленности статистика изучает динамику социально–экономических явлений.

Статистическая методология – это разнообразные методы, применяемые для изучения своего предмета.

В настоящее время знание статистики необходимо каждому специалисту для того, чтобы принять решения в условиях стохастики, проанализировать элементы рыночной экономики.

Статистика опирается на диалектические категории случайного и необходимого, единичного и массового, индивидуального и общего, качественного и количественного.

Согласно диалектическому методу познания общественные явления и процессы, протекающие в общественной жизни страны, рассматриваются в развитии, взаимной связи и причинной обусловленности. Знание законов общественного развития даёт нам фундамент для правильности толкования явлений, подлежащих статистическому исследованию, а также помогает выбрать надлежащую методику их изучения и анализа.

Основные этапы экономико–статистического исследования:

1 этап: с помощью массового научноорганизованного наблюдения получают первичную информацию об отдельных фактах (единицах) изучаемого явления.

Собранная в ходе массового наблюдения информация представляет собой исходный материал для статистического обобщения, для получения объективных выводов об изучаемом явлении.

Для того чтобы освободиться от влияния случайных причин и установить характерные черты изучаемого объекта, нужно получить сведения о достаточно большом числе единиц.

2 этап: это группировка и сводка материалов, которые представляют собой расчленение всей массы единиц на однородные группы и подгруппы, и оформление полученных результатов в виде статистических таблиц. Для того чтобы выделить из состава всех случаев единицы разного состава, показать особенности явлений нужно использовать группировку.

После группировки нужно обобщить данные наблюдения, которые были получены в ходе статистического исследования.

3 этап: проводится анализ полученной при сводке статистической информации на основе применения обобщающих статистических показателей: абсолютных и относительных величин, средних величин, статистических коэффициентов и индексов.

Табличные и графические методы имеют широкое применение при изучении статистической информации.

4. Органы статистики в Российской Федерации

В настоящее время центральным органом единой централизованной системы государственной статистики является Государственный комитет Российской Федерации по статистике (Госкомстат России) – это федеральный орган исполнительной власти.

Единую систему государственной статистики Российской Федерации при Госкомстате России составляют органы в республиках, автономных областях и округах, краях, городах и районах, а также подведомственные им организации и учреждения.

Мощными вычислительными ресурсами обладает Главный межрегиональный центр обработки и распространения статистической информации. Этот центр необходим для обработки поступающих из регионов статистических данных.

Статистическими стандартами Российской Федерации, установленными Госкомстатом России, являются формы и методы сбора и обработки статистических данных, методология расчёта статистических показателей.

Основная деятельность Госкомстата – это разработка федеральных статистических программ, финансируемых из госбюджета.

Основные задачи Госкомстата:

1) координация статистической деятельности в государстве;

2) разработка статистической методологии, которая будет соответствовать потребностям общества на определённом этапе и международным стандартам;

3) предоставление официальной статистической информации Президенту, Правительству, общественности, Федеральному Собранию Российской Федерации, федеральным органам исполнительной власти, международным и иным организациям.

В структуре Госкомстата России для решения задач по сбору, обработке и анализу статистических данных выделены следующие управления:

* статистического планирования и организации статистического наблюдения,

* статистики строительства и основных фондов,

* статистики окружающей среды и сельского хозяйства,

* статистики цен и финансов,

* статистики населения и ряд других по отраслям экономики и социальной сферы.

В соответствии с федеральными программами определённые виды статистических работ ведутся не только Госкомстатом России, но и иными органами государственного управления: Банком России, Министерством образования России, Министерством финансов России, Министерством труда России, Министерством Внутренних Дел и другими (общее количество министерств и ведомств составляет более 200).

«СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ И ЭТАПЫ ЕГО ПРОВЕДЕНИЯ»

1. Понятие о статистическом наблюдении

Любое экономико-статистическое исследование начинается со статистического наблюдения.

Статистическое наблюдение - это предварительная стадия статистического исследования, которая представляет собой планомерный, научно организованный учёт (сбор) первичных статистических данных о массовых социально-экономических явлениях и процессах.

Не всякий сбор данных можно назвать статистическим наблюдением.

Наблюдение будет статистическим, когда:

1) оно сопровождается регистрацией изучаемых фактов в соответствующих учётных документах для дальнейшего их обобщения,

2) носит массовый характер. Это обеспечивает охват значительного числа случаев проявления того или иного процесса, необходимого и достаточного для того, чтобы получить данные, которые касаются не только отдельных единиц совокупности, но и всей совокупности в целом.

Статистическое наблюдение должно отвечать ряду важнейших требований:

а) проводиться непрерывно и систематически;

б) учёт массовых данных должен быть таким, чтобы не только обеспечивалась полнота данных, но и учитывалось их постоянное изменение;

- в) данные должны быть максимально достоверны и точны;
- г) исследуемые явления должны иметь не только научную, но и практическую ценность.

Сбор статистических данных может проводиться как органами государственной статистики, научно-исследовательскими институтами, другими государственными структурами, так и экономическими службами банков, бирж, предприятий, фирм. Только в этом случае исследователи получают достоверную и достаточно разнообразную статистическую информацию, позволяющую всесторонне изучать социально-экономические явления.

2. Этапы проведения статистического наблюдения

Статистическое наблюдение (сбор первичного статистического материала) состоит из трёх основных этапов:

- 1) подготовка статистического наблюдения;
- 2) организация и производство наблюдения;
- 3) контроль полученных первичных данных.

На первом этапе статистического наблюдения определяется цель, устанавливаются объект и единица наблюдения, разрабатываются инструментарий и программа наблюдения.

Общей целью статистического наблюдения является получение достоверной информации о тенденциях развития явлений и процессов для последующего принятия управленческих решений. Она должна быть конкретной и чёткой. Нечётко поставленная цель может привести к сбору не тех данных, которые необходимы для решения конкретной задачи.

Цель определяет объект статистического наблюдения.

Объект наблюдения есть некоторая исследуемая статистическая совокупность или физических лиц (население, работники), или юридических лиц (предприятия, фирмы, учебные заведения), или физических единиц (производственное оборудование, средства передвижения и транспортировки, жилые дома), т.е. исследуемая статистическая совокупность состоит из отдельных единиц.

Единица наблюдения - это первичный элемент объекта статистического наблюдения, который является носителем признаков, подлежащих регистрации.

Указание важнейших признаков позволяет установить границы исследуемой совокупности. Скажем, если необходимо провести исследование рентабельности полиграфических предприятий, то необходимо определить формы собственности этих предприятий, организационно-правовые основы, количество работников предприятия, объём реализации продукции, т.е. то, что отличает как государственные и негосударственные предприятия, так и малые и крупные предприятия. Только в этом случае мы получим достоверную статистическую информацию.

Единицу наблюдения следует отличать от отчётной единицы.

Отчётная единица - такая единица, от которой поступают отчётные данные. Она может совпадать или не совпадать с единицей наблюдения.

Обоснование цели, выбор единиц наблюдения, отчётных единиц, отбор существенных признаков, период времени проведения статистического наблюдения, формы отчётности излагаются в программе статистического наблюдения.

Обычно **программой наблюдения** называют перечень вопросов, которые подлежат регистрации при проведении наблюдения. Чтобы программа наблюдения была научно обоснована и правильно составлена, к ней предъявляются следующие требования:

- 1) чёткая и конкретная формулировка главной цели наблюдения;
- 2) определение места и времени наблюдения, где определяются критический момент (дата или интервал времени, по состоянию на который проводится регистрация признаков) и срок (период заполнения статистического формуляра);
- 3) выделение ряда наиболее существенных признаков объекта наблюдения;
- 4) комплексное определение типа, основных черт и свойств изучаемого явления;
- 5) вопросы, сформулированные в программе, не должны носить двусмысленный характер;
- 6) соблюдение логического принципа последовательности вопросов;
- 7) включение в программу вопросов контрольного характера для проверки собираемых статистических данных;
- 8) сочетание закрытых и открытых вопросов программы.

Программа оформляется в виде документа, так называемого статистического формуляра, который обеспечивает единообразие получаемых сведений от каждой отчётной единицы. Формуляр имеет титульную часть (сведения о тех, кто проводит наблюдение) и адресную часть (адрес и подчинённость отчётной единицы). Программа имеет приложение - инструкцию (инструментарий статистического наблюдения), которая определяет порядок проведения наблюдения и порядок заполнения формы отчётности.

На втором этапе решаются важнейшие организационные вопросы статистического наблюдения. Они заключаются в том, чтобы выбрать соответствующие целям и задачам конкретного статистического наблюдения организационные формы наблюдения, виды наблюдения и способы получения статистической информации.

На третьем этапе собранный статистический материал должен пройти контроль. Как показывает практика, даже при чётко организованном статистическом наблюдении встречаются погрешности и ошибки, которые требуют исправления. Поэтому целью этого этапа является как счётный, так и логический контроль полученных первичных данных. Расхождение между расчётным и действительным значениями исследуемой величины в статистике называют ошибкой наблюдения. В зависимости от причин возникновения различают ошибки регистрации и ошибки репрезентативности.

Ошибки регистрации могут быть случайными и систематическими.

Случайные ошибки не имеют определённой направленности и возникают под действием случайных факторов (перестановка цифр, смещение строк и граф при заполнении статистического формуляра). При обобщении массового материала эти ошибки взаимопогашаются.

Систематические ошибки регистрации имеют определённую направленность, могут либо завышать, либо занижать конкретное значение показателя, что в итоге приводит к искажению действительного положения.

Примерами систематической статистической ошибки при регистрации служат округление возраста населения на цифрах, заканчивающихся на 5 и 0, преуменьшение доходов в документации для налоговых органов, элементы недостоверности, которые вносят предприятия в те характеристики, от которых зависит расчёт с кредиторами, и т.д.

Для выявления ошибок используется счётный контроль, особенно для проверки итоговых сумм. Помимо счётного используется и логический контроль, который может поставить под сомнение правильность полученных данных, поскольку основан на логической взаимосвязи между признаками.

Например, при переписи населения полученный факт, что пятилетний ребенок имеет среднее образование, ставится под сомнение и в этом случае ясно, что при заполнении формуляра допущена ошибка.

Если ошибки регистрации свойственны любому наблюдению (сплошному и несплошному), то **ошибки репрезентативности** - только несплошному наблюдению. Они характеризуют расхождения между значениями показателя, полученного в обследуемой совокупности, и его значением по исходной (генеральной) совокупности.

Ошибки репрезентативности также могут быть случайными и систематическими. Случайные ошибки возникают, если отобранная совокупность не полностью воспроизводит все признаки генеральной совокупности и величину этих ошибок можно оценить. Систематические ошибки репрезентативности могут возникать, если нарушен сам принцип отбора единиц из исходной совокупности. В этом случае проводятся проверка полноты собранных данных, арифметический контроль точности информации на предмет её достоверности, проверка логической взаимосвязи показателей.

Контрольной проверкой собранных данных завершается статистическое наблюдение.

«ФОРМЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ. СТАТИСТИЧЕСКАЯ ОТЧЕТНОСТЬ И ЕЕ ВИДЫ»

1. Формы статистического наблюдения

В отечественной статистике используются три организационные формы (типы) статистического наблюдения (рис. 1).

Основной формой статистического наблюдения является отчётность. Если первичный учёт (первичный учётный документ) регистрирует различные факты, то отчётность является обобщением первичного учёта.

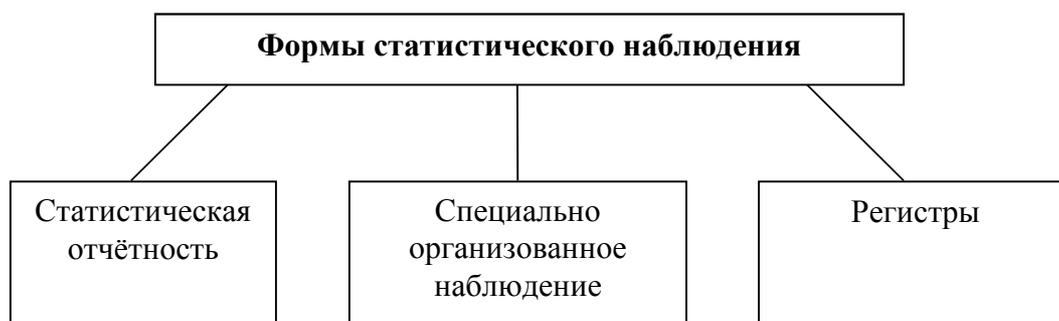


Рис. 1 Классификация форм статистического наблюдения

Отчётность - официальный документ, который скрепляется подписями лиц, ответственных за предоставление и достоверность собранных сведений, и утверждается органами государственной статистики.

Кроме годовой может иметь место ежедневная, недельная, двухнедельная, месячная и квартальная отчётность. Отчётность может быть представлена по почте, телеграфу, телетайпу, факсу.

К **специально организованному статистическому наблюдению** можно отнести перепись. На практике проводится перепись населения, материальных ресурсов, зелёных насаждений, незавершенных строительных объектов, оборудования и т.д.

Перепись - это специально проводимые широкомасштабные работы по сбору необходимой статистической информации об изучаемых объектах в границах отрасли, региона или страны в целом, повторяющееся через равные промежутки времени, задачей которого является не только определение численности и состава исследуемой совокупности, но и анализ количественных изменений в период между двумя обследованиями.

Из всех переписей наиболее известны переписи населения. Цель последних состоит в установлении численности и размещения населения по территории страны, в получении характеристики состава населения по полу, возрасту, занятиям и другим показателям.

В период подготовки всеобщей переписи для уточнения и апробации программно-методических и организационных вопросов наблюдения проводят пробную перепись.

Пробная перепись населения является неотъемлемым этапом подготовки всеобщей переписи населения.

Примером также могут служить переписи крупного рогатого скота, которые проводятся в конце календарного года и позволяют получить информацию о численности и структуре поголовья крупного рогатого скота у сельхозпроизводителей.

Органами статистики также проводятся переписи многолетних насаждений, жилого фонда, незавершенного строительства и пр.

Кроме переписей, к специально организованному наблюдению также относятся и другие единовременные работы по сбору необходимой статистиче-

ской информации, в частности, в рамках социологических или маркетинговых исследований.

Формой непрерывного статистического наблюдения является **регистрационное наблюдение (регистр)**, объектами которого являются долговременные процессы, имеющие фиксированное начало, стадию развития и фиксированное время завершения.

Регистр основан на системе отслеживания состояния переменных и постоянных показателей. В статистической практике различают регистры населения и регистры предприятий.

В настоящее время в России существует Единый государственный регистр предприятий всех форм собственности (ЕГРПО), информационный фонд, которого содержит: регистровый код, сведения о территориальной и отраслевой принадлежности, форме подчинённости, виде собственности, справочные сведения и экономические показатели (среднесписочная численность работников; средства, направляемые на потребление; остаточная стоимость основных средств; балансовая прибыль или убыток; уставный фонд). При закрытии предприятия в десятидневный срок ликвидационная комиссия информирует об этом службу ведения регистра.

2. Способы статистического наблюдения

Статистическая информация может быть получена различными способами, важнейшими из которых являются непосредственное наблюдение, документальный учёт фактов и опрос (рис. 2):



Рис. 2 Способы статистического наблюдения

Непосредственным называют такое наблюдение, при котором сами регистраторы путём непосредственного замера, взвешивания или подсчёта устанавли-

ливают значение признака и на этом основании производят запись в формуляре наблюдения. Этим способом проводится инвентаризации основных средств на предприятиях.

Документальное наблюдение предполагает запись ответов на вопросы формуляра на основании соответствующих документов. Примером такого наблюдения является сбор данных об успеваемости студентов на основе зачётно-экзаменационных ведомостей, заполнение форм статистической отчётности на основании данных бухгалтерского учёта и т.п.

Опрос - это наблюдение, при котором ответы на вопросы формуляра записываются со слов опрашиваемого (респондента). Этим способом проводятся переписи населения, опросы общественного мнения.

Его разновидности: устный (экспедиционный), анкетный, корреспондентский, явочный опрос и саморегистрация.

Устный опрос может быть как прямым (непосредственное общение счётчика с респондентом), так опосредованным (например, по телефону).

При анкетном способе определённое число респондентов получают специальные вопросники либо лично, либо через средства печати. Данный вид опроса применяется в исследованиях, где нужны ориентировочные результаты, не претендующие на высокую точность (изучение общественного мнения).

Явочный способ используется в сплошном наблюдении, когда необходимо личное присутствие (регистрация браков, разводов, рождений и т.д.).

При корреспондентском способе сведения сообщаются штатом добровольных корреспондентов, в силу чего полученный материал не всегда носит качественный характер.

При способе саморегистрации формуляры заполняются самими респондентами, а счётчики консультируют и собирают формуляры. В статистической практике различные виды статистических наблюдений могут сочетаться, дополняя друг друга.

3. Виды статистического наблюдения

Статистическое наблюдение можно разбить по следующим признакам (рис. 3):

- по времени регистрации фактов;
- охвату единиц совокупности.

По времени регистрации фактов наблюдение бывает непрерывным (текущим), периодическим и единовременным.

Непрерывное (текущее) статистическое наблюдение - это систематическая регистрация фактов или явлений по мере их поступления с целью изучения их динамики. Например, регистрации актов гражданского состояния (рождения, браки, смерти), регистрация страховыми компаниями всех несчастных случаев и других неблагоприятных событий по мере их возникновения.

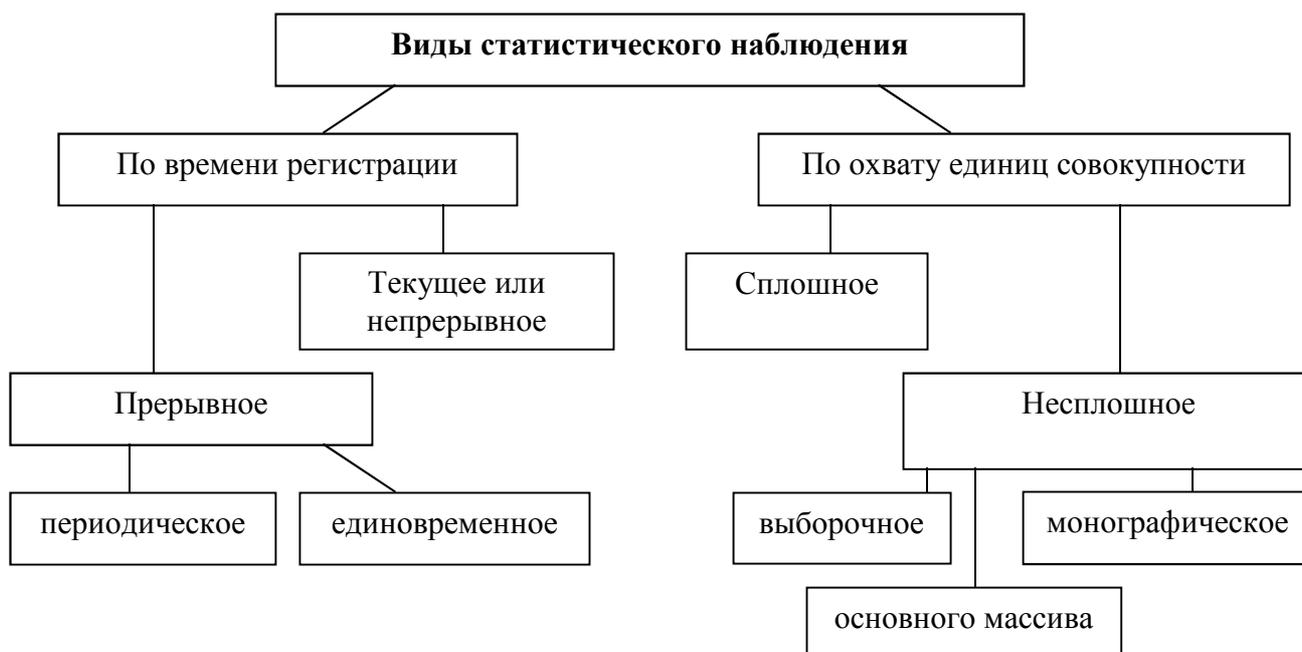


Рис. 3 Классификация видов статистического наблюдения

Видами прерывного наблюдения являются единовременное и периодическое.

Единовременное наблюдение - разовое сплошное наблюдение для сбора количественных характеристик явления или процесса в момент его исследования.

Периодическое наблюдение проводится через определённые промежутки времени по схожим программе и инструментарию. Например, периодическое исследование пассажиропотоков в общественном транспорте, периодическая регистрация цен производителей по отдельным товарам (один раз в месяц или в квартал).

По охвату единиц совокупности статистическое наблюдение бывает сплошным и несплошным.

Сплошное наблюдение охватывает все единицы исследуемой совокупности (например, общая перепись населения).

В свою очередь, **несплошное наблюдение** охватывает только часть исследуемой совокупности.

В зависимости от того, как выбрана эта часть, несплошное наблюдение можно подразделить на:

- выборочное (основано на принципе случайного отбора);
- метод основного массива (исследуются самые существенные или наиболее крупные единицы изучаемой совокупности).;
- монографическое наблюдение (подробное исследование отдельных единиц изучаемой совокупности с целью выявления намечающихся тенденций).

«СТАТИСТИЧЕСКАЯ СВОДКА»

1. Точность статистического исследования

Под точностью статистического наблюдения понимают степень соответствия значения наблюдаемого показателя, вычисленного по материалам обследования, его действительной величине. Расхождение, или разница, между ними называется ошибкой статистического наблюдения.

Различают две группы ошибок:

- 1) ошибки регистрации;
- 2) ошибки репрезентативности.

Ошибки регистрации присущи любому статистическому наблюдению, как сплошному, так и несплошному. Они делятся на ошибки регистрации и систематические ошибки регистрации. Случайными ошибками регистрации называют ошибки, возникающие вследствие действия случайных факторов. К ним можно различного рода непреднамеренные описки. Систематические ошибки регистрации чаще всего имеют искажения: они либо увеличивают, либо уменьшают статистический показатель, и, что характерно, подобная ситуация повторяется от обследования к обследованию. Например: человек округляет свой возраст — вместо 32 лет говорит вместо 79 — 80 и т.п. (это явление широко известно и даже получило название — «аккумуляция возрастов»).

