

Специальная математика и основы статистики

Статистика как наука. Статистическое наблюдение. Статистическая сводка и группировка

Вопрос 1. Предмет и метод статистики

Изучаемые статистикой явления и процессы многообразны. В первую очередь статистика изучает все, что связано с экономической деятельностью общества – производство и реализация промышленной и сельскохозяйственной продукции, строительство и реконструкция основных фондов, работа транспорта и связи, формирование и движение финансовых потоков. Статистические методы широко используются в анализе социальных процессов и явлений – занятости и безработицы, доходов населения, изучении общественного мнения и т.д.

В настоящее время под термином «статистика» чаще всего понимается следующее.

Статистика – это искусство и наука сбора и анализа данных.

Поскольку данными называется любой вид зарегистрированной информации, статистика играет важную роль во всех сферах деятельности человека.

Статистика – это одна из форм практической деятельности человека, цель которой – сбор, обработка и анализ данных о различных социальных, общественных и экономических явлениях.

Исторически данную функцию статистики реализует государство. Во всех странах мира деятельностью по сбору и обобщению информации для внутренних и внешних потребностей экономики занимаются специально учреждаемые государствами органы – статистические комитеты и комиссии, статистическое бюро. В нашем государстве эта работа выполняется Федеральной службой государственной статистики. В задачу таких государственных и ведомственных органов входят сбор, обобщение и публикация статистических данных в целом по стране, например, в России издается Статистический ежегодник «Россия в цифрах», а также по отдельным регионам и отраслям экономики.

Статистикой называют также числовые (цифровые) или количественные данные, характеризующие различные объекты и явления.

Но, прежде всего, **статистика** – это наука, которая изучает количественную сторону массовых социально-экономических явлений и процессов в неразрывной связи с их качественной стороной, выявляет основные характеристики и закономерности развития этих явлений и процессов в конкретных условиях места и времени.

Предметом статистики является количественная сторона массовых качественно определенных социально-экономических явлений и процессов, отображаемых посредством статистически показателей.

Особенности статистики как науки;

1) В отличие от других общественных наук статистика **изучает количественную сторону общественных явлений**. То есть статистика выявляет основные характеристики различных явлений общества и экономики, характеризует их, сравнивает между собой, анализирует происходящее с этими явлениями изменения, используя **числа** или **цифры**.

2) Статистика **изучает массовые явления и выявляет основные закономерности их развития**. Очевидно, что для установления закономерности развития процесса в целом необходимо проанализировать множество единичных случаев. Например, выявить, какая из двух групп колледжа лучше успевает в обучении невозможно, если взять по одному студенту из этих групп и сравнить полученные ими оценки. Для ответа на этот вопрос следует обобщить экзаменационные оценки всех студентов в каждой группе и сравнить полученные результаты. Поэтому статистика изучает не единичные факты, а массовые социально-экономические явления.

3) Для статистики **исключительно важна качественная определенность используемых числовых данных**.

Например, для математики действие $20 + 15 = 35$ уже является результатом. А статистик, прежде чем суммировать числа «20» и «15», обязательно бы выяснил, что они характеризуют и в каких единицах измерения они выражены. Ведь если числа 20 и 15 имеют разную природу, единицы измерения их просто нельзя суммировать, хотя арифметически, без учёта их качественного содержания, это возможно.

Отрасли статистики

Статистика изучает различные аспекты жизни общества. В зависимости от специфики объекта исследования различают отдельные отрасли статистики.