Ошибки репрезентативности присущи только несплошному обследованию. Они также делятся на случайные и систематические ошибки.

Случайные ошибки репрезентативности возникают из-за того, обследованию подвергается не вся совокупность в целом, а только часть, и, следовательно, при несплошном наблюдении они присутствуют всегда. В теории статистики разработаны специальные методы оценки величин таких ошибок, на их основе для наблюдаемых показателей строят доверительные интервалы, т.е. эти ошибки вычисляются и находятся как бы «под контролем». Хуже обстоит дело, если наряду со случайными ошибками имеются и ошибки систематические

Систематические ошибки репрезентативности возникают, если при несплошном наблюдении кардинально нарушаются технологии отбора единиц из генеральной совокупности объектов, но чаще - если в ходе обследования не удается получить информацию обо всех отобранных для наблюдения единицах, например, вследствие отказа отвечать на вопросы анкеты, или если человека не удалось застать дома.

Для повышения точности наблюдения необходимо:

- 1) правильно разработать формуляр статистического наблюдения: вопросы должны быть четкими, однозначными, не допускающими двойного толкования;
- 2) иметь хорошо обученный персонал для проведения обследования;
- 3) строго придерживаться выбранной технологии обследования.

2. Задачи и виды статистической сводки

На основе информации, собранной в ходе статистического наблюдения нельзя, как правило, выявить и охарактеризовать закономерности социально-экономических явлений. Это связано с тем, что наблюдение даёт сведения по каждой единице исследуемого объекта. Полученные данные не являются обобщающими показателями. С их помощью нельзя сделать выводы в целом об объекте без предварительной обработки данных.

Важнейшим этапом исследования социально-экономических явлений и процессов является систематизация первичных данных и получение на этой основе сводной характеристики всего объекта при помощи обобщающих показателей, что достигается путём сводки и группировки первичного статистического материала.

Сводка – это научная обработка первичных данных с целью получения обобщённых характеристик изучаемого социально-экономического явления по ряду существенных для него признаков с целью выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

Сводка может быть различной в зависимости от ряда характеризующих её признаков (рис. 4):

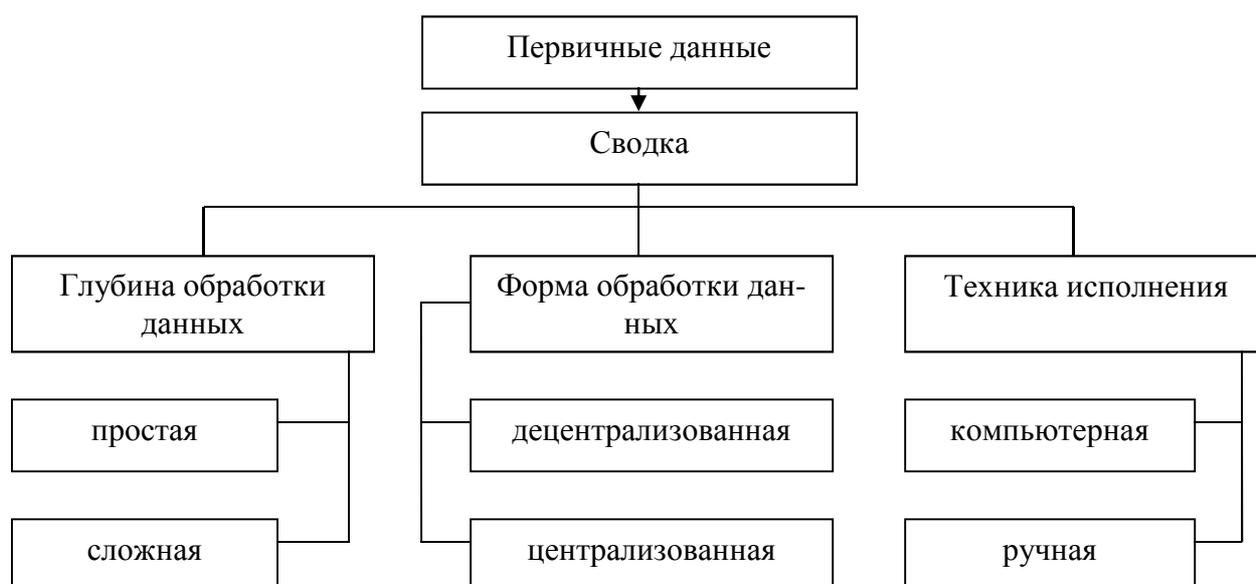


Рис. 4. Классификация видов статистической сводки

По глубине и точности обработки материала различают сводку простую и сложную.

Простая сводка – это операция по подсчёту общих итогов по совокупности единиц наблюдения и оформление этого материала в статистических таблицах.

Сложная сводка – это комплекс последовательных операций, включающих группировку полученных при наблюдении материалов, составление системы показателей для характеристики типичных групп и подгрупп изучаемой со-

вокупности явлений, подсчёт числа единиц и итогов по каждой группе и подгруппе, и по всему объекту и представление результатов в виде статистических таблиц.

Проведению сводки предшествует разработка её программы, которая включает следующие этапы:

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

По форме обработки материала сводка бывает:

- централизованная, когда весь первичный материал поступает в одну организацию, подвергается в ней обработке от начала до конца;
- децентрализованная, когда отчёты предприятий сводятся статистическими органами субъектов РФ, а полученные итоги поступают в Госкомстат РФ и там определяются итоговые показатели в целом по народному хозяйству страны.

По технике исполнения сводка может быть компьютерной и ручной.

Компьютерная сводка – это способ выполнения сводки статистических данных, при котором все операции осуществляются с помощью использования компьютеров и программных продуктов, позволяющих обработать любые объёмы информации в различной степени детализации.

При ручной сводке все основные операции (подсчёт групповых и общих итогов) осуществляются вручную.

Для проведения сводки составляется план, в котором излагаются организационные вопросы: кем и когда будут осуществляться все операции, порядок её проведения, состав сведений, подлежащих опубликованию в периодической печати.

Тема урока: «**МЕТОД ГРУППИРОВОК В СТАТИСТИКЕ**»

1. Виды статистических группировок

Группировкой называется разбиение общей совокупности единиц объекта наблюдения по одному или нескольким существенным признакам на однородные группы, различающиеся между собой в количественном и качественном отношении и позволяющие выделить социально-экономические типы, изучить структуру совокупности и проанализировать связи между отдельными признаками.

Группировки являются важнейшим статистическим методом обобщения статистических данных, основой для правильного исчисления статистических показателей.

С помощью метода группировок решаются следующие задачи:

- выделение социально-экономических типов явлений;

- изучение структуры явления и структурных сдвигов, происходящих в нем;
- выявление взаимосвязи и взаимозависимости между явлениями.

В соответствии с познавательными задачами, решаемыми в ходе построения статистических группировок, различают следующие их виды: типологические, структурные, аналитические.

Типологическая группировка – это разбиение разнородной совокупности единиц наблюдения на отдельные качественно однородные группы и выявление на этой основе социально-экономических типов явлений. При построении группировки этого вида основное внимание должно быть уделено идентификации типов и выбору группировочного признака. Решение вопроса об основании группировки должно осуществляться на основе анализа сущности изучаемого социально-экономического явления.

Примером типологической группировки по атрибутивному признаку является группировка предприятий и организаций по формам собственности (табл. 1).

Данные табл. 1 показывают, что наибольший удельный вес принадлежит предприятиям и организациям с частной формой собственности, а наименьший – прочим формам собственности.

Структурная группировка предназначена для изучения состава однородной совокупности по какому-либо варьирующему признаку, а также структуры и структурных сдвигов, происходящих в нем.

Примером структурной группировки по количественному признаку является группировка населения по возрастным группам (табл. 2).

Таблица 1

Группировка предприятий и организаций по формам собственности за 2011 г.

№ п/п	Группы предприятий по формам собственности	Число предприятий	
		всего, тыс.	% к итогу
	А	1	2
1	Государственная	1645	3,1
2	Муниципальная	4994	9,4
3	Частная	41713	78,9
4	Собственность общественных и религиозных организаций (объединений)	3339	6,3
5	Прочие формы собственности, включая смешанную российскую, иностранную, совместную российскую и иностранную	1203	2,3
	Всего	52894	100,0

Группировка постоянного населения по возрастным группам
(на начало года)

Возрастные группы населения	Тысяч человек		В % к итогу	
	2009 г.	2010 г.	2009 г.	2010 г.
А	1	2	3	4
Всё население	2589,9	2610,2	100,0	100,0
в том числе в возрасте, лет				
0 – 4	135,9	140,7	5,2	5,4
5 – 9	118,3	123,1	4,6	4,7
10 – 19	271,1	275,0	10,5	10,5
20 – 29	435,1	430,7	16,8	16,5
30 – 39	371,5	375,1	14,3	14,4
40 – 49	367,5	357,3	14,2	13,7
50 – 59	380,0	386,8	14,7	14,8
60 – 69	225,2	230,9	8,7	8,8
70 лет и старше	285,3	289,9	11,0	11,1

Численность населения (табл. 2) за 12 месяцев увеличилась. При этом возросло число населения моложе и старше трудоспособного возраста. Наибольший удельный вес приходится на население в трудоспособном возрасте 60,5%, что на 0,5 % ниже начала года.

Аналитическая группировка - выявляет взаимосвязи между изучаемыми явлениями и признаками, их характеризующими.

В статистике при изучении связей социально-экономических явлений признаки делятся на факторные и результативные.

Факторными называются признаки, под воздействием которых изменяются другие результативные признаки. Взаимосвязь проявляется в том, что с возрастанием или убыванием значения факторного признака систематически возрастает или убывает значение признака результативного и наоборот.

Особенностями построения аналитической группировки являются:

1) единицы статистической совокупности группируются по факторному признаку;

2) каждая выделенная группа характеризуется средними величинами результативного признака.

Пример аналитической группировки представлен в табл. 3.

Данные табл. 3. позволяют выявить зависимость между размером процентной ставки и суммы выданного кредита. Чем выше ставка, тем меньше сумма кредита, приходящаяся в среднем на один банк и чем меньше ставка, тем выше размер кредита.

Таблица 3

Группировка зависимости суммы выданного кредита коммерческими банками от размера процентной ставки (цифры условные)

№ группы	Размер процентной ставки	Число банков	Сумма выданного кредита, млн. руб.	
			всего	в среднем на один банк
	А	1	2	3
1	11 – 14	4	103,5	25,88
2	14 – 17	7	134,8	19,26
3	17 – 20	10	127,3	12,73
4	20 – 23	6	42,8	7,13
5	23 – 26	3	6,8	2,27
	Итого	30	415,2	13,84

По способу построения группировки бывают простые и комбинационные.

Простая - группировка, в которой группы образованы только по одному признаку. Среди простых особо выделяют ряды распределения.

Комбинационная - группировка, в которой разбиение совокупности на группы производится по двум и более признакам, взятым в сочетании (комбинации).

Сначала группы формируются по одному признаку, затем группы делятся на подгруппы по другому признаку, а эти в свою очередь делятся по третьему и так далее. Таким образом, комбинационные группировки дают возможность изучить единицы совокупности одновременно по нескольким взаимосвязанным признакам.

При построении комбинационной группировки возникает вопрос о последовательности разбиения единиц объекта по признакам. Как правило, рекомендуется сначала производить группировку по атрибутивным признакам, значения которых имеют ярко выраженные качественные различия.

По упорядоченности исходных данных группировки бывают первичные и вторичные.

Процесс образования новых групп на основе группировки, произведённой по первичным данным, называется вторичной группировкой.

Необходимость во вторичной группировке возникает в случаях:

1) когда в результате первоначальной группировки нечётко проявился характер распределения изучаемой совокупности (в этом случае производят укрупнение или уменьшение интервалов);

2) когда требуется сопоставить между собой данные, имеющие различное число выделенных групп или неодинаковые границы интервалов.

2. Принципы построения статистических группировок

Построение статистических группировок осуществляется по следующим этапам:

1. Определение группировочного признака.
2. Определение числа групп.
3. Расчёт ширины интервала группировки.
4. Определение признаков, которые в комбинации друг с другом будут характеризовать каждую выделенную группу.

Построение группировки начинается с определения группировочного признака.

Группировочным признаком называется признак, по которому проводится разбиение единиц совокупности на отдельные группы. От правильного выбора группировочного признака зависят выводы статистического исследования. В качестве основания группировки необходимо использовать существенные, теоретически обоснованные признаки.

В основание группировки могут быть положены как количественные, так и качественные признаки. **Количественные признаки** – это признаки, которые имеют числовое выражение (объём выпускаемой продукции, возраст человека, доход сотрудника фирмы и т. д.). **Качественные признаки** отражают состояние единицы совокупности (пол, отраслевая принадлежность предприятия, форма собственности фирмы и т.д.).

После того, как определено основание группировки, следует решить вопрос о количестве групп, на которые необходимо разбить исследуемую совокупность единиц наблюдения.

Число групп зависит от задач исследования и вида показателя, положенного в основание группировки, объёма изучаемой совокупности и степени вариации признака. Вид показателя особенно существенен при анализе качественных признаков. Так, например, группировка сотрудников фирмы по полу учитывает только две градации: «мужской» и «женский».

В случае группировки единиц наблюдения по количественному признаку особое внимание необходимо обратить на число единиц исследуемого объекта, объём совокупности и степень колеблемости группировочного признака.

При небольшом объёме совокупности ($n < 50$) не следует образовывать большого количества групп, так как группы будут включать недостаточное число единиц объекта.

Показатели, рассчитанные для таких групп, не будут представительными и не позволят получить адекватную характеристику исследуемого явления.

Часто группировка по количественному признаку имеет задачу отразить распределение единиц совокупности по этому признаку. В этом случае количество групп зависит, в первую очередь, от степени колеблемости группировочного признака: чем больше его колеблемость, тем больше можно образовать групп. Поэтому при определении числа групп необходимо принять во внимание размах вариации признака (R), который позволяет оценить вариацию признака

между крайними значениями признака – максимальным (X_{\max}) и минимальным (X_{\min}) и определяется по следующей формуле:

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

Чем больше размах вариации признака, положенного в основание группировки, тем, как правило, может быть образовано большее число групп. При этом может возникнуть проблема получения пустых групп, т.е. групп, не содержащих ни одной единицы наблюдения.

Построение большого числа групп позволит, с одной стороны, точнее воспроизвести характер исследуемого объекта. Однако, с другой стороны, слишком большое число групп затрудняет выявление закономерностей при исследовании социально-экономических явлений и процессов. Поэтому в каждом конкретном случае при определении числа групп следует исходить не только из степени колеблемости признака, но и из особенностей объекта и показателей, его характеризующих, а также цели исследования.

Определение числа групп можно осуществить несколькими способами. Формально-математический способ предполагает использование формулы Стерджесса:

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

где: n – число групп;

N – число единиц совокупности.

Согласно этой формуле выбор числа групп зависит только от объёма изучаемой совокупности. Когда определено число групп, то следует определить интервалы группировки.

Интервал – это значения варьирующего признака, лежащие в определённых границах. Каждый интервал имеет верхнюю и нижнюю границы или одну из них. **Нижней границей** интервала называется наименьшее значение признака в интервале. **Верхней границей** интервала называется наибольшее значение признака в интервале. **Величина интервала** - разность между верхней и нижней границами интервала.

Интервалы группировки бывают:

- равные и неравные;
- открытые и закрытые.

В зависимости от величины интервалы группировки бывают: равные и неравные. В свою очередь, неравные интервалы подразделяются на прогрессивно возрастающие, прогрессивно убывающие, произвольные и специализированные.

Равные интервалы применяются в случае, если изменение количественного признака внутри изучаемой совокупности единиц наблюдения происходит равномерно и его вариация проявляется в сравнительно узких границах.

Ширина равного интервала определяется по следующей формуле:

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

где: X_{\max} , X_{\min} - максимальное и минимальное значения признака в совокупности;

n - число групп.

Если максимальные или минимальные значения сильно отличаются от смежных с ними значений вариантов в упорядоченном ряду значений группировочного признака, то для определения величины интервала следует использовать не максимальное или минимальное значения, а значения, несколько превышающие минимум, и несколько меньше, чем максимум.

Полученную по формуле величину округляют, и она будет являться шириной интервала.

Существуют следующие **правила определения ширины интервала**.

Если величина интервала, рассчитанная по формуле представляет собой величину, которая имеет один знак до запятой (например: 0,67; 1,487; 3,82), то в этом случае полученные значения целесообразно округлить до десятых и их использовать в качестве ширины интервала.

Если рассчитанная величина интервала имеет две значащие цифры до запятой и несколько после запятой (например 14,876), то это значение необходимо округлить до целого числа (до 15).

В случае, когда рассчитанная величина интервала представляет собой трехзначное, четырехзначное и так далее число, то эту величину следует округлить до ближайшего числа, кратного 100 или 50. Например, 652 следует округлить до 650 или до 700.

Если размах вариации признака в совокупности велик и значения признака варьируют неравномерно, то надо использовать группировку с неравными интервалами. Неравные интервалы могут быть получены в процессе объединения пустых, не содержащих ни одной единицы совокупности, равных интервалов. Это происходит в том случае, если после построения равных интервалов по изучаемому признаку образуются группы, содержащие мало или не содержащие вообще ни одной единицы, т.е. группы, не отражающие определённых типов изучаемого явления по признаку. В этом случае возникает необходимость в увеличении интервалов группировки.

Также неравные интервалы могут быть прогрессивно-возрастающие или прогрессивно-убывающие в арифметической или геометрической прогрессии. Величина интервалов, изменяющихся в арифметической и геометрической прогрессии, определяется следующим образом:

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

а в геометрической прогрессии:

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

где: q – константа: для прогрессивно-возрастающих интервалов имеет знак «+», а при прогрессивно-убывающих – знак «-».

q – константа: для прогрессивно-возрастающих – больше «1»; для прогрессивно-убывающих – меньше «1».

Применение неравных интервалов обусловлено тем, что в первых группах небольшая разница в показателях имеет большое значение, а в последних группах эта разница не существенна.

Например, при построении группировки строительных компаний города, по показателю численности работающих, который варьирует от 500 человек до 3500 человек, нецелесообразно рассматривать равные интервалы, т. к. учитываются как малые, так и крупнейшие строительные фирмы города. Поэтому следует образовывать неравные интервалы: 500-1000, 1000-2000, 2000-3500, т. е. величина каждого последующего интервала больше предыдущего на 500 человек и увеличивается в арифметической прогрессии.

Выбор исследователя в построении равных или неравных интервалов зависит от степени заполнения каждой выделенной группы, т.е. от числа единиц в них. Если величина интервала существенна и содержит большое число единиц совокупности, то эти интервалы необходимо дробить, а в противном случае – объединять.

Интервалы группировок могут быть закрытыми и открытыми.

Закрытыми называются интервалы, у которых имеются обе границы: верхняя и нижняя границы.

Открытые – это интервалы, у которых указана только одна граница: как правило, верхняя – у первого интервала и нижняя – у последнего. Например, группы страховых компаний по числу работающих в них сотрудников (чел.): до 50, 50-100, 100-150, 150 и более.

Применение открытых интервалов целесообразно в тех случаях, когда в совокупности встречается незначительное число единиц наблюдения с очень малыми или очень большими значениями вариантов, которые резко, в несколько раз, отличаются от всех остальных значений изучаемого признака.

При группировке единиц совокупности по количественному признаку границы интервалов могут быть обозначены по-разному, в зависимости от того, непрерывный или дискретный признак положен в основание группировки.

Если основанием группировки служит непрерывный признак (например, группы строительных фирм по объёму строительно-монтажных работ, выполненных собственными силами (тыс. руб.): 1200-1400, 1400-1600, 1600-1800, 1800-2000), то одно и то же значение признака выступает и верхней и нижней границами двух смежных интервалов. В данном случае объём работ 1400 тыс. руб. составляет верхнюю границу первого интервала и нижнюю границу

второго, 1600 тыс. руб. – соответственно второго и третьего и т.д., т.е. верхняя граница i – го интервала равна нижней границе $(i+1)$ – го интервала.

При таком обозначении границ может возникнуть вопрос, в какую группу включать единицы наблюдения, значения признака у которых совпадают с границами интервалов.

Например, во вторую или третью группу должна войти строительная фирма с объёмом строительно-монтажных работ 1600 тыс. рублей? Если верхняя граница формируется по принципу «исключительно», то фирма должна быть отнесена к третьей группе, в противном случае – ко второй.

Для того чтобы правильно отнести к той или иной группе единицу совокупности, значение признака которой совпадает с границами интервалов, можно ориентироваться на открытые интервалы (по нашему примеру группы строительных фирм по объёму строительно-монтажных работ преобразуются в следующие: до 1400, 1400-1600, 1600-1800, 1800 и более). В данном случае, вопрос отнесения отдельных единиц совокупности, значения которых являются граничными, к той или иной группе решается на основе анализа последнего открытого интервала.

Возможны два случая обозначения последнего открытого интервала: 1) 1800 тыс. руб. и более; 2) более 1800 тыс. руб. В первом случае, строительные фирмы с объёмом строительно-монтажных работ 1600 тыс. руб. попадут в третью группу; во втором случае – во вторую группу.

Если в основании группировки лежит дискретный признак, то нижняя граница i -го интервала равна верхней границе $i-1$ -го интервала, увеличенной на 1.

Например, группы строительных фирм по числу занятого персонала (чел.) будут иметь вид: 100-150, 151-200, 201-300.

При определении границ интервалов статистических группировок иногда исходят из того, что изменение количественного признака приводит к появлению нового качества.

В этом случае граница интервала устанавливается там, где происходит переход от одного качества к другому.

Строя такую группировку, следует дифференцированно устанавливать границы интервалов для разных отраслей народного хозяйства. Это достигается путём использования группировок со специализированными интервалами. **Специализированные интервалы** – это такие интервалы, которые применяются для выделения из совокупности одних и тех же типов по одному и тому же признаку для явлений, находящихся в различных условиях.

При изучении социально-экономических явлений на макроуровне часто применяют группировки, интервалы которых не будут ни прогрессивно-возрастающими, ни прогрессивно-убывающими. Такие интервалы называются произвольными и, как правило, используются при группировке предприятий, например, по уровню рентабельности.

«РЯДЫ РАСПРЕДЕЛЕНИЯ В СТАТИСТИКЕ»

1. Ряды распределения

Ряды распределения представляют собой простейшую группировку, в которой каждая выделенная группа характеризуется только частотой.

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

Атрибутивными называют ряды распределения, построенные по качественным признакам, то есть признакам, характеризующим состояние изучаемого явления и не имеющим числового выражения.

Атрибутивные ряды распределения характеризуют состав совокупности по тем или иным существенным признакам. Взятые за несколько периодов, эти данные позволяют исследовать изменение структуры.

Вариационными рядами называют ряды распределения, построенные по количественному признаку, т.е. признаку, имеющему числовое выражение у отдельных единиц совокупности.

Вариационный ряд состоит из двух элементов: вариантов и частот. Вариантами называются отдельные значения признака, которые он принимает в вариационном ряду, то есть конкретное значение варьирующего признака. Частотами называются численности отдельных вариантов или каждой группы вариационного ряда. Частоты показывают, как часто встречаются те или иные значения признака в изучаемой совокупности.

Сумма всех частот определяет численность всей совокупности, её объём. Частостями называются частоты, выраженные в долях единицы или в процентах к итогу. Соответственно сумма частостей равна 1 или 100%.

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

Дискретный вариационный ряд – это ряд распределения в котором группы составлены по признаку, изменяющемуся прерывно, т.е. через определенное число единиц и характеризуют распределение единиц совокупности по дискретному признаку, принимающему только целые значения. Например, группы студентов по баллу в сессию по предмету: 5,4,3,2.

Интервальный вариационный ряд распределения – это ряд распределения, в котором группировочный признак, составляющий основание группировки, может принимать в интервале любые значения, отличающиеся друг от друга на сколь угодно малую величину.

Построение интервальных вариационных рядов целесообразно, прежде всего, при непрерывной вариации признака, а также если дискретная вариация признака проявляется в широких пределах, то есть число вариантов дискретного признака достаточно велико.

Правила построения рядов распределения аналогичны правилам построения группировки.

2. Графическое изображение рядов распределения

Анализ рядов распределения наглядно можно проводить на основе их графического изображения. Для этой цели строят полигон, гистограмму, огиву и кумуляту распределения.

Полигон используется при изображении дискретных вариационных рядов. Для его построения в прямоугольной системе координат по оси абсцисс в одинаковом масштабе откладываются ранжированные значения варьирующего признака, а по оси ординат наносится шкала для выражения величины частот. Полученные на пересечении оси абсцисс (x) и оси ординат (y) точки соединяются прямыми линиями, в результате чего получают ломаную линию, называемую полигоном частот. Иногда для замыкания полигона предлагается крайние точки (слева и справа на ломаной линии) соединить с точками на оси абсцисс, в результате чего получается многоугольник.

Например, изобразим графически распределение жилого фонда по типу квартир, представленных в табл. 4.

Гистограмма применяется для изображения интервального вариационного ряда.

При построении гистограммы на оси абсцисс откладываются величины интервалов, а частоты изображаются прямоугольниками, построенным на соответствующих интервалах.

Таблица 4

Распределение жилого фонда городского района по типу квартир
(цифры условные)

№ п/п	Группы квартир по числу комнат	Число квартир, тыс. ед.
1	1	10
2	2	35
3	3	30
4	4	15
5	5	5
	Всего:	95



Рис. 5. Полигон распределения жилого фонда

Высота столбиков должна быть пропорциональна частотам. В результате получается график, на котором ряд распределения изображен в виде смежных друг с другом столбиков.

Гистограмма может быть преобразована в полигон распределения, если середины верхних сторон прямоугольников соединить прямыми линиями.

При построении гистограммы распределения вариационного ряда с неравными интервалами по оси ординат наносят не частоты, а плотность распределения признака в соответствующих интервалах. Это необходимо сделать для устранения влияния величины интервала на распределение интервала и получения возможности сравнивать частоты.

Плотность распределения – это частота, рассчитанная на единицу ширины интервала, то есть, сколько единиц в каждой группе приходится на единицу величины интервала.

Изобразим графически интервальный ряд распределения, приведенный в табл. 5.

Для графического изображения вариационных рядов может использоваться кумулятивная кривая. При помощи **кумуляты** изображается ряд накопленных частот.

Таблица 5

Распределение семей по размеру жилой площади, приходящейся на одного человека (цифры условные)

№ п/п	Группы семей по размеру жилой площади, приходящейся на одного человека, м ²	Число семей с данным размером жилой площади	Накопленное число семей
1	3 – 5	10	10
2	5 – 7	20	30
3	7 – 9	40	70
4	9- 11	30	100
5	11 – 13	15	115
	Всего:	115	

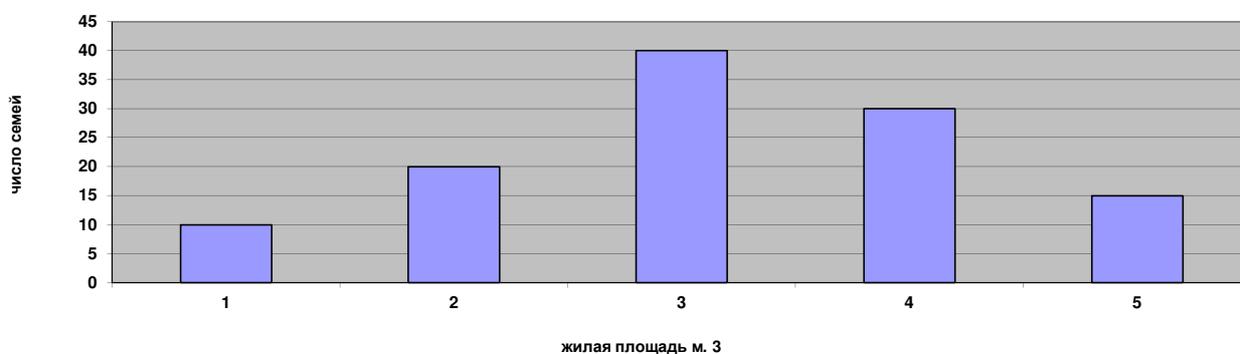


Рис. 6. Гистограмма распределения семей по размеру жилой площади

Накопленные частоты определяются путем последовательного суммирования частот по группам. Накопленные частоты показывают, сколько единиц

совокупности имеют значения признака не больше, чем рассматриваемое значение.

При построении кумуляты интервального вариационного ряда по оси абсцисс (x) откладываются варианты ряда, а по оси ординат (y) накопленные частоты, которые наносят на поле графика в виде перпендикуляров к оси абсцисс в верхних границах интервалов. Затем эти перпендикуляры соединяют и получают ломаную линию, то есть кумуляту.

Используя данные накопленного ряда (табл. 5), построим кумуляту распределения (рис. 7).

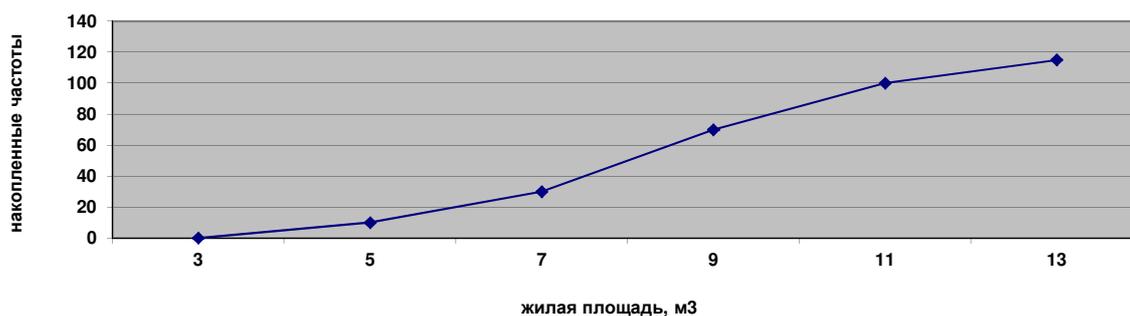


Рис. 2.7. Кумулята распределения семей по размеру жилой площади

Если при графическом изображении вариационного ряда в виде кумуляты оси x и y поменять местами, то получим **огиву**.

«СПОСОБЫ НАГЛЯДНОГО ПРЕДСТАВЛЕНИЯ СТАТИСТИЧЕСКИХ ДАННЫХ»

1. Статистическая таблица и её элементы

Результаты сводки и группировки материалов статистического наблюдения, как правило, представляются в виде таблиц. Таблица является наиболее рациональной, наглядной и компактной формой представления статистического материала. Однако не всякая таблица является статистической. Таблица умножения, опросный лист социологического обследования и так далее могут носить табличную форму, но еще не являются статистическими таблицами.

Статистическая таблица - это цифровое выражение итоговой характеристики всей наблюдаемой совокупности или её составных частей по одному или нескольким существенным признакам.

Основные элементы статистической таблицы, составляющие как бы её остов (основу), показаны на рис. 8.

Название таблицы (общий заголовок)

Содержание строк	Наименование граф (верхние заголовки)					
А	1	2	3	3	5	...
Наименование строк (боковые заголовки)						
Итоговая строка						Итоговая графа

* примечания к таблице

Рис. 8 Остов (основа) статистической таблицы

Статистическая таблица содержит три вида заголовков: общий, верхние и боковые. Общий заголовок отражает содержание всей таблицы (к какому месту и времени она относится), располагается над макетом таблицы по центру и является внешним заголовком. Верхние заголовки характеризуют содержание граф (заголовки сказуемого), а боковые (заголовки подлежащего) – строк. Они являются внутренними заголовками.

Остов таблицы, заполненный заголовками, образует макет таблицы; если на пересечении граф и строк записать цифры, то получается полная статистическая таблица.

Цифровой материал может быть представлен абсолютными (численность населения РФ), относительными (индексы цен на продовольственные товары) и средними (среднемесячный доход сотрудника коммерческого банка) величинами.

Таблицы могут сопровождаться примечанием, используемым с целью пояснения, в случае необходимости, заголовков, методики расчета некоторых показателей, источников информации и так далее.

По логическому содержанию таблица представляет собой «статистическое предложение», основными элементами которого являются подлежащее и сказуемое.

Подлежащим статистической таблицы называется объект, который характеризуется цифрами. Это может быть одна или несколько совокупностей, отдельные единицы совокупности в порядке их перечня или сгруппированные по каким-либо признакам, территориальные единицы и так далее. Обычно подлежащее таблицы дается в левой части, в наименовании строк.

Сказуемое статистической таблицы образует система показателей, которыми характеризуется объект изучения, то есть подлежащее таблицы. Сказуемое формирует верхние заголовки и составляет содержание граф с логически последовательным расположением показателей слева направо.

Различают простые, групповые и комбинационные таблицы.

В **простых** таблицах, как правило, содержится справочный материал, где дается перечень групп или единиц, составляющих объект изучения. При этом части подлежащего не являются группами одинакового качества, отсутствует систематизация изучаемых единиц. Сказуемое этих таблиц содержит абсолютные величины, отражающие объёмы изучаемых процессов.

Примером простой перечневой таблицы является информация о наличии строительных машин в строительных управлениях региона (табл. б).

Таблица 6

Наличие строительных машин в строительных управлениях региона по состоянию на 01.01.2012 году (цифры условные)

Вид машин	Количество машин данного вида, тыс. шт.
Экскаваторы	32,6
Скреперы	8,7
Бульдозеры	31,6
Краны передвижные	40,3
Всего	113,2

Подлежащим таблицы является вид машин, сказуемым - количество разных их видов.

Групповые и комбинационные таблицы предназначены для научных целей, где, в отличие от простых таблиц, в сказуемом - средние и относительные величины на основе абсолютных величин.

Групповая таблица - это таблица, где статистическая совокупность разбивается на отдельные группы по какому-либо одному существенному признаку, при этом каждая группа характеризуется рядом показателей (табл. 2.7).

Комбинационная таблица - это таблица, где подлежащее представляет собой группировку единиц совокупности по двум и более признакам, которые распределяются на группы сначала по одному признаку, а затем на подгруппы по другому признаку внутри каждой из уже выделенных групп. Комбинационная таблица устанавливает существенную связь между факторами группировки (табл. .8).

Таблица 7

Группировка магазинов по уровню производительности труда работников за отчетный период

Уровень производительности труда магазинов, тыс. д. е.	Количество магазинов	Фондоотдача на 1 д. е. активной части основных фондов, д. е.	Рентабельность активной части основных фондов %
А	1	2	3
До 60	4	40,42	2,3
60-70	4	43,1	2,8
70-80	7	75,8	4,7
80-90	7	65,9	4,0
90-100	3	93,1	5,1
Более 100	7	109,3	6,4
Всего	32	X	X
В среднем	X	75	4,4

В приведённой групповой таблице подлежащим являются магазины города, которые разделены на группы по уровню производительности труда, сказуемым - показатели этих организаций (количество магазинов, фондоотдача, рентабельность).

В примере комбинированной таблицы подлежащим являются магазины, распределённые на группы и подгруппы по части площади торгового зала и длительности рабочего дня; в сказуемом приведены показатели, которые наиболее полно характеризуют эффективность работы магазинов.

При составлении таблиц необходимо соблюдать **общие правила**:

- 1) таблица должна быть легко обозримой;
- 2) общий заголовок должен кратко выражать основное содержание;
- 3) наличие строк «общих итогов»;
- 4) наличие нумерации строк, которые заполняются данными;
- 5) соблюдение правила округления чисел.

По **структурному строению** сказуемого различают статистические таблицы с простой и сложной его разработкой.

При **простой разработке** сказуемого показатель, определяющий его, не подразделяется на подгруппы и итоговые значения получаются путём простого суммирования значений по каждому признаку отдельно, независимо друг от друга.

Сложная разработка сказуемого предполагает в сказуемом комбинацию одного признака с другим.

Таблица 8

Группировка продовольственных магазинов города по части площади торгового зала и длительности рабочего дня за отчетный период

Группы и подгруппы магазинов по части площади торгового зала (%) и длительности рабочего дня (в часах)	Количество магазинов	Фондоотдача на 1 д.е. активной части основных фондов, д. е.	Рентабельность активной части основных фондов %
А	1	2	3
К 35%	13	48,5	3,10
В том числе: 8-10 часов	4	41,2	2,20
более 10 часов	9	57,5	4,02
35-45%	21	69,8	5,20
В том числе: 8-10 часов	6	54,6	3,08
более 10 часов	15	77,4	7,10
45-55%	18	90,6	6,40
В том числе: 8-10 часов	5	68,9	4,17
более 10 часов	13	18,7	7,98
Всего	52	X	X
В среднем	X	73,5	4,70

2. Статистический график и его элементы

Графический метод - метод условных изображений статистических данных при помощи геометрических фигур, линий, точек и разнообразных символических образов.

Главное достоинство графиков - наглядность. При правильном построении графика статистические показатели привлекают к себе внимание, становятся выразительными, лаконичными и запоминающимися.

Графики прочно вошли в повседневную работу экономистов, статистиков и работников бухгалтерского учета.

Для построения графика необходимо знать, для каких целей составляется график, изучить исходный материал и владеть методикой графических изображений.

Современную науку невозможно представить без применения графических методов, настолько прочно они вошли в арсенал средств научного общения и в методику научного исследования.

Графиками в статистике называются условные изображения числовых величин и их соотношений в виде различных геометрических образов - точек, линий, плоских фигур и т. п. Использование графиков для изложения статистических показателей позволяет придать последним наглядность и выразительность, облегчить их восприятие, а во многих случаях помогает уяснить сущность изучаемого явления, его закономерности и особенности, увидеть тенденции его развития, взаимосвязь характеризующих его показателей.

Каждый график состоит из графического образа и вспомогательных элементов. **Графический образ** - совокупность точек, линий и фигур, с помощью которых изображаются статистические данные. Эти знаки образуют собственно языковую ткань графика, его основу.

Вспомогательными элементами графика являются: поле графика - пространство, на котором размещаются образующие график геометрические фигуры. Размер графика зависит от его назначения.

Размер поля графика и пропорции его сторон в каждом случае определяются исполнителем. Однако не следует строить сильно удлинённые в горизонтальном или вертикальном направлении графики. Такие графики эстетически невыразительны. **Геометрические знаки или образы** - многообразные знаки, с помощью которых изображают статистические величины. В статистических графиках в качестве геометрических знаков используются точки, отрезки прямой линии, квадраты, прямоугольники, а также фигуры в виде рисунков или силуэтов изображаемых предметов.

Знак составляет основу графика, его язык. Одни и те же данные графически должны быть изображены различными знаками в зависимости от того, какой аспект явления должен подчеркнуть график, на что нацелить внимание его читателя.

Масштабные ориентиры статистических графиков - масштаб, масштабные шкалы и масштабные знаки. **Масштаб** - условная мера перевода числовой величины в графическую и обратно. При построении графика масштаб должен быть таким, чтобы подлежащие нанесению на график данные поместились на поле графика. На вертикальной шкале графика должна быть нулевая точка. В тех случаях, когда минимальное значение признака намного выше нуля, нецелесообразно вести отсчет от нулевой точки, так как поле графика бу-

дет заполнено неравномерно. В таких случаях рекомендуется делать разрыв вертикальной шкалы.

Масштабная шкала - линия, разделенная на отрезки точками. Наиболее часто в статистических графиках используются располагающиеся по осям координат равномерные прямолинейные масштабные шкалы, в которых отрезки между двумя соседними точками (графические интервалы) строго пропорциональны размерам и периодам времени изображаемых на графике данных. В секторных диаграммах используются криволинейные масштабные шкалы. Площадь круга делится на сектора пропорционально изображаемому на графике числам.

Масштабные знаки - эталоны величин, изображаемых на графике в виде отдельных графических знаков: квадратов, кругов, рисунков, силуэтов и др. Ими пользуются для сравнения графических знаков со знаком-эталоном.

Экспликация графика - пояснения, раскрывающие содержание графика: заголовок, единицы измерения, условные обозначения.

Пояснительные надписи к отдельным элементам графика могут быть помещены либо на поле графика, либо в форме условных обозначений за пределами поля графика. Все надписи рекомендуется выполнять горизонтально. Не следует использовать для закрашки графиков слишком пестрые и яркие цвета. Заголовок должен кратко и точно ответить на три вопроса - что, где, когда?

Графики, применяемые для изображения статистических данных, чрезвычайно разнообразны. В данной главе будут рассмотрены наиболее часто применяемые в статистической практике графики.

Графические изображения используются чаще всего для сравнения между собой статистических величин, определения роли отдельных факторов во всей их совокупности, изучения структуры и структурных сдвигов, связи между признаками, изменения явлений во времени, определения степени распространения явления в пространстве и т. д.

Основными элементами графиков, изображающих количественные соотношения, являются шкала, масштаб, оси координат и числовая (координатная) сетка. График должен иметь заглавие, отражающее содержание изображаемого явления, время и место, к которому относятся данные, и расшифровку условных обозначений. Для большей наглядности графика применяют различную штриховку, окраску и т. д.

По способу построения графики делятся на диаграммы, картограммы, картодиаграммы.

Столбиковые диаграммы являются наиболее простым видом диаграмм. При их построении данные изображаются в виде столбиков от числовых значений изображаемых величин по определенному масштабу.

Примером применения столбиковой диаграммы могут служить данные о численности постоянного населения (рис. 9)

Простая столбиковая диаграмма

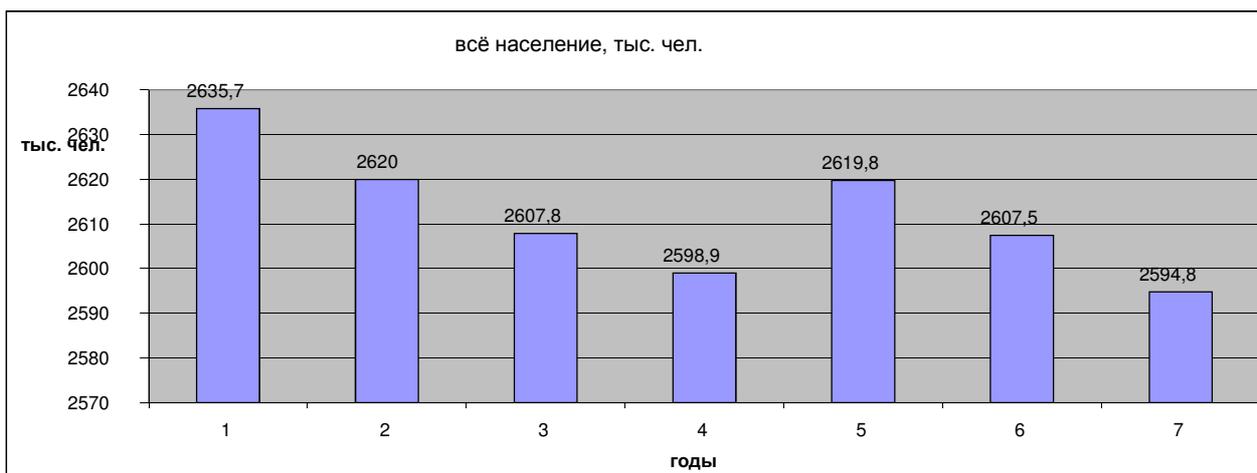


Рис. 9. Динамика численности постоянного населения Волгоградской области на начало 2006 – 2012 гг. тыс. чел.

На масштабной шкале проставляются круглые или округлённые значения изображаемых величин. Такая диаграмма называется простой, так как столбики не имеют внутренних долей. Если же они делятся на части, то диаграмма называется сложной (рис. 10).