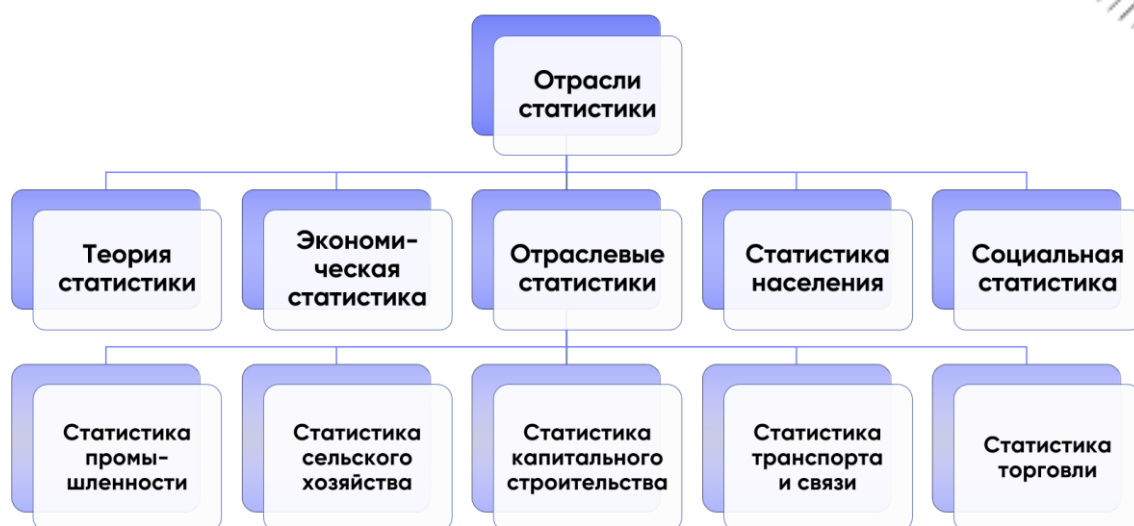


Рис. 1. Отрасли статистики

Теория статистики рассматривает общие понятия и методы сбора, обработки и анализа данных, разрабатывает общие показатели и методы изучения структуры, взаимосвязи и динамики социально-экономических процессов и явлений. Разрабатываемые ей методы и показатели используются другими отраслевыми статистиками.

Экономическая статистика изучает количественные закономерности происходящих в экономике явлений и процессов, выявляет основные тенденции экономического развития и их воздействие на уровень жизни населения на макроуровне, т.е. на уровне крупного региона или страны в целом. К основным показателям экономической статистики относятся валовой внутренний продукт, валовой региональный продукт, такие элементы национального богатства как основные фонды, материальные и оборотные средства, имущество домашних хозяйств.

Отраслевые статистики изучают основные показатели, закономерности и тенденции развития отраслей экономики.

Статистика населения изучает численный и национальный состав, а также возрастно-половую структуру населения, его размещение и воспроизводство как по стране в целом, так и в разрезе территориальных единиц. Одной из основных задач статистики населения является построение краткосрочных и долгосрочных демографических прогнозов.

Социальная статистика изучает социальную структуру населения, его уровень жизни и, в частности, доходы, уровень образования и культуры, состояние здоровья и медицинского обслуживания, использование свободного времени, общественное мнение, уровень преступности и другие социальные аспекты жизнедеятельности общества.

Метод статистики

Для того чтобы получить общее представление о статистической методологии, необходимо рассмотреть сам *процесс статистического исследования*, который включает четыре основных этапа.

1 этап. Сбор первичного статистического материала, проверка его полноты и достоверности. С этой целью применяются специальные методы *статистического наблюдения*. От качества полученных исходных данных во многом зависят окончательные результаты всего исследования.

2 этап. Предварительная обработка данных. Собранную информацию необходимо упорядочить, систематизировать, т.е. подготовить для дальнейшей работы. Основным методом, используемый на данном этапе – *метод группировок*. В результате его реализации от больших массивов данных исследователь переходит к компактным и удобным для анализа статистическим таблицам.

3 этап. Расчёт и интерпретация обобщающих статистических показателей. На данном этапе рассчитываются показатели *среднего уровня и вариации, структуры, взаимосвязи и динамики* изучаемых процессов и явлений. Полученные результаты анализируются и формулируются выводы.

4 этап. *Моделирование взаимосвязей* между социально-экономическими процессами и явлениями.

Используемые в процессе реализации всех этапов статистические приемы и методы в целом составляют методологию статистики.

Вопрос 2. Основные категории статистики

Как любая наука, статистика имеет свои специальные термины-категории, которые находятся в постоянном обиходе и используются при объяснении других понятий и методов. К важнейшим категориям статистики относятся: признак; статистическая совокупность, единица статистической совокупности, вариация, статистический показатель.

Признак - это объективная характеристика какого-либо объекта или явления, характерная черта или свойство, которое может быть определено или измерено. Именно значения различных признаков наблюдаются и регистрируются на первом этапе исследования – наблюдении.

Например, признаками, характеризующими промышленное предприятие, являются выручка от реализации продукции, прибыль, стоимость основных фондов, численность персонала и др. Признаками человека являются возраст, пол, место жительства, профессия, доход.

Значение, которое может принимать признак, называется *вариантом*. Например, существуют всего четыре варианта значений признака «экзаменационная оценка»: «2», «3», «4», «5». Если же учитывать оценки, проставляемые в зачетную книжку, то таких вариантов

остается три, так как оценка «2» в зачетку не проставляется. У отдельно взятого учащегося в зачетке может быть указано множество оценок по разным предметам, но вариантов будет по-прежнему три, а возможно, два или один, если студент учится без троек и четверок. Признаки подразделяются на количественные и качественные (рис. 2).



Рис. 2. Классификация видов признаков

Количественным является признак, отдельные варианты которого имеют числовое выражение и отражают размеры, масштабы изучаемого объекта или явления. К количественным признакам, например, относятся доход домохозяйства, площадь жилого помещения, цена товара, стаж работы. Количественные признаки в статистике преобладают над другими видами признаков, они наиболее информативны, аналитичны, именно на их использование нацелена большая часть статистического инструментария.

Непрерывные количественные признаки могут принимать любые числовые значения; например, прибыль предприятия может быть и положительной, причем от одного рубля до триллионов рублей, и нулевой, и даже отрицательной, если предприятие терпит убытки.