Сложная столбиковая диаграмма

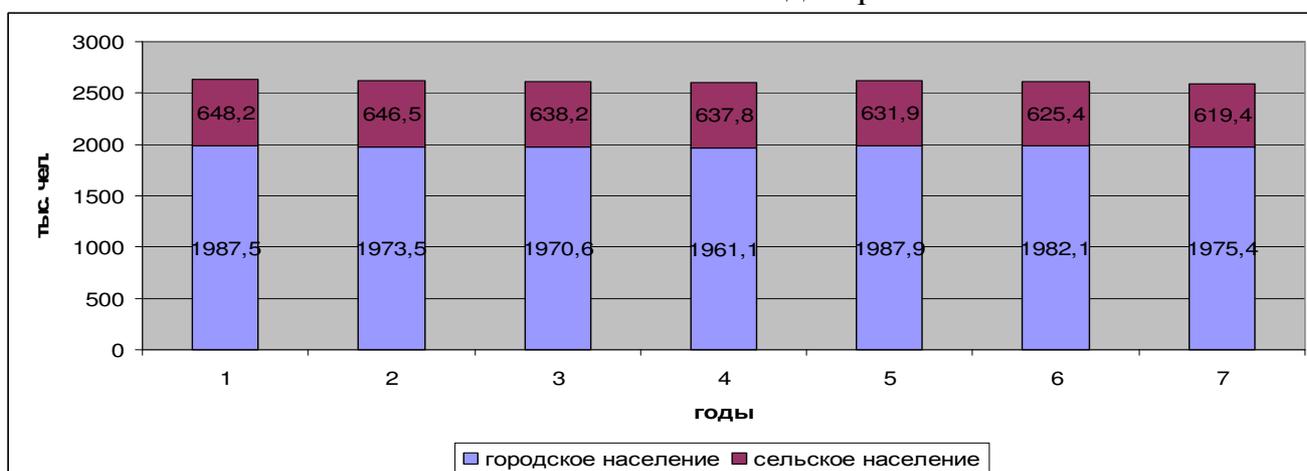


Рис. 10. Динамика численности постоянного городского и сельского населения Волгоградской области на начало 2006 – 2012 гг., тыс. чел.

Разновидностью столбиковых диаграмм являются ленточные диаграммы. Они изображают размеры признака в виде расположенных по горизонтали прямоугольников одинаковой ширины, но различной длины, пропорционально изображаемым величинам. Начало полос должно находиться на одной и той же вертикальной линии (рис. 11).

Простая ленточная диаграмма

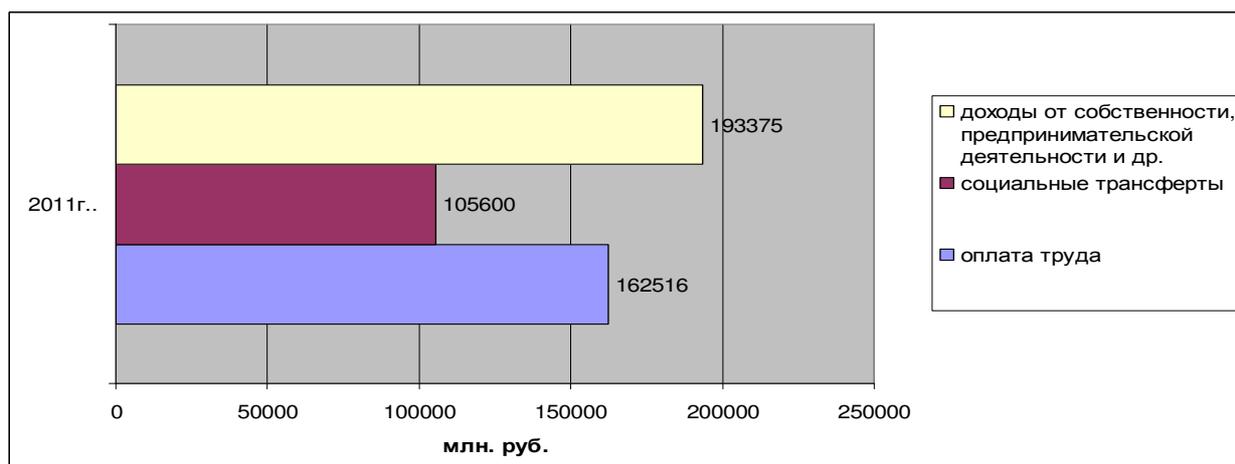


Рис. 2.11. Состав денежных доходов населения Волгоградской области за 2011 г

Квадратные и круговые диаграммы относятся к типу плоскостных диаграмм. Они представляют собой различные по размерам квадраты или круги, площади которых пропорциональны величине изображаемых статистических данных.

Если числа обозначить буквой d , то стороны квадратов будут равны \sqrt{d} . Известно, что площадь круга $S = \eta R^2$. Поэтому радиусы отдельных кругов будут равны квадратному корню из значений изображаемых величин.

Недостаток квадратных и круговых диаграмм заключается в том, что они менее наглядны, чем столбиковые, так как сравниваются площади, а не высоты, и строить их несколько сложнее.

Нередко состав, структура того или иного явления изображаются с помощью кругов, разделённых на сектора, пропорциональные долям частей явлений. Круг принимается за целое (100%) и разбивается на сектора, дуги которых пропорциональны значениям отдельных частей изображаемых величин. Дуга каждого сектора круга рассчитывается по формуле:

$$360^\circ \cdot d / 100,$$

где 360° - весь круг (100%),

d - величина изображаемого явления в процентах.

Такие диаграммы называются секторными (рис. 12).

Секторная диаграмма

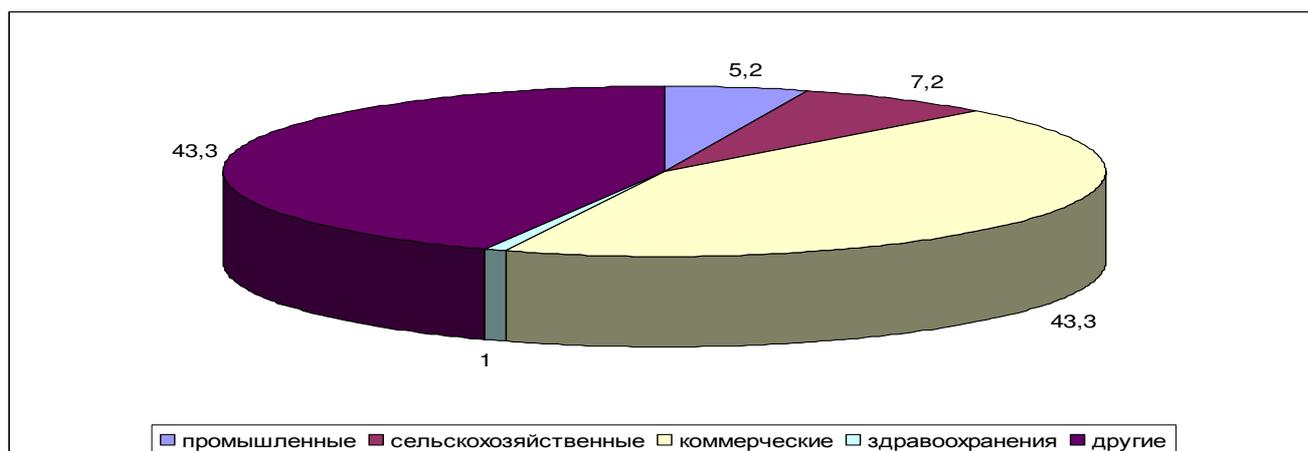


Рис. 12. Структура введенных в действие зданий нежилого назначения за 2011 г. (в процентах к итогу)

Секторные диаграммы следует применять лишь в тех случаях, когда совокупность делится не более чем на 4-5 частей, а также при условии значительных различий сравниваемых структур, иначе они теряют свою выразительность.

Наиболее распространенным видом диаграмм являются линейные. Чаще всего они используются для изображения динамических рядов и при изучении связи между явлениями. При построении линейных диаграмм применяют координатную или числовую сетку. На оси абсцисс системы прямоугольных координат на равном расстоянии друг от друга наносятся точки, соответствующие числу членов динамического ряда, а на оси ординат - показатели по принятому масштабу. После этого наносят данные и, соединив концы перпендикуляров, получают ломаную линию, характеризующую изображаемый динамический ряд (рис. 13).

Линейная диаграмма

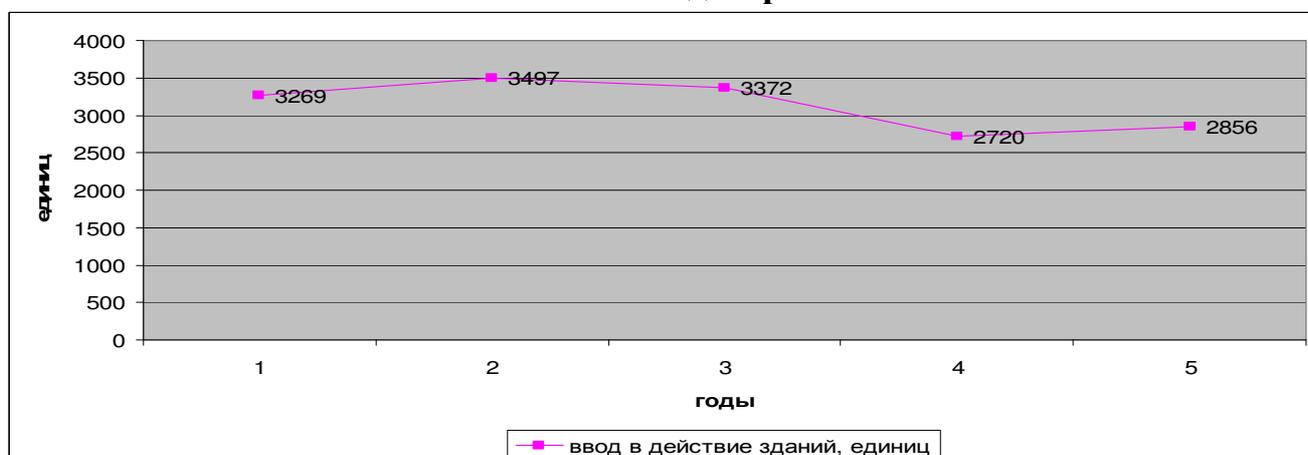


Рис. 13. Динамика ввода в действие зданий за 2007-2011 гг. в Волгоградской области

Общий вид графика зависит от правильного соотношения масштабов на осях абсцисс и ординат. В противном случае колебания будут либо малозаметными, либо слишком резкими. Если данные относятся к различным периодам времени, интервалы между ними при нанесении на оси абсцисс должны быть пропорциональны длительности периодов. При помощи линейных диаграмм можно выразить одновременно ряд показателей, что даёт возможность сравнивать их друг с другом.

«АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ В СТАТИСТИКЕ»

1. Понятие абсолютного показателя. Виды абсолютных показателей

Первичная статистическая информация выражается, прежде всего, в виде абсолютных показателей, которые являются количественной базой всех форм учёта.

Абсолютные показатели характеризуют итоговую численность единиц совокупности или её частей, размеры (объёмы, уровни) изучаемых явлений и процессов, выражают временные характеристики. Абсолютные показатели могут быть только именованными числами, где единица измерения выражается в конкретных цифрах.

В зависимости от сущности исследуемого явления и поставленных задач единицы измерения могут быть натуральными, условно-натуральными, стоимостными и трудовыми.

Натуральные единицы измерения соответствуют потребительским или природным свойствам товара или предмета и оцениваются в физических мерах массы, длины, объёма (килограмм, тонна, метр и т.д.).

Разновидностью натуральных единиц выступают **условно-натуральные**, которые используются в тех случаях, если продукт, имея несколько разновидностей, должен переводиться в условный продукт с помощью специальных коэффициентов (молочные продукты с разным содержанием сливочной основы, мыло с разным содержанием жирных кислот и т.д.).

Стоимостные единицы измерения оценивают социально-экономические процессы и явления в денежном выражении (цены, сопоставимые цены), что очень важно в условиях рыночной экономики.

Трудовые единицы измерения призваны отражать затраты труда, трудоёмкость технологических операций в человеко-днях, человеко-часах.

Вся совокупность абсолютных величин включает:

индивидуальные показатели - характеризуют значения отдельных единиц совокупности,

суммарные показатели - характеризуют итоговое значение нескольких единиц совокупности или итоговое значение существенного признака по той или иной части совокупности.

Абсолютные показатели следует также подразделить на моментные и интервальные.

Моментные абсолютные показатели характеризуют факт наличия явления или процесса, его размер (объем) на определенную дату времени.

Интервальные абсолютные показатели характеризуют итоговый объем явления за тот или иной период времени (например, выпуск продукции за квартал или за год и т. д.), допуская при этом последующее суммирование.

Абсолютные показатели не могут дать исчерпывающего представления об изучаемой совокупности или явлении, поскольку не могут отразить структуру, взаимосвязи, динамику. Данные функции выполняют относительные показатели, которые определяются на основе абсолютных показателей.

2. Относительные показатели, их роль и типология

В статистике относительные показатели используют в сравнительном анализе, в обобщении и синтезе.

Относительные показатели - это цифровые обобщающие показатели, они есть результат сопоставления двух статистических величин.

По своей природе относительные величины производны от деления текущего (сравниваемого) абсолютного показателя на базисный показатель.

Относительные показатели могут быть получены или как соотношения одноименных статистических показателей, или как соотношения разноименных статистических показателей. В первом случае получаемый относительный показатель рассчитывается или в процентах, или в относительных единицах, или в промилле (в тысячных долях). Если соотносятся разноименные абсолютные показатели, то относительный показатель в большинстве случаев бывает именованным.

В зависимости от величин числителя и знаменателя этой дроби относительные величины могут быть выражены в таких формах: коэффициентах (частях), процентах (%), промиллях (‰), продецимиллях (‱), когда за базу сравнения принимают соответственно 1, 100, 1000, 10000 единиц.

Относительные величины, используемые в статистической практике:

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

Относительная величина структуры (ОВС) характеризует структуру совокупности, определяет долю (удельный вес) части в общем объеме совокупности.

ОВС рассчитывают как отношение объема части совокупности к абсолютной величине всей совокупности, определяя тем самым удельный вес части в общем объеме совокупности (%):

$$ОВС = m_i / M,$$

где m_i - объём исследуемой части совокупности;
 M - общий объём исследуемой совокупности.

Относительная величина координации (ОВК) характеризует соотношение между двумя частями исследуемой совокупности, одна из которых выступает как база сравнения (%):

$$\text{ОВК} = m_i / m_6 * 100, \%$$

где m_i - одна из частей исследуемой совокупности;
 m_6 - часть совокупности, которая является базой сравнения. За базу сравнения принимают наибольшее значение.

Относительная величина планового задания (ОВПЗ) используется для расчёта в процентном отношении увеличения (уменьшения) величины показателя плана по сравнению с его базовым уровнем в предшествующем периоде, для чего используется формула:

$$\text{ОВПЗ} = P_{\text{пл}} / P_6 * 100, \%$$

где $P_{\text{пл}}$ - плановый показатель;
 P_6 - базовый показатель в предшествующем периоде.

Относительная величина выполнения плана (ОВВП) характеризует степень выполнения планового задания за отчётный период (%) и рассчитывается по формуле:

$$\text{ОВВП} = P_{\text{ф}} / P_{\text{пл}} * 100, \%$$

где $P_{\text{ф}}$ – фактическая величина выполнения плана за отчётный период;
 $P_{\text{пл}}$ - величина плана за отчётный период.

Относительная величина динамики (ОВД) характеризует изменение объёма одного и того же явления во времени в зависимости от принятого базового уровня.

ОВД рассчитывают как отношение уровня анализируемого явления или процесса в текущий момент времени к уровню этого явления или процесса за прошедший период времени. В результате мы получаем коэффициент роста, который выражается кратным отношением. При исчислении этой величины в процентах (результат умножается на 100) получаем темп роста.

Темпы роста можно просчитывать как с постоянным базовым уровнем (базисные темпы роста - ОВД_6), так и с переменным базовым уровнем (цепные темпы роста - $\text{ОВД}_ц$):

$$\text{ОВД}_6 = P_T / P_6 * 100, \%$$

где P_T - уровень текущий;
 P_6 - уровень базисный;

$$\text{ОВДц} = P_T / P_{T-1} * 100, \%$$

где P_T - уровень текущий;
 P_{T-1} - уровень, предшествующий текущему.

Относительная величина сравнения (ОВСр) - соотношение одноимённых абсолютных показателей, относящихся к разным объектам, но к одному и тому же времени (например, соотносятся темпы роста населения в разных странах за один и тот же период времени):

$$\text{ОВСр} = M_A / M_B * 100, \%$$

где M_A - показатель первого одноимённого исследуемого объекта;
 M_B - показатель второго одноимённого исследуемого объекта (база сравнения).

Все предыдущие показатели относительных величин характеризовали соотношения одноименных статистических объектов.

Однако есть группа относительных величин, которые характеризуют соотношение разноименных, но связанных между собой статистических показателей. Эту группу называют группой **относительных величин интенсивности (ОВИ)**, которые выражаются, как правило, именованными числами.

В статистической практике относительные величины интенсивности применяются при исследовании степени объёмности явления по отношению к объёму среды, в которой происходит распространение этого явления. ОВИ здесь показывает, сколько единиц одной совокупности (числитель) приходится на одну, на десять, на сто единиц другой совокупности (знаменатель).

Примерами относительных величин интенсивности могут служить показатели уровня технического развития производства, уровня благосостояния граждан, показатели обеспеченности населения средствами массовой информации, предметами культурно-бытового назначения и т.д.

ОВИ рассчитывается по формуле:

$$\text{ОВИ} = A / V_A * 100,$$

где A - распространение явления;
 V_A - среда распространения явления A .

При расчёте относительных величин интенсивности может возникнуть проблема выбора адекватной явлению базы сравнения (среды распространения явления). Например, при определении показателя плотности населения нельзя брать в качестве базы сравнения общий размер территории того или иного гос-

ударства, в этом случае базой сравнения может быть лишь территория в 1 км². Критерием правильности расчёта является сопоставимость по разработанной методологии расчёта сравниваемых показателей, применяющихся в статистической практике.

«СРЕДНИЕ ВЕЛИЧИНЫ В СТАТИСТИКЕ»

1. Сущность и значение средней величины.

Область применения средних величин в статистическом исследовании

Средние величины используются на этапе обработки и обобщения полученных первичных статистических данных. Потребность определения средних величин связана с тем, что у различных единиц исследуемых совокупностей индивидуальные значения одного и того же признака, как правило, неодинаковы.

Средней величиной называют показатель, который характеризует обобщённое значение признака или группы признаков в исследуемой совокупности.

Если исследуется совокупность с качественно однородными признаками, то средняя величина выступает здесь как типическая средняя. Например, для групп работников определённой отрасли с фиксированным уровнем дохода определяется типическая средняя расходов на предметы первой необходимости, т.е. типическая средняя обобщает качественно однородные значения признака в данной совокупности, каковым является доля расходов у работников данной группы на товары первой необходимости.

При исследовании совокупности с качественно разнородными признаками на первый план может выступить нетипичность средних показателей. Такими, к примеру, являются средние показатели произведённого национального дохода на душу населения (разные возрастные группы), средние показатели урожайности зерновых культур по всей территории России (районы разных климатических зон и разных зерновых культур), средние показатели рождаемости населения по всем регионам страны, средние температуры за определённый период и т.д. Здесь средние величины обобщают качественно разнородные значения признаков или системных пространственных совокупностей (международное сообщество, континент, государство, регион, район и т.д.) или динамических совокупностей, протяженных во времени (век, десятилетие, год, сезон и т.д.). Такие средние величины называют системными средними.

Таким образом, значение средних величин состоит в их обобщающей функции. Средняя величина заменяет большое число индивидуальных значений признака, обнаруживая общие свойства, присущие всем единицам совокупности. Это, в свою очередь, позволяет избежать случайных причин и выявить общие закономерности, обусловленные общими причинами.

2. Виды средних величин и методы их расчёта

На этапе статистической обработки могут быть поставлены самые различные задачи исследования, для решения которых нужно выбрать соответ-

ствующую среднюю. При этом необходимо руководствоваться следующим правилом: величины, которые представляют собой числитель и знаменатель средней, должны быть логически связаны между собой.

Используются две категории средних величин (рис. 14):

- * степенные средние;
- * структурные средние.

Первая категория степенных средних включает: среднюю арифметическую, среднюю гармоническую, среднюю квадратическую, среднюю геометрическую и средняя кубическая.



Рис. 14. Виды средних величин в статистике

Вторая категория (структурные средние) - это мода и медиана. Эти виды средних будут рассмотрены в [теме: «Структурные характеристики вариационного ряда распределения»](#).

Введём следующие условные обозначения:

\bar{x} - средняя, где черта сверху свидетельствует о том, что имеет место осреднение индивидуальных значений;

x_i - варианты (значение) осредняемого признака или срединное значение интервала, в котором измеряется вариант;

n - число вариантов;

f - частота (повторяемость индивидуальных значений признака).

k - показатель степени.

Степенные средние в зависимости от представления исходных данных могут быть простыми и взвешенными.

Простая средняя считается по не сгруппированным данным и имеет следующий вид:

$$\bar{x} = \sqrt[k]{\frac{\sum x_i^k}{n}},$$

Взвешенная средняя считается по сгруппированным данным и имеет общий вид:

$$\bar{x} = \sqrt[k]{\frac{\sum x_i^k f_i}{\sum f_i}}.$$

В зависимости от того, какое значение принимает показатель степени, различают следующие виды степенных средних:

- средняя арифметическая, если $k = 1$;
- средняя гармоническая, если $k = -1$;
- средняя геометрическая, если $k = 0$;
- средняя квадратическая, если $k = 2$;
- средняя кубическая, если $k = 3$.

Формулы степенных средних приведены в табл. 9.

Таблица 9

Виды степенных средних

Вид степенной средней	Показатель степени (m)	Формула расчета	
		простая	взвешенная
гармоническая	- 1	$\bar{x} = \frac{n}{\sum \frac{1}{x_i}}$	$\bar{x} = \frac{\sum w_i}{\sum \frac{w_i}{x_i}}$
геометрическая	0	$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{nx_i}$	$\bar{x} = \sqrt[\sum f_i]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}} = \sqrt[\sum f_i]{nx_i^{f_i}}$
арифметическая	1	$\bar{x} = \frac{\sum x_i}{n}$	$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$
квадратическая	2	$\bar{x} = \sqrt{\frac{\sum x_i^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}}$
кубическая	3	$\bar{x} = \sqrt[3]{\frac{\sum x_i^3}{n}}$	$\bar{x} = \sqrt[3]{\frac{\sum x_i^3 f_i}{\sum f_i}}$

Если рассчитать все виды средних для одних и тех же исходных данных, то значения их окажутся неодинаковыми. Здесь действует правило мажорантности средних: с увеличением показателя степени k увеличивается и соответствующая средняя величина.

В статистической практике чаще, чем остальные виды средних взвешенных, используется средняя арифметическая и средняя гармоническая взвешен-

ные. Выбор вида степенной средней определяется экономическим содержанием задачи и наличием данных.

Рассмотрим среднюю арифметическую простую и взвешенную.

Пример: Студент Петров по результатам учебного семестра имеет следующие оценки: теория бухгалтерского учета - 4, экономическая статистика - 5, финансы, денежное обращение и кредит - 3, экономика фирмы - 2. Какова его средняя оценка по результатам семестра?

Поскольку каждая оценка встречается один раз, для расчета средней применяем формулу арифметической простой.

Перечисленные дисциплины студент Петров сдал в среднем на 3,5 балла.

Пример: Имеются следующие данные о распределении бригад по уровню выработки продукции (табл. 10).

Таблица 10

Распределение бригад по уровню выработки продукции за смену

Бригады	Выработка продукции в среднем на одного человека, шт. (x)	Число рабочих, чел. (f)
1	110	12
2	120	10
3	130	14
4	140	8
Итого	-	44

Определим сменную выработку рабочего в среднем по четырём бригадам. Введём строку условных обозначений, приняв за x значения осредняемого признака, f - число рабочих с данным значением x.

Исходные данные представлены в виде дискретного ряда распределения; каждое x встречается несколько раз, следовательно, применяем формулу средней арифметической взвешенной.

В смену рабочий данных четырёх бригад изготавливает в среднем 124 единицы продукции.

Расчёт средней по интервальному ряду

Если исходные данные заданы в виде интервального ряда, то:

- 1) закрывают открытые интервалы, приняв их равными ближайшим закрытым;
- 2) за значения осредняемого признака x берут середины интервалов и строят условный дискретный ряд распределения:

$$x = \frac{x_{н.г.} + x_{в.г.}}{2}$$

где $x_{н.г.}$ - значение нижней границы интервала («от»); $x_{в.г.}$ - значение верхней границы интервала («до»).

- 3) расчёт средней производится по средней арифметической взвешенной.

Пример: Имеются данные о распределении рабочих по стажу работы (табл. 11):

Таблица 11

Стаж работы, лет	Доля рабочих, % к итогу
До 5	10
5-10	44
10-15	30
15-20	10
20 и выше	6

Каков средний стаж работы рабочего?

Строим расчётную таблицу, обозначив долю рабочих через f .

Закрываем открытый интервал «до 5». Ширина ближайшего закрытого интервала равна 5 годам (5-10), следовательно, наш интервал примет вид от 0 до 5. Аналогично открытый интервал «20 и выше» примет вид 20-25, поскольку ширина ближайшего закрытого (15-20) равна 5.

Стаж работы, лет	f	x	xf
До 5	10	2,5	25
5-10	44	7,5	330
10-15	30	12,5	375
15-20	10	17,5	175
20 и выше	6	22,5	135
Итого	100		1040

Находим середину каждого интервала и принимаем ее за значение x .

Исчисляем значения $x \cdot f$ и сумму этих значений, необходимую для расчета средней арифметической взвешенной, заносим результаты в расчетную таблицу.

Рабочий данного цеха отработал в среднем 10,4 года. Расчет средней по интервальному ряду распределения дает приближенный результат за счет того, что за значения x берутся не точные данные, а осредненные значения (середины интервалов).

Средняя гармоническая имеет более сложную конструкцию, чем средняя арифметическая. Её чаще всего применяют для расчётов тогда, когда в качестве весов используются не единицы совокупности – носители признака, а произведения этих единиц на значения признака, т.е. $w = xf$.

Результат получился тот же.

3. Свойства средней арифметической

Средняя арифметическая обладает рядом свойств, знание которых необходимо для понимания сущности средних, а также для упрощения их вычисления.

1. Средняя арифметическая суммы варьирующих величин равна сумме средних арифметических величин:

Если $x_i = y_i + z_i$, то

$$\bar{x} = \frac{\sum x_i}{n} = \frac{\sum (y_i + z_i)}{n} = \frac{\sum y_i}{n} + \frac{\sum z_i}{n} = \bar{y} + \bar{z} \quad \text{Это}$$

правило показывает, в каких случаях можно суммировать средние величины. Если, например, выпускаемые изделия состоят из двух деталей у и z и на изготовление каждой из них расходуется в среднем $y = 3$ ч, $z = 5$ ч, то средние затраты времени на изготовление одного изделия (х), будут равны: $3 + 5 = 8$ ч., т.е. $x = y + z$.

2. Алгебраическая сумма отклонений индивидуальных значений признака от средней равна нулю, так как сумма отклонений в одну сторону погашается суммой отклонений в другую, т.е., $\sum (x - \bar{x}) = 0$, потому что

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

Это правило показывает, что средняя является равнодействующей.

3. Если все варианты ряда уменьшить или увеличить на одно и тоже число а, то средняя уменьшится или увеличится на это же число а:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum (x \pm a)}{\sum f_i} \pm a.$$

4. Если все варианты ряда уменьшить или увеличить в А раз, то средняя также уменьшится или увеличится в А раз:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum \frac{x_i}{A} f_i}{\sum f_i} \cdot A = \frac{\sum A x_i f_i}{\sum f_i} \div A.$$

5. Если все частоты ряда разделить или умножить на одно и тоже число d, то средняя не изменится:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\sum x_i \frac{f_i}{d_i}}{\sum \frac{f_i}{d_i}} = \frac{\sum x_i f_i d}{\sum f_i d}.$$

Это свойство показывает, что средняя зависит не от размеров весов, а от соотношения между ними. Следовательно, в качестве весов могут выступать не только абсолютные, но и относительные величины.

«ПОКАЗАТЕЛИ ВАРИАЦИИ В СТАТИСТИКЕ»

1. Понятие вариации. Показатели вариации

Вариацию можно определить как количественное различие значений одного и того же признака у отдельных единиц совокупности.

Термин «вариация» имеет латинское происхождение - variatio, что означает различие, изменение, колеблемость.

Изучение вариации в статистической практике позволяет установить зависимость между изменением, которое происходит в исследуемом признаке, и теми факторами, которые вызывают данное изменение.

Для измерения вариации признака используют как абсолютные, так и относительные показатели.

К абсолютным показателям вариации относят: размах вариации, среднее линейное отклонение, среднее квадратическое отклонение, дисперсию.

К относительным показателям вариации относят: коэффициент осцилляции, линейный коэффициент вариации, относительное линейное отклонение и др.

Размах вариации R. Это самый доступный по простоте расчёта абсолютный показатель, который определяется как разность между самым большим и самым малым значениями признака у единиц данной совокупности:

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

Размах вариации (размах колебаний) - важный показатель колеблемости признака, но он даёт возможность увидеть только крайние отклонения, что ограничивает область его применения.

Для более точной характеристики вариации признака на основе учёта его колеблемости используются другие показатели.

Среднее линейное отклонение \bar{d} , которое вычисляют для того, чтобы учесть различия всех единиц исследуемой совокупности. Эта величина определяется как средняя арифметическая из абсолютных значений отклонений от средней. Так как сумма отклонений значений признака от средней величины равна нулю, то все отклонения берутся по модулю.

Формула среднего линейного отклонения (простая):

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

Формула среднего линейного отклонения (взвешенная):

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i},$$

При использовании показателя среднего линейного отклонения возникают определенные неудобства, связанные с тем, что приходится иметь дело не только с положительными, но и с отрицательными величинами, что побудило искать другие способы оценки вариации, чтобы иметь дело только с положительными величинами. Таким способом стало возведение всех отклонений во вторую степень. Обобщающие показатели, найденные с использованием вторых степеней отклонений, получили очень широкое распространение. К таким показателям относятся среднее квадратическое отклонение σ и среднее квадратическое отклонение в квадрате σ^2 , которое называют дисперсией.

Средняя квадратическая простая:

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

Средняя квадратическая взвешенная:

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

Дисперсия есть не что иное, как средний квадрат отклонений индивидуальных значений признака от его средней величины.

Формулы дисперсии взвешенной:

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

и простой:

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

Расчёт дисперсии можно упростить. Для этого используется способ отсчёта от условного нуля (способ моментов), если имеют место равные интервалы в вариационном ряду.

Кроме показателей вариации, выраженных в абсолютных величинах, в статистическом исследовании используются показатели вариации (V), выраженные в относительных величинах, особенно для целей сравнения колеблемости различных признаков одной и той же совокупности или для сравнения колеблемости одного и того же признака в нескольких совокупностях.

Данные показатели рассчитываются как отношение:

- размаха вариации к средней величине признака (коэффициент осцилляции):

$$V_R = \frac{R}{\bar{X}} \times 100\% ,$$

- отношение среднего линейного отклонения к средней величине признака (линейный коэффициент вариации):

$$V_d = \frac{\bar{d}}{\bar{X}} \times 100\% ,$$

- отношение среднего квадратического отклонения к средней величине признака (коэффициент вариации):

$$V_\sigma = \frac{\sigma}{\bar{X}} \times 100\% ,$$

и, как правило, выражаются в процентах.

Из приведенных формул видно, что чем больше коэффициент V приближен к нулю, тем меньше вариация значений признака.

В статистической практике наиболее часто применяется коэффициент вариации. Он используется не только для сравнительной оценки вариации, но и для характеристики однородности совокупности. Совокупность считается однородной, если коэффициент вариации не превышает 33% (для распределений, близких к нормальному).

2. Характеристика структурных средних показателей

Для получения более полной характеристики вариационного ряда помимо средней величины рассчитываются так называемые структурные показатели. К ним относятся мода, медиана, квартили, децили и т.д.

Мода (Mo) — это наиболее часто встречающееся значение признака, или иначе говоря, значение варианты с наибольшей частотой (f).

Медиана - это значение признака, которое делит статистическую совокупность на две равные части: половина единиц совокупности имеет значения признака не меньше медианы, другая половина — значения признака не больше медианы.

Значения изучаемого признака всех единиц статистической совокупности можно расположить в порядке возрастания (или убывания). В этом случае мы получим ранжированный ряд. Если число единиц совокупности нечетное, то значение признака, находящееся в середине ранжированного ряда, будет являться медианой. Если число единиц совокупностей четное, то медианой будет средняя величина из двух значений признака, находящихся в середине ряда.

Квартили (Q) делят ранжированный ряд на четыре равные части: первый квартиль (Q_1) — значение признака, которое не превышает 25% единиц совокупности; второй квартиль (Q_2) совпадает с медианой (Me), т.е. 50% единиц совокупности по своему значению меньше второго квартиля и 50% больше

него; третий квартиль (Q_3) — значения признака в совокупности, 75% единиц в которой меньше его по своему значению.

Децили (D) делят ранжированный ряд на десять равных частей: первым децилем (D_1) является значение признака, которое не превышает 10% единиц совокупности, вторым (D_2) 20%, третьим (D_3) — 30% и т.д. При этом пятый дециль (D_5) совпадает с медианой и вторым квартилем (D_2).

Медиана, квартили и децили относятся к группе квантилей.

Квантили — это показатели, которые делят вариационные ряды на определенное количество равных частей, Среди них, помимо названных, также имеются квинтили, которые делят ряд на пять равных частей, перцентили — на сто и т.д.

Структурные показатели не зависят от того, имеются ли в статистической совокупности аномальные (резко выделяющиеся) наблюдения. И если средняя величина при их наличии теряет свою практическую значимость, то информативность медианы наоборот усиливается — она начинает выполнять функции средней, т.е. характеризовать центр совокупности.

Способы расчета рассматриваемых структурных показателей зависят от вида вариационного ряда.

«РЯДЫ ДИНАМИКИ. ПОКАЗАТЕЛИ ИЗМЕНЕНИЯ УРОВНЕЙ РЯДОВ ДИНАМИКИ»

1. Ряды динамики и их виды

Изменение социально-экономических явлений во времени изучается статистикой методом построения и анализа динамических рядов.

Ряды динамики - это значения статистических показателей, которые представлены в определённой хронологической последовательности.

Каждый динамический ряд содержит две составляющие:

- 1) показатели периодов времени (годы, кварталы, месяцы, дни или даты);
- 2) показатели, характеризующие исследуемый объект за временные периоды или на соответствующие даты, которые называют уровнями ряда.

Уровни ряда выражаются как абсолютными, так и средними или относительными величинами.

В зависимости от характера показателей строят динамические ряды абсолютных, относительных и средних величин. Ряды динамики из относительных и средних величин строят на основе производных рядов абсолютных величин.

Различают интервальные и моментные ряды динамики.

Динамический **интервальный ряд** содержит значения показателей за определённые периоды времени. В интервальном ряду уровни можно суммировать, получая объём явления за более длительный период, или так называемые накопленные итоги.

Динамический **моментный ряд** отражает значения показателей на определённый момент времени (дату времени). В моментных рядах исследователя может интересовать только разность явлений, отражающая изменение уровня

ряда между определёнными датами, поскольку сумма уровней здесь не имеет реального содержания. Накопленные итоги здесь не рассчитываются.

Важное аналитическое отличие моментных рядов от интервальных состоит в том, что сумма уровней интервального ряда вполне реальный показатель - общий выпуск продукции за определенные периоды времени, общие затраты рабочего времени, общий объем продаж акций и т.д., сумма же уровней моментного ряда не имеет смысла.

По расстоянию между датами или интервалами времени выделяют полные и неполные хронологические ряды. В **полных рядах** даты или периоды времени указываются через равный интервал. В **неполных рядах** расстояние между датами или периодами не равны.

По числу показателей можно выделить изолированные и комплексные (многомерные) ряды динамики. Если ведётся анализ во времени одного показателя, то ряд динамики изолированный (например, данные о добыче угля). В многомерном ряду представлена динамика нескольких показателей, характеризующих одно явление.

На рис. 15 показана рассмотренная классификация рядов динамики.

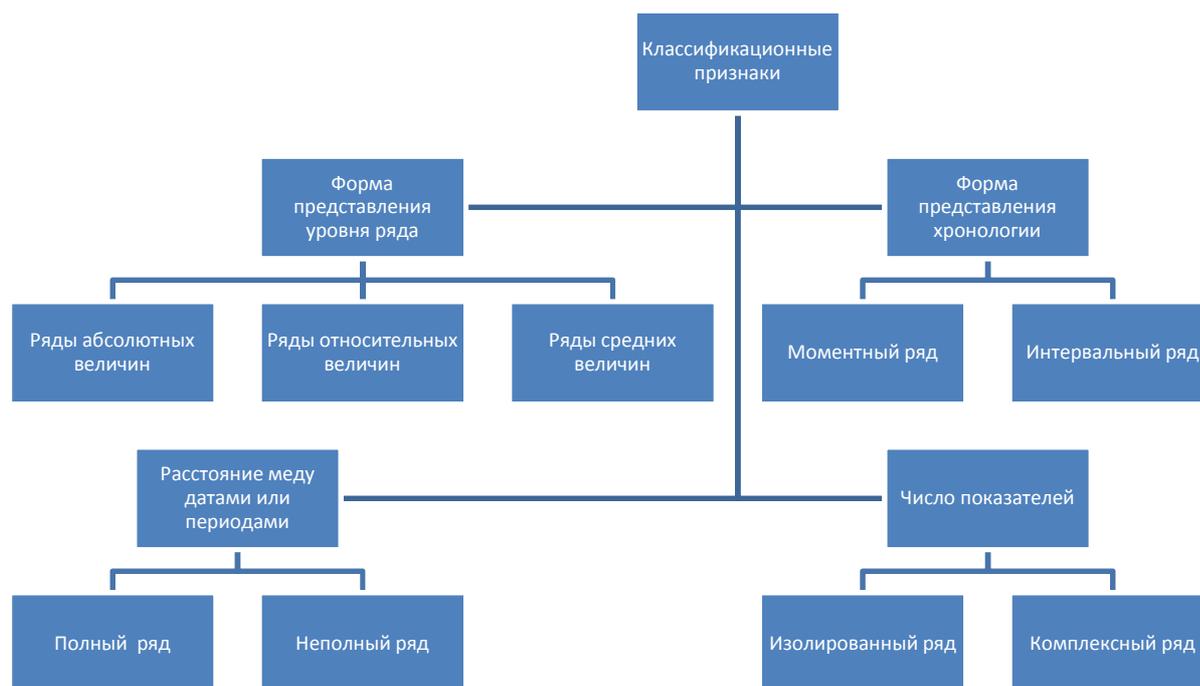


Рис. 15. Виды рядов динамики

Статистические данные должны быть сопоставимы по территории, кругу охватываемых объектов, единицам измерения, времени регистрации, ценам, методологии расчёта.

Сопоставимость по территории означает, что данные по странам и регионам, границы которых изменились, должны быть пересчитаны в старых пределах.

Сопоставимость по кругу охватываемых объектов означает сравнение уровней, характеризующих определённую совокупность. Территориальная и

объёмная сопоставимость обеспечивается смыканием рядов динамики, при этом либо абсолютные уровни заменяются относительными, либо делается пересчёт в условные абсолютные уровни.

Не возникает особых сложностей при обеспечении сопоставимости данных по единицам измерения; стоимостная сравнимость достигается системой сопоставимых цен. При изменении масштаба цен необходимо стоимостные показатели пересчитывать в новый масштаб иен.

Единство методологии расчёта важно для средних и относительных уровней ряда.

Для приведения уровней ряда к сопоставимому виду используется **смыкание рядов динамики**, методика которого включает:

1. Для даты или периода, когда произошло изменение, необходимо зафиксировать два уровня ряда, например в прежних и новых территориальных границах, ценах, единицах измерения.

2. Рассчитывается коэффициент пересчета делением уровня ряда после изменения на уровень той же даты или периода до изменения.

3. Строится сопоставимый ряд, в котором уровни до даты или периода, в котором произошли изменения, умножаются на коэффициент пересчета.

2. Показатели изменений уровней динамических рядов

Для характеристики интенсивности развития во времени используются статистические показатели, получаемые сравнением уровней между собой, в результате чего получаем систему абсолютных и относительных показателей динамики: абсолютный прирост, коэффициент роста, темп роста, темп прироста, абсолютное значение 1% прироста.

Для характеристики интенсивности развития за длительный период рассчитываются средние показатели: средний уровень ряда, средний абсолютный прирост, средний коэффициент роста, средний темп роста, средний темп прироста, среднее абсолютное значение 1% прироста.

Если в ходе исследования необходимо сравнить несколько последовательных уровней, то можно получить или сравнение с постоянной базой (базисные показатели), или сравнение с переменной базой (цепные показатели).

Базисные показатели характеризуют итоговый результат всех изменений в уровнях ряда от периода базисного уровня до данного (i-го) периода.

Цепные показатели характеризуют интенсивность изменения уровня от одного периода к другому в пределах того промежутка времени, который исследуется.

Абсолютный прирост выражает абсолютную скорость изменения ряда динамики и определяется как разность между данным уровнем и уровнем, принятым за базу сравнения.

Абсолютный прирост (базисный):

$$\Delta y_{\text{б}} = y_i - y_0,$$

где y_i - уровень сравниваемого периода;
 y_0 - уровень базисного периода.

Абсолютный прирост с переменной базой (цепной), который называют скоростью роста,

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

где y_i - уровень сравниваемого периода;
 y_{i-1} - уровень предшествующего периода.

Коэффициент роста K_i определяется как отношение данного уровня к предыдущему или базисному, показывает относительную скорость изменения ряда. Если коэффициент роста выражается в процентах, то его называют темпом роста.

Коэффициент роста базисный

$$K_b = \frac{y_i}{y_0},$$

Коэффициент роста цепной

$$K_u = \frac{y_i}{y_{i-1}},$$

Темп роста

$$Tr = K \cdot 100\%,$$

Темп прироста T_{Π} определяется как отношение абсолютного прироста данного уровня к предыдущему или базисному.

Темп прироста базисный

$$T_{\Pi_b} = \frac{y_i - y_0}{y_0} \cdot 100\%,$$

Темп прироста цепной

$$T_{\Pi_u} = \frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100\%,$$

Темп прироста можно рассчитать и иным путём: как разность между темпом роста и 100 % или как разность между коэффициентом роста и 1 (единицей):

$$T_{\pi} = T_p - 100\%,$$

$$T_{\pi} = K_i - 1,$$

Абсолютное значение одного процента прироста A_i . Этот показатель служит косвенной мерой базисного уровня. Представляет собой одну сотую часть базисного уровня, но одновременно представляет собой и отношение абсолютного прироста к соответствующему темпу роста.

Данный показатель рассчитывают по формуле:

$$A_i = \frac{y_i - y_{i-1}}{T_{\pi_{i(i-1)}}} \frac{y_{i-1}}{100} = 0.01 y_{i-1},$$

3. Средние показатели ряда динамики

Для характеристики динамики изучаемого явления за продолжительный период рассчитывают группу средних показателей динамики. Можно выделить две категории показателей в этой группе:

- а) средние уровни ряда;
- б) средние показатели изменения уровней ряда.

Средние уровни ряда рассчитываются в зависимости от вида временного ряда.

Для интервального ряда динамики абсолютных показателей средний уровень ряда рассчитывается по формуле простой средней арифметической:

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

где n - число уровней ряда.

Для моментного динамического ряда средний уровень определяется следующим образом.

Средний уровень моментного ряда с равными интервалами рассчитывается по формуле средней хронологической:

$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + \frac{y_n}{2}}{n-1},$$

где n - число дат.

Средний уровень моментного ряда с неравными интервалами рассчитывается по формуле средней арифметической взвешенной, где в качестве весов берётся продолжительность промежутков времени между временными моментами изменений в уровнях динамического ряда:

$$\bar{y} = \frac{\sum y_t}{\sum t},$$

где t - продолжительность периода (дни, месяцы), в течение которого уровень не изменялся.

Средний абсолютный прирост (средняя скорость роста) определяется как средняя арифметическая из показателей скорости роста за отдельные периоды времени:

$$\overline{\Delta y} = \frac{\sum \Delta y_u}{n},$$

где $\sum \Delta y_u$ - сумма цепных абсолютных приростов;
 n - число приростов.

Или

$$\overline{\Delta y} = \frac{y_n - y_1}{n - 1},$$

где y_n - конечный уровень ряда;
 y_1 - начальный уровень ряда.

Средний коэффициент роста рассчитывается по формуле средней геометрической из показателей коэффициентов роста за отдельные периоды:

$$\overline{Kp} = \sqrt[n]{K_1 \cdot K_2 \cdot \dots \cdot K_n},$$

где K_1, K_2, \dots, K_n - коэффициенты роста по сравнению с предыдущим периодом;
 n - число коэффициентов роста.

Средний коэффициент роста можно определить иначе:

$$\overline{Kp} = \sqrt[n-1]{\frac{y_n}{y_0}},$$

Средний темп роста, %. Это средний коэффициент роста, который выражается в процентах:

$$\overline{Tp} = \overline{Kp} \cdot 100\%$$

Средний темп прироста, %. Для расчёта данного показателя первоначально определяется средний темп роста, который затем уменьшается на 100%. Его также можно определить, если уменьшить средний коэффициент роста на единицу:

$$\overline{T_{\Pi}} = \overline{Tp} - 100\%,$$

или

$$\overline{T_{\Pi}} = (\overline{Kp} - 1) \cdot 100\%,$$

Среднее абсолютное значение 1% прироста можно рассчитать по формуле:

$$\overline{A} = \frac{\overline{\Delta y}}{\overline{T_{\Pi}}},$$

Тема урока: **«ОСНОВНЫЕ КОМПОНЕНТЫ ДИНАМИЧЕСКОГО РЯДА. МЕТОДЫ АНАЛИЗА ОСНОВНОЙ ТЕНДЕНЦИИ В РЯДАХ ДИНАМИКИ»**

1. Способы обработки динамического ряда

В ходе обработки динамического ряда важнейшей задачей является выявление основной тенденции развития явления (тренда) и сглаживание случайных колебаний. Для решения этой задачи в статистике существуют особые способы, которые называют методами выравнивания.

Выделяют три основных способа обработки динамического ряда:

а) укрупнение интервалов динамического ряда и расчет средних для каждого укрупненного интервала;

б) метод скользящей средней;

в) аналитическое выравнивание (выравнивание по аналитическим формулам).

Укрупнение интервалов - наиболее простой способ. Он заключается в преобразовании первоначальных рядов динамики в более крупные по продолжительности временных периодов, что позволяет более чётко выявить действие основной тенденции (основных факторов) изменения уровней.

По интервальным рядам итоги исчисляются путём простого суммирования уровней первоначальных рядов. Для других случаев рассчитывают средние величины укрупнённых рядов (переменная средняя). Переменная средняя рассчитывается по формулам простой средней арифметической.

Скользящая средняя - это такая динамическая средняя, которая последовательно рассчитывается при передвижении на один интервал при заданной продолжительности периода. Если, предположим, продолжительность периода равна 3, то скользящие средние рассчитываются следующим образом:

$$\bar{y}_1 = \frac{y_1 + y_2 + y_3}{3}; \bar{y}_2 = \frac{y_2 + y_3 + y_4}{3}; \bar{y}_3 = \frac{y_3 + y_4 + y_5}{3} \text{ и т.д.}$$

При чётных периодах скользящей средней можно центрировать данные, т.е. определять среднюю из найденных средних. К примеру, если скользящая исчисляется с продолжительностью периода, равной 2, то центрированные средние можно определить так:

$$\bar{y}_1 = \frac{\bar{y}_1 + \bar{y}_2}{2}; \bar{y}_2 = \frac{\bar{y}_2 + \bar{y}_3}{2}; \bar{y}_3 = \frac{\bar{y}_3 + \bar{y}_4}{2} \text{ и т.д.}$$

Первую рассчитанную центрированную относят ко второму периоду, вторую - к третьему, третью - к четвертому и т.д. По сравнению с фактическим, сглаженный ряд становится короче на $(m - 1)/2$, где m - число уровней интервала.

Суть **аналитического выравнивания** динамического ряда заключается в том, что фактические уровни ряда заменяются плавными уровнями, вычисленными на основе определённой линии (прямой или кривой), выбранной в предположении, что она точнее всего отображает общую тенденцию явления.

В основе метода лежит установление функциональной зависимости уровней ряда от времени $Y_t = f(t)$ с использованием корреляционно-регрессионного анализа. При этом на практике чаще всего применяются математические функции такого вида:

а) линейная $Y_t = a_0 + a_1 t$,

б) параболическая $Y_t = a_0 + a_1 t^2$,

в) гиперболическая $Y_t = a_0 + a_1 \frac{1}{t}$,

г) степенная $Y_t = a_0 t^a$,

где a_0, a_1 - параметры, которые находятся методом наименьших квадратов;
 t - порядковый номер периода.

На основе теоретического анализа выявляется характер развития явления во времени, и на этой основе выбирается тот или другой вид аналитической функции. Практикой статистических исследований установлено, что принятие соответствующей аналитической функции осуществляется при таких условиях:

1) выравнивать динамические ряды по уравнению прямой линии целесообразно тогда, когда более или менее постоянны цепные абсолютные приросты, то есть тогда, когда уровни ряда изменяются приблизительно в арифметической прогрессии;

2) выравнивание динамических рядов по уравнению квадратичной параболы необходимо выполнять в тех случаях, когда изменение уровней ряда происходит с приблизительно равномерным ускорением или замедлением цепных абсолютных приростов;

3) выравнивание по степенной функции целесообразно использовать тогда, когда уровни ряда динамики выявляют тенденцию постоянства цепных темпов роста, то есть в случае изменения уровней ряда динамики в геометрической прогрессии.

Расчёт параметров математических функций осуществляется методом наименьших квадратов. Он даёт возможность получить такую зависимость, которая наиболее близко проходит к точкам фактических данных на графике в осях координат «t - y», то есть даёт наименьшую сумму квадратов отклонений фактических значений результативного признака y от уровней (теоретических) значений Y:

$$\sum (y - \hat{y})^2 = \min ,$$

На основании этого условия получают систему нормальных уравнений для расчёта параметров a_0 и a_1 , где в качестве фактора x выступает время t.

Выравнивание рядов динамики по методу наименьших квадратов, как и выравнивание, посредством других приёмов, должно осуществляться в пределах качественно однородных периодов. Если в динамическом ряду есть качественно неоднородные периоды, то выявлять тенденцию целесообразно в пределах каждого из них.

Расчёт параметров a_0 и a_1 можно значительно упростить, если отсчёт времени $t = 0$ осуществлять с середины динамического ряда. Тогда значения t, размещённые выше середины, будут отрицательными, а ниже - положительными. В обоих случаях $\sum t = 0$. Для этого уровень, который будет пребывать в середине ряда динамики, берут за условное начало отсчёта или нулевое значение.

Для того чтобы сумма показателей времени равнялась нулю, условные обозначения нужно давать таким образом:

при нечётном числе уровней ряда динамики, чтобы выполнить условие $\sum t = 0$, уровень, который будет прибывать в середине ряда, приравнивают к нулю, а уровни, расположенные выше его, помечают числами со знаком «минус» (-1; -2; -3 и т. п.), а ниже - числами со знаком «плюс» (+1; +2; +3 и т. д.);

при парном числе уровней ряда динамики уровни, которые лежат выше среднего значения (оно находится в середине между двумя средними датами), помечают натуральными числами со знаком «минус» (-1; -3; -5 и т. п.), а уровни, которые лежат ниже среднего значения, — натуральными числами со знаком «плюс» (+1; +2; +3 и т. д.).

При условии, что $\sum t = 0$, система нормальных уравнений упрощается, приобретая в случае линейной зависимости такой вид:

$$\begin{cases} a_0 n = \sum y \\ a_1 \sum t^2 = \sum ty \end{cases}$$

Откуда

$$a_0 = \frac{\sum y}{n},$$
$$a_1 = \frac{\sum yt}{\sum t^2},$$

Парабола второго порядка ($Y_t = a_0 + a_1t + a_2t^2$) используется для описания рядов динамики, в которых меняется направление развития: со снижения показателей на их рост и наоборот.

Параметр a_2 называется коэффициентом регрессии и характеризует изменение интенсивности развития в единицу времени.

При $a_2 > 0$ наблюдается ускоренное развитие, при $a_2 < 0$ – замедленное.

$$a_0n + a_2 \sum t^2 = \sum y$$
$$a_1 \sum t^2 = \sum yt$$
$$a_0 \sum t^2 + a_2 \sum t^4 = \sum yt^2$$

Отсюда

$$a_1 = \frac{\sum yt}{\sum t^2}.$$
$$a_0 = \frac{\sum t^4 \cdot \sum y - \sum t^2 \cdot \sum yt^2}{n \sum t^4 - (\sum t^2)^2},$$
$$a_2 = \frac{n \sum yt^2 - \sum t^2 \cdot \sum y}{n \sum t^4 - (\sum t^2)^2},$$

Задачей аналитического выравнивания является определение не только общей тенденции развития явления, но и некоторых недостающих значений как внутри периода, так и за его пределами. **Способ определения неизвестных значений внутри динамического ряда называют интерполяцией.** Эти неизвестные значения можно определить:

- 1) используя полусумму уровней, расположенных рядом с интерполируемыми;
- 2) по среднему абсолютному приросту;
- 3) по темпу роста.

Способ определения количественных значений за пределами ряда называют экстраполяцией. Экстраполирование используется для прогнозирова-

ния тех факторов, которые не только в прошлом и настоящем обуславливают развитие явления, но и могут оказать влияние на его развитие в будущем.

Экстраполировать можно по средней арифметической, по среднему абсолютному приросту, по среднему темпу роста.

На практике результат экстраполяции прогнозируемых уровней социально-экономических явлений обычно выполняют интервальными оценками. Для определения границ интервалов используется интервальное неравенство:

$$Y_i - t_g \sigma_\varepsilon \leq Y_{np} \leq Y_i + t_g \sigma_\varepsilon,$$

где t_g - коэффициент доверия g. распределения Стьюдента;

$$Y_e - \text{остаточное среднее квадратическое отклонение } \sigma = \sqrt{\frac{\sum (y - Y_i)^2}{(n - m)}};$$

n - количество уровней рассматриваемого (базисного) ряда динамики;

m - количество параметров геометрической зависимости тренда;

$(n - m)$ – число степеней свободы;

Y_i - дискретное (точечное) значение прогнозного уровня.

Коэффициент доверия выбирается из статистических таблиц распределения Стьюдента в зависимости от числа степеней свободы $(n - m)$ и уровня значимости α (0,01 или 0,05).

Тогда окончательно с вероятностью $p=1-\alpha$ прогнозный уровень тренда в будущем Y_i будет находиться в пределах: верхний предел составит $(Y_i + t_k \sigma_\varepsilon)$, нижний предел - $(Y_i - t_k \sigma_\varepsilon)$.

При аналитическом выравнивании может иметь место **автокорреляция**, под которой понимается **зависимость между соседними членами динамического ряда**. Автокорреляцию можно установить с помощью перемещения уровня на одну дату.

Коэффициент автокорреляции вычисляется по формуле:

$$r_a = \frac{\overline{y_t \cdot y_{t+1}} - \overline{y_t} \cdot \overline{y_{t+1}}}{\sigma_{y_t} \sigma_{y_{t+1}}},$$

2. Изучение сезонных колебаний

Анализ рядов динамики предполагает и исследование сезонной неравномерности (**сезонных колебаний**), под которыми понимают устойчивые внутри-годовые колебания, причиной которых являются многочисленные факторы, в том числе и природно-климатические. Сезонные колебания измеряются с помощью индексов сезонности, которые рассчитываются двумя способами в зависимости от характера динамического развития.

При относительно неизменном годовом уровне явления индекс сезонности можно рассчитать как процентное отношение средней величины из фактических уровней одноименных месяцев к общему среднему уровню за исследуемый период:

$$I_c = \frac{\overline{y_i}}{y_0} \cdot 100\% ,$$

В условиях изменчивости годового уровня индекс сезонности определяется как процентное отношение средней величины из фактических уровней одноимённых месяцев к средней величине из выровненных уровней одноименных месяцев:

$$I_c = \frac{\overline{y_i}}{\hat{y}_i} \cdot 100\% ,$$

Для того, чтобы выявить устойчивую сезонную волну, на которой не отражались бы случайные условия одного года, индексы сезонности вычисляют по данным за несколько лет (не менее трех), распределенным по месяцам.

Совокупность рассчитанных индексов отражает сезонную волну и позволяет определить, каким образом может быть влияние факторов на изменение уровней временного ряда.

«Понятие и классификация индексов»

1. Индексы, их общая характеристика и сфера применения

В статистической практике индексный метод имеет такое же широкое распространение, как и метод средних величин.

Индексами называют сравнительные относительные величины, которые характеризуют изменение сложных социально-экономических показателей (показатели, состоящие из несуммируемых элементов) во времени, в пространстве, по сравнению с планом.

Индекс - это результат сравнения двух одноименных показателей, при исчислении которого следует различать числитель индексного отношения (сравниваемый или отчётный уровень) и знаменатель индексного отношения (базисный уровень, с которым производится сравнение).

Выбор базы зависит от цели исследования. Если изучается динамика, то за базисную величину может быть взят размер показателя в периоде, предшествующем отчётному. Если необходимо осуществить территориальное сравнение, то за базу можно принять данные другой территории. За базу сравнения могут приниматься плановые показатели, если необходимо использовать индексы как показатели выполнения плана.

Индексы формируют важнейшие экономические показатели национальной экономики и её отдельных отраслей. Индексные показатели позволяют осуществить анализ результатов деятельности предприятий и организаций, выпускающих самую разнообразную продукцию или занимающихся различными видами деятельности.

С помощью индексов можно проследить роль отдельных факторов при формировании важнейших экономических показателей, выявить основные резервы производства. Индексы широко используются в сопоставлении международных экономических показателей при определении уровня жизни, деловой активности, ценовой политики и т.д.

Существует два подхода в интерпретации возможностей индексных показателей: обобщающий (синтетический) и аналитический, которые в свою очередь определяются разными задачами.

Суть обобщающего подхода - в трактовке индекса как показателя среднего изменения уровня исследуемого явления. В этом случае основной задачей, решаемой с помощью индексных показателей, будет характеристика общего изменения многофакторного экономического показателя.

Аналитический подход рассматривает индекс как показатель изменения уровня результативной величины, на которую оказывает влияние величина, изучаемая с помощью индекса. Отсюда и иная задача, которая решается с помощью индексных показателей: выделить влияние одного из факторов в изменении многофакторного показателя.

От содержания изучаемых показателей, методологии расчёта первичных показателей, целей и задач исследования зависят и способы построения индексов.

По степени охвата элементов явления индексы делят на индивидуальные и общие (сводные).

Индивидуальные индексы (i) - это индексы, которые характеризуют изменение только одного элемента совокупности.

Общий (сводный) индекс (I) характеризует изменение по всей совокупности элементов сложного явления.

Если индексы охватывают только часть явления, то их называют групповыми.

В зависимости от способа изучения общие индексы могут быть построены или как **агрегатные** (от лат. aggrega - присоединяю) индексы, или как **средние взвешенные индексы** (средние из индивидуальных).

Способ построения агрегатных индексов заключается в том, что при помощи так называемых соизмерителей можно выразить итоговые величины сложной совокупности в отчётном и базисном периодах, а затем первую сопоставить со второй.

В статистике имеют большое значение индексы переменного и фиксированного состава, которые используются при анализе динамики средних показателей.

Индексом переменного состава называют отношение двух средних уровней.

Индекс фиксированного состава есть средний из индивидуальных индексов. Он рассчитывается как отношение двух стандартизованных средних, где влияние изменения структурного фактора устранено, поэтому данный индекс называют ещё индексом постоянного состава.

В зависимости от характера и содержания индексируемых величин различают индексы **количественных (объемных) показателей** и индексы **качественных показателей**.

Классификация индексов представлена на рис. 3.2

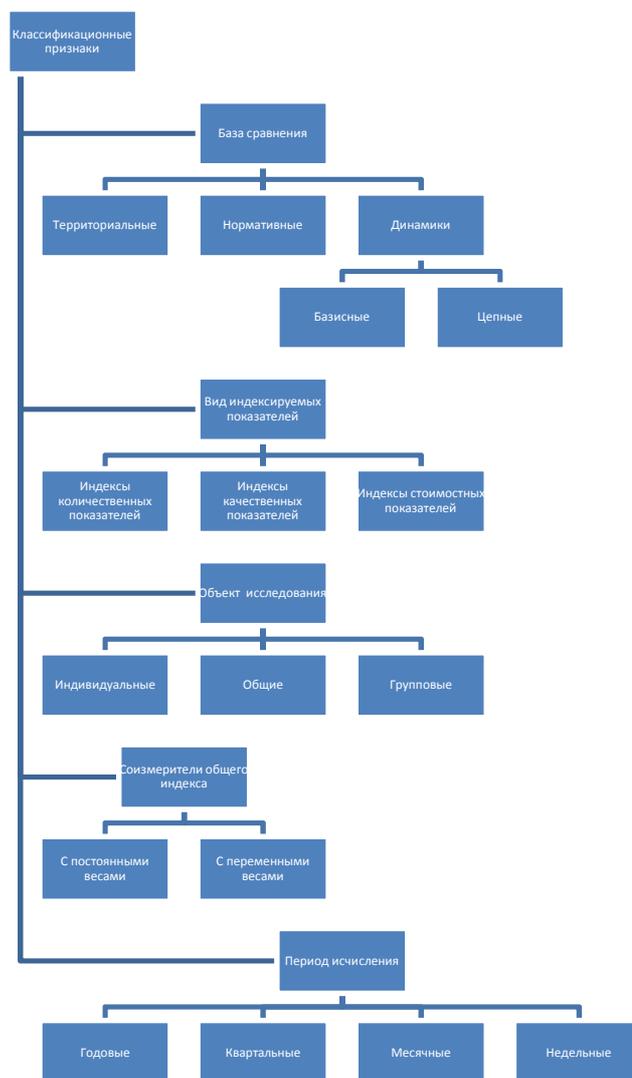


Рис. 16. Виды индексов

2. Индексы количественных показателей

К индексам количественных (объемных) показателей относятся такие индексы, как индексы физического объема производства продукции, затрат на выпуск продукции, стоимости продукции, а также индексы показателей, размеры которых определяются абсолютными величинами. Используются различные виды индексов количественных показателей.

Индекс физического объёма продукции (ФОП) (товарооборота) отражает изменение выпуска (реализации) продукции.

Индивидуальный индекс ФОП отражает изменение выпуска продукции одного вида и определяется по формуле:

$$i_q = q_1 \div q_0,$$

где q_1 и q_0 - количество продукции данного вида в натуральном выражении в текущем и базисном периодах.

Агрегатный индекс ФОП (предложен Э. Ласпейресом) отражает изменение выпуска всей совокупности продукции, где индексируемой величиной является количество продукции q , а соизмерителем - цена p :

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

где q_1 и q_0 - количество выработанных единиц отдельных видов продукции соответственно в отчётном и базисном периодах;

p_0 - цена единицы продукции (отдельного вида) в базисном периоде.

При вычислении индекса ФОП в качестве соизмерителей может выступать также себестоимость продукции или трудоёмкость.

Средние взвешенные индексы ФОП используются в том случае, если известны индивидуальные индексы объёма по отдельным видам продукции и стоимость отдельных видов продукции (или затраты) в базисном или отчётном периоде.

Средний взвешенный арифметический индекс ФОП определяется по формуле:

$$I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0},$$

где i_q - индивидуальный индекс по каждому виду продукции;

$q_0 p_0$ - стоимость продукции каждого вида в базисном периоде.

Средний взвешенный гармонический индекс ФОП:

$$I_q = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_q}},$$

где $q_1 p_1$ - стоимость продукции каждого вида в текущем периоде.

Аналогично рассчитывается индекс затрат на выпуск продукции (ЗВП), который отражает изменение затрат на производство и может быть как индивидуальным, так и агрегатным.

Индивидуальный индекс ЗВП отражает изменение затрат на производство одного вида и определяется по формуле:

$$I_z = \frac{z_1 q_1}{z_0 q_0},$$

где z_1 и z_0 - себестоимость единицы продукции искомого вида в текущем и базисном периодах;

$q_1 z_1$ и $q_0 z_0$ - суммы затрат на выпуск продукции искомого вида в текущем и базисном периодах.

Агрегатный индекс ЗВП характеризует изменение общей суммы затрат на выпуск продукции за счёт изменения количества выработанной продукции и её себестоимости и определяется по формуле:

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_0},$$

где $q_1 z_1$ и $q_0 z_0$ - затраты на выпуск продукции каждого вида соответственно в отчетном и базисном периодах.

Рассмотрим построение **индекса стоимости продукции** (СП), который может определяться и как индивидуальный, и как агрегатный.

Индивидуальный индекс СП характеризует изменение стоимости продукции данного вида и имеет вид:

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0},$$

где p_1 и p_0 - цена единицы продукции данного вида в текущем и базисном периодах;

$q_1 p_1$ и $q_0 p_0$ - стоимость продукции данного вида в текущем и базисном периодах.

Агрегатный индекс СП (товарооборота) характеризует изменение общей стоимости продукции за счёт изменения количества продукции и цен и определяется по формуле:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0},$$

3. Индексы качественных показателей. Факторный анализ

Качественные показатели определяют уровень исследуемого итогового показателя и определяются путём соотношения итогового показателя и определённого количественного показателя (например, средняя заработная плата определяется путём соотношения фонда заработной платы и количества работников).

К индексам качественных показателей относятся индексы цен, себестоимости, средней заработной платы, производительности труда.

Самым распространённым индексом в этой группе является индекс цен.

Индивидуальный индекс цен характеризует изменение цен по одному виду продукции и определяется по формуле:

$$i_p = p_1 \div p_0,$$

где p_1 и p_0 - цена за единицу продукции в текущем и базисном периодах.

Соответственно определяются индексы себестоимости и затрат рабочего времени по каждому виду продукции.

Агрегатный индекс цен определяет среднее изменение цены p по совокупности определённых видов продукции q .

Для характеристики среднего изменения цен на потребительские товары используют индекс цен, предложенный Э. Ласпейресом (индекс Ласпейреса):

$$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0},$$

где q_0 - потребительская корзина (базовый период);

p_0 и p_1 - соответственно цены базисного и отчётного периодов.

Если количество набора продуктов принимается на уровне отчётного периода (q_1), то в этом случае индекс цен именуется индексом Пааше:

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

Если известны индивидуальные индексы цен по отдельным видам продукции и стоимость отдельных видов продукции, то применяются средние взвешенные индексы цен (средний взвешенный арифметический и средний взвешенный гармонический индексы цен).

Формула *среднего взвешенного арифметического индекса цен*:

$$I_p = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0},$$

где i - индивидуальный индекс по каждому виду продукции;
 p_0q_0 - стоимость продукции каждого вида в базисном периоде.

Формула *среднего взвешенного гармонического индекса цен*:

$$I_p = \frac{\sum p_1 q_1}{\sum \frac{p_1 q_1}{i_p}},$$

где $p_1 q_1$ - стоимость продукции каждого вида в текущем периоде.

В статистической практике очень широко используется **агрегатный территориальный индекс цен**, который может быть рассчитан по следующей формуле:

$$I_{A/B} = \frac{\sum p_A q_A}{\sum p_B q_A},$$

где $p_A p_B$ - цена за единицу продукции каждого вида соответственно на территории А и В;

q_A - количество выработанной или реализованной продукции каждого вида по территории А (в натуральном выражении).

Из формулы видно, что в данном индексе в качестве фиксированного показателя (веса) принят объём продукции территории А. При расчёте данного индекса в качестве веса можно принять также объём продукции территории В или суммарный объём продукции двух территорий.

Возможны два способа расчёта индексов: цепной и базисный.

Цепные индексы получают путём сопоставления текущих уровней с предшествующим, при этом база сравнения постоянно меняется.

Базисные индексы получают путём сопоставления с тем уровнем периода, который был принят за базу сравнения.

В качестве примера можно привести цепные и базисные индексы цен.

Цепные индивидуальные индексы цен имеют следующий ряд расчёта:

$$i_{p_{1/0}} = \frac{p_1}{p_0}, \quad i_{p_{2/1}} = \frac{p_2}{p_1}, \quad i_{p_{3/2}} = \frac{p_3}{p_2} \text{ и т.д.}$$

Базисные индивидуальные индексы цен:

$$i_{p_{1/0}} = \frac{p_1}{p_0}, \quad i_{p_{2/0}} = \frac{p_2}{p_0}, \quad i_{p_{3/0}} = \frac{p_3}{p_0}, \text{ и т.д.}$$

Следует помнить, что *произведение цепных индивидуальных индексов цен равно последнему базисному индексу*.

Цепные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad I_{p_{2/1}} = \frac{\sum p_2 q_2}{\sum p_1 q_2}; \quad I_{p_{3/2}} = \frac{\sum p_3 q_3}{\sum p_2 q_3}; \text{ и т.д.}$$

Базисные агрегатные индексы цен:

$$I_{p_{1/0}} = \frac{\sum p_1 q_1}{\sum p_0 q_1}; \quad I_{p_{2/0}} = \frac{\sum p_2 q_2}{\sum p_0 q_2}; \quad I_{p_{3/0}} = \frac{\sum p_3 q_3}{\sum p_0 q_3}; \text{ и т.д.}$$

Между индексами существует также взаимосвязь и взаимозависимость, как и между самими экономическими явлениями, что позволяет проводить факторный анализ. Благодаря индексному методу можно рассматривать все факторы независимо друг от друга, что даёт возможность определить размер абсолютного изменения сложного явления за счёт каждого фактора в отдельности.

Для выявления роли каждого фактора в отдельности индекс сложного показателя разлагают на частные (факторные) индексы, которые характеризуют роль каждого фактора. При этом используют два метода:

- 1) метод обособленного изучения факторов;
- 2) последовательно-цепной метод.

При первом методе сложный показатель берётся с учётом изменения лишь того фактора, который взят в качестве исследуемого, все остальные остаются неизменными на уровне базисного периода.

Последовательно-цепной метод предполагает использование системы взаимосвязанных индексов, которая требует определённого расположения факторов. Как правило, на первом месте в цепи располагают качественный фактор.

При определении влияния первого фактора все остальные сохраняются в числителе и знаменателе на уровне базисного периода, при определении второго факторного индекса первый фактор сохраняется на уровне базисного периода, а третий и все последующие - на уровне отчётного периода, при определении третьего факторного индекса первый и второй факторы сохраняются на уровне базисного периода, четвёртый и все остальные - на уровне отчётного периода и т.д.

4. Индекс постоянного состава. Индекс переменного состава.

Индекс структурных сдвигов

Индексы переменного состава характеризуют изменение средних величин, то есть сложных качественных показателей (средней цены, средней себестоимости по однородной продукции, производительности труда и т. д.):

$$I_p = \frac{\sum p_1 q_1}{\sum q_1} \div \frac{\sum p_0 q_0}{\sum q_0} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum p_1 d_1}{\sum p_0 d_0},$$

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} \div \frac{\sum z_0 q_0}{\sum q_0} = \frac{\bar{z}_1}{z_0} = \frac{\sum z_1 d_1}{\sum z_0 d_0},$$

где d - относительная величина структуры (удельный вес производства однородной продукции на отдельном участке в общем объеме производимой продукции).

Индекс постоянного состава выражает степень влияния на изменение средней величины только изменения отдельных величин усредняемого показателя при условии элиминирования влияния структурных сдвигов;

$$I_p = \frac{\sum p_1 q_1}{\sum q_1} \div \frac{\sum p_0 q_1}{\sum q_1} = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 d_1}{\sum p_0 d_1}$$

Индекс структурных сдвигов выражает степень влияния структуры экономических явлений на изменение средних уровней при условии неизменности индивидуальных показателей признака.

$$I_{\text{структ. сдвигов}} = \frac{\sum p_0 q_1}{\sum q_1} \div \frac{\sum p_0 q_0}{\sum q_0} = \frac{\sum p_0 d_1}{\sum p_0 d_0}$$

Индексы структурных сдвигов удобнее рассчитывать в табличной форме, а также на основе взаимосвязи индексов.

Очень часто приходится сравнивать данные за ряд периодов. В этих случаях следует правильно выбрать базу сравнения. В зависимости от базы сравнения различают индексы с постоянной базой (базисные) и переменной базой сравнения (цепные). Базисные и цепные индексы могут быть индивидуальными и общими. Они представляют собой разновидность относительных величин динамики.

«ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ. СПОСОБЫ ФОРМИРОВАНИЯ ВЫБОРОЧНОЙ СОВОКУПНОСТИ. МЕТОДЫ ОЦЕНКИ РЕЗУЛЬТАТОВ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ»

1. Понятие о выборочном наблюдении

Выборочное наблюдение относится к разновидности несплошного наблюдения. Оно охватывает отобранную часть единиц генеральной совокупности.

Цель выборочного наблюдения - по отобранной части единиц дать характеристику всей совокупности единиц.

Чтобы отобранная часть была репрезентативна (т.е. представляла всю совокупность единиц), выборочное наблюдение должно быть специально организовано. Следовательно, в отличие от генеральной совокупности, представляющей всю совокупность исследуемых единиц, выборочная совокупность представляет ту часть единиц генеральной совокупности, которая является объектом непосредственного наблюдения.

Выборочный метод может широко использоваться органами государственной статистики. Он позволяет при значительной экономии средств и затрат получать необходимую достоверную информацию. Гарантия репрезентативности обеспечивается применением научно обоснованных способов отбора единиц, которые подлежат обследованию.

Следует сразу же иметь в виду, что при сопоставлении показателей по результатам выборочного исследования с характеристиками для всей генеральной совокупности могут иметь место отклонения. Величина этих отклонений называется ошибкой наблюдения, которая может быть или ошибкой регистрации (несовершенство технических условий), или ошибкой репрезентативности (случайное или систематическое нарушение правил при отборе единиц).

В статистике приняты следующие условные обозначения:

μ_x - средняя ошибка репрезентативности при определении среднего размера изучаемого признака;

μ_p - средняя ошибка репрезентативности при определении доли данного признака;

N - объём генеральной совокупности;

n - объём выборочной совокупности;

\bar{X} - средняя в генеральной совокупности;

\tilde{X} - средняя в выборочной совокупности;

p - доля единиц в генеральной совокупности;

w - доля единиц в выборочной совокупности;

σ^2 - дисперсия в выборке;

σ^2 - средняя из групповых дисперсий;

$w_i(1 - w_i)$ - средняя долей из групповых дисперсий;

$\delta^2 = \frac{\sum (x - \tilde{x})^2}{S}$ - межсерийная выборочная дисперсия;

S - число серий в генеральной совокупности;

s - число отобранных серий;

$w_s(1 - w_s)$ - межсерийная выборочная дисперсия долей.

Δ - предельная ошибка выборки.

2. Виды выборки, способы отбора и ошибки выборочного наблюдения

По способу отбора (способу формирования) выборки единиц из генеральной совокупности распространены следующие виды выборочного наблюдения:

* простая случайная выборка (собственно-случайная);

* типическая (стратифицированная);

- * серийная (гнездовая);
- * механическая;
- * комбинированная;
- * ступенчатая.

Простая случайная выборка (собственно-случайная) есть отбор единиц из генеральной совокупности путём случайного отбора, но при условии вероятности выбора любой единицы из генеральной совокупности. Отбор проводится методом жеребьёвки или по таблице случайных чисел.

Типическая (стратифицированная) выборка предполагает разделение неоднородной генеральной совокупности на типологические или районированные группы по какому-либо существенному признаку, после чего из каждой группы производится случайный отбор единиц.

Для **серийной (гнездовой) выборки** характерно то, что генеральная совокупность первоначально разбивается на определённые равновеликие или неравновеликие серии (единицы внутри серий связаны по определённому признаку), из которых путём случайного отбора отбираются серии и затем внутри отобранных серий проводится сплошное наблюдение.

Механическая выборка представляет собой отбор единиц через равные промежутки (по алфавиту, через временные промежутки, по пространственному способу и т.д.). При проведении механического отбора генеральная совокупность разбивается на равные по численности группы, из которых затем отбирается по одной единице.

Комбинированная выборка основана на сочетании нескольких способов выборки.

Многоступенчатая выборка есть образование внутри генеральной совокупности вначале крупных групп единиц, из которых образуются группы, меньшие по объёму, и так до тех пор, пока не будут отобраны те группы или отдельные единицы, которые необходимо исследовать.

Выборочный отбор может быть повторным и бесповторным. При повторном отборе вероятность выбора любой единицы не ограничена. При бесповторном отборе выбранная единица в исходную совокупность не возвращается.

Для отобранных единиц рассчитываются обобщённые показатели (средние или относительные) и в дальнейшем результаты выборочного исследования распространяются на всю генеральную совокупность.

При выборочном наблюдении решаются две задачи:

- 1) определение среднего размера изучаемого признака;
- 2) определение доли единиц, обладающих данным признаком.

Для определения средней ошибки выборки (ρ) применяются следующие формулы (табл. 13.):

Наименование показателей	Повторная выборка	Бесповторная выборка
Собственный случайный и механический отбор		
а) при определении среднего размера изучаемого признака	$\mu_x = \sqrt{\frac{\sigma^2}{n}}$	$\mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
б) при определении доли	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$
Типический отбор		
а) при определении среднего размера изучаемого признака	$\mu_x = \sqrt{\frac{\sigma_i^2}{n}}$	$\mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$
б) при определении доли данного признака	$\mu_p = \sqrt{\frac{w_i(1-w_i)}{n}}$	$\mu_p = \sqrt{\frac{w_i(1-w_i)}{n} \left(1 - \frac{n}{N}\right)}$
Серийный отбор		
а) при определении среднего размера изучаемого признака	$\mu_x = \sqrt{\frac{\delta^2}{S}}$	$\mu_x = \sqrt{\frac{\delta^2}{S} \left(\frac{S-s}{S-1}\right)}$
б) при определении доли	$\mu_p = \sqrt{\frac{w_s(1-w_s)}{S}}$	$\mu_p = \sqrt{\frac{w_s(1-w_s)}{S} \left(\frac{S-s}{S-1}\right)}$

Предельная ошибка выборки (Δ) определяется по формулам:

$$\Delta_x = t \times \mu_x,$$

$$\Delta_p = t \times \mu_p,$$

где t - коэффициент доверия.

В статистических исследованиях с помощью формулы предельной ошибки можно решать ряд задач.

1. Определять возможные пределы нахождения характеристики генеральной совокупности на основе данных выборки.

Доверительные интервалы для генеральной средней можно установить на основе соотношений:

$$\bar{X} = \tilde{X} \pm \Delta_x \quad \text{или} \quad \tilde{X} + \Delta_x \geq \bar{X} \geq \tilde{X} - \Delta_x,$$

Доверительные интервалы для генеральной доли устанавливаются на основе соотношений:

$$P = W \pm \Delta_p \quad \text{или} \quad W + \Delta_p \geq P \geq W - \Delta_p,$$

2. Определять доверительную вероятность, которая означает, что характеристика генеральной совокупности отличается от выборочной на заданную величину.

Доверительная вероятность является функцией от t , где:

$$t = \frac{\Delta_x}{\mu_{\bar{x}}}$$

Доверительная вероятность по величине t определяется по специальной таблице.

При обобщении результатов выборочного наблюдения наиболее часто используют следующие уровни вероятности и соответствующие им значения t :

P	0,683	0,950	0,954	0,997
t	1	1,96	2	3

3. Определять необходимый объем выборки с помощью допустимой величины ошибки.

Определение необходимой численности выборки (n) производится на основе алгебраического преобразования формы предельных ошибок выборки.

1. При определении среднего размера признака

$$n = \frac{t^2 \sigma^2}{\Delta_x^2} \text{ - повторный отбор,}$$

$$n = \frac{t^2 \sigma^2 N}{\Delta_x^2 N + t^2 \sigma^2} \text{ - бесповторный отбор}$$

2. При определении доли признака

$$n = \frac{t^2 w(1-w)}{\Delta_w^2} \text{ - повторный отбор,}$$

$$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)} \text{ - бесповторный отбор.}$$

3. Методы распространения выборочного наблюдения на генеральную совокупность

Основными методами распространения выборочного наблюдения на генеральную совокупность являются прямой пересчёт и способ коэффициентов.

Прямой пересчёт есть произведение среднего значения признака на объём генеральной совокупности. Однако большое число факторов не позволяет в полной мере использовать точечную оценку прямого пересчёта при распространении результатов выборки на генеральную совокупность.

На практике чаще пользуются интервальной оценкой, которая даёт возможность учитывать размер предельной ошибки выборки, которая рассчитана для средней или для доли признака.

Способ коэффициентов используется в тех случаях, когда выборочное наблюдение проводится для проверки и уточнения данных сплошного наблюдения.

При этом рекомендуется использовать формулу:

$$Y_1 = Y_0 \frac{y_1}{y_0},$$

где Y_1 - численность совокупности с поправкой на недоучёт;

Y_0 - численность совокупности без этой поправки;

y_0 - численность совокупности в контрольных точках по первоначальным данным;

y_1 - численность совокупности в тех же точках по данным контрольных мероприятий.

Если нужно уточнить данные сплошного наблюдения при осуществлении контроля за выборочными исследованиями, необходимо определить поправку на недоучёт.

Метод расчёта этой поправки широко применяется при исследовании небольших совокупностей, когда можно рассчитать коэффициент недоучёта по каждой категории работников и, уточнив данные, распространить результаты на всю совокупность.

Например: имеются данные о количестве скота, находящегося в личном пользовании согласно переписи, а также согласно контрольному обходу (табл. 14).

Чтобы определить процент недоучёта, нужно найти разность между данными контрольного обхода и данными сплошного наблюдения, а затем полученную величину разделить на данные сплошного наблюдения.

При переписи недоучтено:

Коров $863 - 850 - 6 + 2 = 9$;

Нетелей $144 - 140 - 4 + 1 = 1$;

Тёлок $87 - 80 - 2 = 5$.

Данные контрольного обхода о количестве тёлок сопоставляют с данными переписи.

Отсюда коэффициент недоучёта коров равен $9 \cdot 100 / 850 = 1,06\%$;

Коэффициент недоучёта нетелей - $1 \cdot 100 / 140 = 0,72\%$;

Коэффициент недоучёта тёлков - $5 \cdot 100 / 80 = 6,25 \%$.

Полученные результаты выборочного наблюдения (проценты недоучёта) распространяются на всю совокупность.

Для этого поправочные коэффициенты (проценты недоучёта) умножаем на данные сплошного наблюдения, полученные в результате переписи скота во всех хозяйствах (табл. 15).

Таблица 14

Количество скота, находящегося в индивидуальном пользовании населения (цифры условные)

Группа скота	Учтено во всех хозяйствах по переписи	Учтено в хозяйствах, подвергнутых контрольному обходу		За время прошедшее от переписи до контрольного обхода в хозяйствах, подвергнутых контрольному обходу	
		по переписи	при контрольном обходе	прибыло	убыло
А	1	2	3	4	5
Коровы	9200	850	863	6	2
Нетели и тёлки, рождённые в прошлом году и старше	1200	140	144	4	1
Тёлки, рождённые в этом году	800	80	87	2	-
Итого	11200	1070	1094	12	3

Таблица 15

Расчёт фактического количества поголовья скота при помощи поправочных коэффициентов (процент недоучёта)

Группы скота	Поправочные коэффициенты	Учтено во всех хозяйствах	Количество скота с поправкой на данные выборочного обследования
А	1	2	3
Коровы	1,06	9200	9752
Нетели	0,72	1200	864
Тёлки	6,25	800	5000
Итого	-	11200	15616

Тема урока: «МЕТОДЫ ИЗУЧЕНИЯ СВЯЗИ МЕЖДУ ЯВЛЕНИЯМИ»

1. Сущность корреляционной связи

Важнейшей целью статистики является изучение объективно существующих связей между явлениями. В ходе статистического исследования этих связей необходимо выявить причинно-следственные зависимости между показателями, т.е. насколько изменение одних показателей зависит от изменения других показателей.

Существует две категории зависимостей (функциональная и корреляционная) и две группы признаков (признаки-факторы и результативные признаки). В отличие от функциональной связи, где существует полное соответствие между факторными и результативными признаками, в корреляционной связи отсутствует это полное соответствие.

Корреляционная связь - это связь, где воздействие отдельных факторов проявляется только как тенденция (в среднем) при массовом наблюдении фактических данных.

Примерами корреляционной зависимости могут быть зависимости между размерами активов банка и суммой прибыли банка, ростом производительности труда и стажем работы сотрудников.

2. Корреляционно-регрессионный метод анализа

Наиболее простым вариантом корреляционной зависимости является парная корреляция, т.е. зависимость между двумя признаками (результативным и факторным или между двумя факторными).

Математически эту зависимость можно выразить как зависимость результативного показателя y от факторного показателя x .

Связи могут быть прямые и обратные. В первом случае с увеличением признака x увеличивается и признак y , при обратной связи с увеличением признака x уменьшается признак y .

Важнейшей задачей является определение формы связи с последующим расчётом параметров уравнения, или, иначе, нахождение уравнения связи (уравнения регрессии).

Могут иметь место различные формы связи:

- прямолинейная

$$Y = a_0 + a_1x,$$

Параметр a_1 называется коэффициентом регрессии и показывает, насколько в среднем отклоняется величина, результативного признака y при отклонении величины факторного признака x на одну единицу.

криволинейная в виде:

- параболы второго порядка (или высших порядков):

$$Y = a_0 + a_1x + a_2x^2,$$

Параметр a_2 характеризует степень ускорения или замедления кривизны параболы и при $a_2 > 0$ парабола имеет минимум, а при $a_2 < 0$ – максимум. Параметр a_1 характеризует кривизну кривой, а параметр a_0 – вершину кривой.
- гиперболы:

$$Y = a_0 + a_1 \frac{1}{x},$$

и т.д.

Точное аналитическое выражение имеет только функциональная связь. Корреляционная связь может быть выражена лишь приближённо, при наличии определённых условий.

Параметры для всех этих уравнений связи, как правило, определяют из системы нормальных уравнений, которые должны отвечать требованию метода наименьших квадратов (МНК), чтобы $\sum (y_i - Y)^2 = \min$:

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases},$$

Если связь выражена параболой второго порядка ($Y = a_0 + a_1x + a_2x^2$), то систему нормальных уравнений для отыскания параметров a_0 , a_1 , a_2 (такую связь называют множественной, поскольку она предполагает зависимость более чем двух факторов) можно представить в виде системы:

$$\begin{cases} a_0n + a_1 \sum x + a_2 \sum x^2 = \sum y \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2y \end{cases},$$

Рассмотрим применение приёмов корреляционного анализа на конкретном примере.

Анализ данных табл. 16 показывает, что с увеличением стоимости основных фондов растёт, как правило, и выпуск продукции. Однако, мы не можем утверждать, что увеличение стоимости основных фондов, например, на 1 млн. руб., приводит к фактическому увеличению выпуска продукции на определённую сумму.

Таблица 16

Стоимость основных фондов и выпуск продукции по группе предприятий

Название предприятий	Стоимость основных фондов, млн. руб. (x)	Выпуск продукции, млн. руб. (y)	xy	x ²	Y
A	1	2	3	4	5
1 ООО «Слава»	6	2,4	14,4	36	2,692
2 ООО «Лидер»	8	4,0	32,0	64	3,537
3 ООО «Олимп»	9	3,6	32,4	81	3,958
4 ООО «СОМ»	10	4,0	40,0	100	4,380
5 ООО «Сюзи»	10	4,5	45,0	100	4,380
6 ООО «Престиж»	11	4,6	50,6	121	4,802
7 ООО «Тандем»	12	5,6	67,2	144	5,224
8 ООО «Рубин»	13	6,5	84,5	169	5,646
9 ООО «Злата»	14	7,0	98,0	196	6,068
10 ООО «Вернисаж»	15	5,0	75,0	225	6,490
Итого	108	47,2	539,1	1236	47,177

Чтобы установить, насколько повышается в среднем выпуск продукции при увеличении основных фондов на 1 млн. руб., прежде всего, определим форму связи.

Допустим, что между стоимостью основных фондов и выпуском продукции существует прямолинейная связь, которая выражается уравнением прямой $Y = a_0 + a_1x$. Параметры уравнения определим при помощи системы двух нормальных уравнений, отвечающих требованию способа наименьших квадратов.

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases}$$

$$\begin{cases} 10a_0 + 108a_1 = 47,2 \\ 108a_0 + 1236a_1 = 539,1 \end{cases}$$

Решим систему нормальных уравнений, для чего каждый член обоих уравнений поделим на коэффициенты при a_0 и из второго уравнения вычтем первое:

$$4,47 = a_0 + 10,8a_1$$

$$4,99 = a_0 + 11,44a_1$$

Определим параметр a_1 : $a_1 = 0,27 / 0,64 = 0,422$.

Подставим значение a_1 в первое уравнение и найдём параметр a_0 : $4,72 = a_0 + 10,8 \cdot 0,422$, откуда $a_0 = 4,72 - 4,56 = 0,16$.

Линейное уравнение корреляционной связи будет иметь следующий вид: $Y = 0,16 + 0,422x$. Параметр a_1 показывает, что с увеличением

стоимости основных фондов в среднем на 1 млн. руб. выпуск продукции увеличивается в среднем на 0,422 млн. руб. Параметр a_0 - свободный член уравнения, $a_0 = 0,16$, когда $x = 0$.

Подставляем значения параметров a_0 и a_1 в уравнение прямой $Y = 0,16 + 0,422x$ и находим теоретические, выровненные значения Y .

$$Y_1 = 0,16 + 0,422 \cdot 6 = 2,692,$$

$$Y_2 = 0,16 + 0,422 \cdot 8 = 3,537 \text{ и т.д. (см. табл. 16 графа 5).}$$

Графически зависимость выпуска продукции от стоимости основных фондов показана на рис. 17

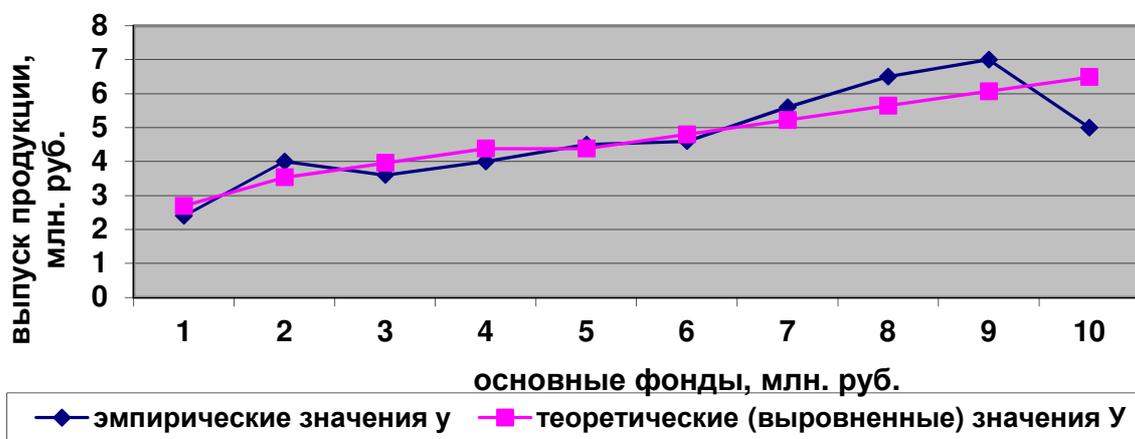


Рис. 17 Зависимость выпуска продукции от стоимости основных фондов по 10 предприятиям

Если в результате качественного анализа установлена криволинейная зависимость, принимающая форму кривой второго порядка, то связь выражается уравнением кривой $Y = a_0 + a_1x + a_2x^2$. Задача сводится к нахождению параметров a_0 , a_1 и a_2 . Для этого необходимо решить систему нормальных уравнений:

$$\begin{cases} a_0n + a_1 \sum x + a_2 \sum x^2 = \sum y \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum x^2y \end{cases}$$

Пример. Имеются данные о возрасте и выработке по группе рабочих предприятия «А».

Возраст, лет (x)	18-22	23-27	28-32	33-37	38-42	43-47	48-52	53-50
Выработка деталей на 1-го рабочего, шт.	5	6	7	8	10	8	6	5

Для решения системы нормальных уравнений составим расчётную таблицу (табл. 3.5).

Таблица 17

Определение зависимости выработки рабочих предприятия «А» от возраста

x	y	xy	x^2	x^3	x^4	x^2y	Y
1	2	3	4	5	6	7	8
20	5	100	400	8000	160000	2000	5,690
25	6	150	625	15625	390625	3750	6,600
30	7	210	900	27000	810000	6300	7,225
35	8	280	1225	42875	1500625	9800	7,565
40	10	400	1600	64000	2560000	16000	7,620
45	8	360	2025	91125	4100625	16200	7,390
50	6	300	2500	125000	6250000	15000	6,875
55	5	275	3025	166375	9150625	15125	6,075
Итого 300	55	2075	12300	540000	24922500	84175	55,04

Подставим данные таблицы в систему нормальных уравнений:

$$\begin{cases} 8a_0 + 300a_1 + 12300a_2 = 55 \\ 300a_0 + 12300a_1 + 540000a_2 = 2075 \\ 12300a_0 + 540000a_1 + 24922500a_2 = 84175 \end{cases}$$

Поделим каждый член уравнения на коэффициенты при a_0 и получим следующее уравнение:

$$\begin{cases} a_0 + 37,5a_1 + 12300a_2 = 6,88 \\ a_0 + 41,0a_1 + 1800,0a_2 = 6,92 \\ a_0 + 43,9a_1 + 2026,2a_2 = 6,84 \end{cases}$$

Вычтем из второго уравнения первое, из третьего – второе и поделим каждый член уравнений на коэффициент при a_1 :

$$\begin{aligned} 3,5a_1 + 262,5a_2 &= 0,04 & a_1 + 75a_2 &= +0,011 \\ 2,9a_1 + 226,2a_2 &= -0,08 & a_1 + 78a_2 &= -0,028 \end{aligned}$$

Вычтем теперь из второго уравнения первое и получим:

$$- 0,017$$

$$- 0,017 = 3a_2, \text{ откуда } a_2 = -0,017/3 = -0,0057.$$

Подставим в уравнение значение:

$$a_2 : 0,011 = a_1 + 75(-0,0057), \text{ откуда } a_1 = 0,4275 + 0,011 = 0,4385.$$

Методом подстановки получаем значение a_0 :

$$6,88 = a_0 + 37,5 \cdot 0,4385 + 1537,5(-0,0057);$$

$$6,88 = a_0 + 16,44375 - 8,76375, \text{ откуда } a_0 = -0,8.$$

Теперь можно записать уравнение параболы:

$$Y = -0,8 + 0,4385x - 0,0057x^2$$

Отрицательное значение a_2 показывает, что после определённого возраста (в данном случае 43 – 47 лет) выработка рабочих начинает снижаться.

Определим теоретические (выровненные) значения Y для чего в уравнение кривой подставим значения x :

$$Y_1 = -0,8 + 0,4385 \cdot 20 - 0,0057 \cdot 400 = 5,690;$$

$$Y_2 = -0,8 + 0,4385 \cdot 25 - 0,0057 \cdot 625 = 6,600 \text{ и т.д. (см табл. 17 графа 8).}$$

Графически зависимость выработки деталей от возраста рабочих представлена на рис. 18.

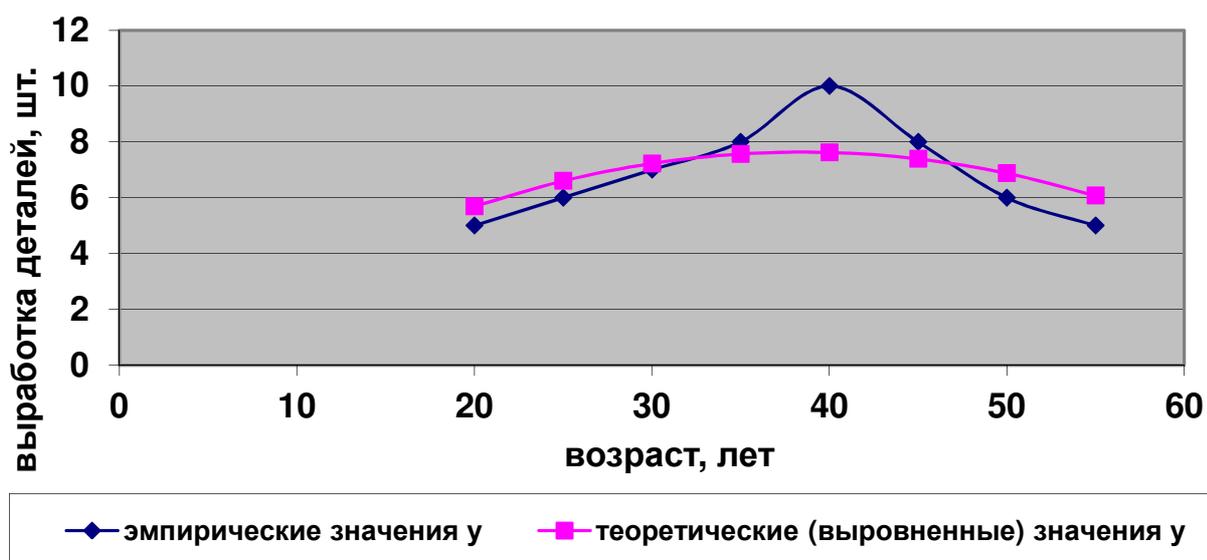


Рис. 18. Зависимость выработки деталей от возраста рабочих предприятия «А»

Другая важнейшая задача - измерение тесноты зависимости - для всех форм связи может быть решена при помощи вычисления эмпирического корреляционного отношения η :

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

где $\delta^2 = \frac{\sum (\bar{y}_i - \bar{y})^2}{n}$ дисперсия в ряду выровненных значений результативного показателя \bar{y}_i ;

$$\sigma^2 = \frac{\sum (y - \bar{y})^2}{n} - \text{дисперсия в ряду фактических значений } y.$$

Для определения степени тесноты парной линейной зависимости служит линейный коэффициент корреляции r , для расчёта которого можно использовать следующие формулы:

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{n \sigma_x \sigma_y},$$

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y},$$

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2}},$$

$$r = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}}$$

Линейный коэффициент корреляции может принимать значения в пределах от -1 до +1 или по модулю от 0 до 1. Чем ближе он по абсолютной величине к 1, тем теснее связь. Знак указывает направление связи: «+» - прямая зависимость, «-» имеет место при обратной зависимости.

Пример. Рассмотрим вычисление коэффициента корреляции по стоимости основных фондов и выпуску продукции по 10 предприятиям (табл. 19).

$$\bar{x} = \frac{\sum x}{n} = \frac{108}{10} = 10,8; \quad \bar{y} = \frac{\sum y}{n} = \frac{47,2}{10} = 4,72$$

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2}} = \frac{29,340}{\sqrt{69,6 \cdot 16,956}} = \frac{29,34}{\sqrt{1180,1376}} = \frac{29,34}{34,35} = 0,854$$

или

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y},$$

где $\overline{xy} = \frac{\sum xy}{n} = \frac{539,1}{10} = 53,91;$

σ_y - среднее квадратическое отклонение результативного признака:

$$\sigma_y = \sqrt{\frac{\sum (y - \bar{y})^2}{n}} = \sqrt{\frac{16,956}{10}} = \sqrt{1,6956} = 1,302;$$

σ_x - среднее квадратическое отклонение факторного признака:

$$\sigma_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{69,6}{10}} = \sqrt{6,96} = 2,638.$$

Подставим необходимые данные в формулу и получим:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y} = \frac{53,91 - 10,8 \cdot 4,72}{1,302 \cdot 2,638} = \frac{2,934}{3,435} = 0,854$$

Таблица 19

Стоимость основных фондов и выпуск продукции по 10 предприятиям

Название предприятий	Стоимость основных фондов, млн. руб. (x)	Выпуск продукции, млн. руб. (y)	$(x - \bar{x})$	$(y - \bar{y})$	$\frac{(x - \bar{x}) \cdot (y - \bar{y})}{(y - \bar{y})}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	xy
А	1	2	3	4	5	6	7	8
«Слава»	6	2,4	- 4,8	- 2,32	11,136	23,04	5,3824	14,4
«Лидер»	8	4,0	- 2,8	- 0,72	2,016	7,84	0,5184	32,0
«Олимп»	9	3,6	- 1,8	- 1,12	2,016	3,24	1,2544	32,4
«СОМ»	10	4,0	- 0,8	- 0,72	0,576	0,64	0,5184	40,0
«Сюзи»	10	4,5	- ,08	- 0,22	0,176	0,64	0,0484	45,0
«Престиж»	11	4,6	0,2	- 0,12	- 0,024	0,04	0,0144	50,6
«Тандем»	12	5,6	1,2	0,88	1,056	1,44	0,7744	67,6
«Рубин»	13	6,5	2,2	1,78	3,916	4,84	3,1684	84,5
«Злата»	14	7,0	3,2	2,28	7,296	10,24	5,1984	98,0
«Вернисаж»	15	5,0	4,2	0,28	1,176	17,64	0,0784	75,0
Итого	108	47,2	-	-	29,34	69,60	16,956	539,1

Таким образом, связь между стоимостью основных фондов и выпуском продукции прямая и высокая.

3. Непараметрические показатели связи

В статистической практике могут встречаться такие случаи, когда качества факторных и результативных признаков не могут быть выражены численно. Поэтому для измерения тесноты зависимости необходимо использовать другие показатели. Для этих целей используются так называемые непараметрические методы.

Наибольшее распространение имеют ранговые коэффициенты корреляции, в основу которых положен принцип нумерации значений статистического ряда. При использовании коэффициентов корреляции рангов коррелируются не сами значения показателей x и y , а только номера их мест, которые они занимают в каждом ряду значений. В этом случае номер каждой отдельной единицы будет её рангом. Если значения признака совпадают, то определяется средний ранг путём деления суммы рангов на число значений.

Коэффициенты корреляции, основанные на использовании ранжированного метода, были предложены К. Спирменом и М. Кендэллом.

Коэффициент корреляции рангов Спирмена (r) основан на рассмотрении разности рангов значений результативного и факторного признаков и может быть рассчитан по формуле:

$$P = \frac{6 \sum d^2}{n(n^2 - 1)},$$

где $d = (R_x - R_y)$, т.е. разность рангов каждой пары значений x и y ,
 n – число наблюдений.

Коэффициент колеблется от -1 до +1. Если ранги по обоим признакам совпадают, то $\sum d^2 = 0$, $r = 1$ и, следовательно, связь полная прямая. Если $r = -1$, связь полная обратная, при $r = 0$ связь между признаками отсутствует.

Ранговый коэффициент корреляции Кендалла (τ) можно определить по формуле:

$$\tau = \frac{S}{n(n-1)} = \frac{2S}{n(n-1)},$$

где $S = P + Q$.

Этот коэффициент измеряется в интервале от -1 до 1 и интерпретируется так же, как и коэффициент Пирсона, но он даёт более строгую оценку связи,

чем коэффициент Спирмена $R \approx \frac{3}{2} \tau$. Это соотношение выполняется при большом числе наблюдений, $n > 30$, и слабых, либо умеренно тесных связях.

К непараметрическим методам исследования можно отнести коэффициент ассоциации K_{ac} и коэффициент контингенции $K_{кон}$, которые используются, если, например, необходимо исследовать тесноту зависимости между качественными признаками, каждый из которых представлен в виде альтернативных признаков.

Для определения этих коэффициентов создается расчётная таблица (таблица «четырёх полей»), где статистическое сказуемое схематически представлено в следующем виде:

Признаки	А (да)	А (нет)	Итого
В (да)	a	b	a + b
В (нет)	c	d	c + d
Итого	a + c	b + d	n

Здесь a, b, c, d - частоты взаимного сочетания (комбинации) двух альтернативных признаков $A - \bar{A}$ и $B - \bar{B}$; n - общая сумма частот.

Коэффициент ассоциации можно рассчитать по формуле:

$$K_{ac} = \frac{ad - bc}{ad + bc},$$

Коэффициент контингенции рассчитывается по формуле:

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

Нужно иметь в виду, что для одних и тех же данных коэффициент контингенции (изменяется от -1 до +1) всегда меньше коэффициента ассоциации.

Связь считается подтверждённой, если $K_{ac} > 0.5$ или $K_{кон} > 0.3$

Если необходимо оценить тесноту связи между альтернативными признаками, которые могут принимать любое число вариантов значений, применяется коэффициент взаимной сопряженности Пирсона ($K_{П}$) и Чупрова ($K_{ч}$).

Для исследования такого рода связи первичную статистическую информацию располагают в форме таблицы:

Признаки	А	В	С	Итого
Д	m_{11}	m_{12}	m_{13}	$\sum m_{1j}$
Е	m_{21}	m_{22}	m_{23}	$\sum m_{2j}$
Ф	m_{31}	m_{32}	m_{33}	$\sum m_{3j}$
Итого	$\sum m_{j1}$	$\sum m_{j2}$	$\sum m_{j3}$	П

Здесь m_{ij} - частоты взаимного сочетания двух атрибутивных признаков; Π - число пар наблюдений.

Коэффициент взаимной сопряженности Пирсона определяется по формуле:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{1 + \varphi^2}},$$

Коэффициент Чупрова определяется по формуле:

$$K_{\text{ч}} = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}},$$

где φ^2 - показатель средней квадратической сопряженности, определяемый путём вычитания единицы из суммы отношений квадратов частот каждой клетки корреляционной таблицы к произведению частот соответствующего столбца и строки:

$$\varphi^2 = \left(\frac{m_{11}^2}{\sum m_{1j} \cdot \sum m_{j1}} + \frac{m_{21}^2}{\sum m_{2j} \cdot \sum m_{j2}} + \frac{m_{33}^2}{\sum m_{3j} \cdot \sum m_{j3}} \right) - 1$$

K_1, K_2 - число групп по каждому из признаков.

Величина коэффициента взаимной сопряженности, отражающая тесноту связи между качественными признаками, колеблется в пределах от 0 до 1.

Коэффициент Фехнера характеризует элементарную степень тесноты связи, который целесообразно использовать для установления факта наличия связи, когда существует небольшой объём исходной информации. Данный коэффициент определяется по формуле:

$$K_{\Phi} = \frac{n_a - n_b}{n_a + n_b},$$

где n_a - количество совпадений знаков отклонений индивидуальных величин от их средней арифметической;

n_b - соответственно количество несовпадений.

Коэффициент Фехнера может изменяться в пределах -1,0 до +1,0. при значении равном 1 он указывает на полную прямую связь, при значении -1 на полную обратную связь, при нулевом значении – на отсутствие связи.

Например, имеются данные о выпуске продукции (x) на 6 однотипных предприятиях и потреблении на них электроэнергии (y):

Выпуск продукции	5	7	10	12	15	17
Потребление электричества	17	22	26	24	30	42

Рассчитаем средние значения для x и y

$$\bar{x} = \frac{5 + 7 + 10 + 12 + 15 + 17}{6} = 11,$$

$$\bar{y} = \frac{17 + 22 + 26 + 24 + 30 + 42}{6} = 26,83.$$

Выпуск продукции	5	7	10	12	15	17
Потребление электричества	17	22	26	24	30	42
$x - \bar{x}$	- 6	- 4	- 1	1	4	6
$y - \bar{y}$	- 9,83	- 4,83	- 0,83	- 2,83	3,17	15,17

$$K_{\phi} = \frac{5 - 1}{6} = 0,67.$$

Судя по полученному значению коэффициента, связь можно считать достаточно сильной.

Недостаток показателя Фехнера состоит в том, что разные по абсолютной величине отклонения имеют одинаковый вес. Более совершенным показателем степени тесноты связи является линейный коэффициент корреляции.

Учебное издание

Л.А. Чиркова

Статистика

Учебное пособие

Рекомендовано Учебно — методическим советом филиала в качестве учебно — методического пособия по специальности 38.02.01 Экономика и бухгалтерский учет (по отраслям)

Редактор Лебедева Е.М.

Подписано к печати 16.09.2015 г. Формат 60x84 ¹/₁₆.

Бумага офсетная. Усл. п. л. 5,23. Тираж 100 экз. Изд. № 3632.

Издательство Брянского государственного аграрного университета
243365 Брянская обл., Выгоничский район, с. Кокино, Брянский ГАУ