Дискретные количественные признаки принимают значения, ограниченные определенным диапазоном, широта которого зависит от особенностей признака. Например, при изучении семей по числу детей типичным вариантом признака будут 0, 1, 2 и 3, хотя есть семьи, в которых растут и воспитываются и 4, и 5 и даже большее количество детей.

Качественные признаки, в отличие от количественных, не поддаются прямому числовому описанию. Они характеризуют различные свойства объектов и явлений, отражают их состояние или отдельные характеристики. Например, к качественным признакам относятся форма собственности предприятия, пол человека, уровень образования, квалификация специалиста.

Качественный признак называется **альтернативным**, если он имеет только два варианта значений. Например, пол человека может быть мужским или женским, население страны или региона обычно делится на городское и сельское. Альтернативный признак может иметь и

числовое выражение. Например, семья может иметь доход в размере «до 15 тыс. руб. в месяц» и «15 тыс.руб. в месяц и более».

В отличие от альтернативного **атрибутивный** признак имеет более двух вариантов, которые выражаются в виде понятий или наименований. Например, район проживания, вид продукции, специальность работника, цвет товара. Такие признаки имеют место в различных областях исследования, но чаще с ними работают маркетологи, социологи и психологи.

Порядковые признаки отличаются от атрибутивных тем, что они имеют несколько ранжированных, т.е. упорядоченных по возрастанию или убыванию, качественных вариантов. Например, уровень образования (начальное, общее среднее и т.д.), уровень квалификации, воинское звание. Отдельные варианты порядкового признака трудно соизмерить количественно. Например, понятно, что образованность человека с высшим образованием выше, чем со средним специальным, но при этом нельзя утверждать, что она выше на 20% или на 30%. Водительская категория «Е» выше, чем категория «В», но количественных пропорций между ними не существует. Порядковый признак может иметь числовое выражение. Например, тарифный разряд рабочего или служащего, рейтинговые и экзаменационные оценки. Однако варианты таких признаков также не имеют количественных пропорций: рабочий 6-го разряда не обязательно в два раза больше вырабатывает продукции и в два раза больше зарабатывает, чем рабочий 3-го разряда.

Статистика изучает социально-экономические явления комплексно, поэтому, приступая к исследованию, отдельные объекты или явления объединяются по их качественному содержанию в специальные группы.

Статистическая совокупность - это реально существующее множество объектов или явлений, объединенных хотя бы одним общим признаком и обладающих внутренней взаимосвязью и целостностью.

Статистика имеет дело с совокупностями промышленных, сельскохозяйственных, строительных и торговых предприятий, с совокупностью коммерческих банков, с совокупностью населения страны или отдельного ее региона. Так, например, всех жителей г. Москвы можно рассматривать как статистическую совокупность, так как один признак - город проживания - будет у всех одинаковым. Индивидуальный объект или явление, составляющее статистическую совокупность, называется **единицей статистической совокупности**.

При проведении исследования именно у единиц совокупности измеряются или фиксируются важные для исследования признаки. Для отрасли единицей совокупности будет являться отдельное предприятие, для банковской системы - отдельный банк. В некоторых случаях для одной и той же совокупности можно выделить разные группы единиц. Например, при изучении половозрастной структуры

населения единицей является отдельный человек, при изучении доходов, обеспеченности жильем и предметами длительного пользования (телевизоры, холодильники и т.п.) единицей будет являться семья или домохозяйство.

Общее число единиц, образующих статистическую совокупность, называется **объемом совокупности**.

Объем совокупности следует отличать от **объема признака**, т.е. суммарного значения признака по всем единицам изучаемой совокупности. Так, число предприятий в отрасли - это объем совокупности, а общий выпуск продукции всеми предприятиями отрасли - объем признака.

Одной из важнейших характеристик статистической совокупности является ее однородность. **Однородной** является совокупность, единицы которой близки между собой по значениям признаков, существенных для данного исследования. Многие методы и приемы статистического исследования применимы лишь к однородным совокупностям.

Внутри статистической совокупности отдельные объекты и явления могут сильно отличаться друг от друга по своим свойствам и характеристикам. Например, у единиц совокупности «Население г. Москвы» кроме общего города проживания есть множество признаков, по которым единицы совокупности - москвичи - отличаются друг от друга: возраст, уровень образования, размер дохода. На языке статистики этот пример звучит так: население Москвы *варьирует* по возрасту, уровню образования, доходу. Или: внутри совокупности «Население г. Москвы» *существует вариация признаков* «возраст», «уровень образования» и «доход».

Вариация - это изменяемость, колеблемость величины признака у различных единиц совокупности, т.е. принятие единицами совокупности разных вариантов значений признака.

Причинами такой изменчивости являются:

Во-первых, *индивидуальные особенности изучаемых объектов*. Например, коммерческие банки варьируют по величине активов в зависимости от размера уставного фонда и реализуемой финансовой политики, работники предприятия имеют разную производительность труда в зависимости от их квалификации и мотивации, но также на производительность труда влияют состояние здоровья, настроение работника и общая атмосфера в трудовом коллективе.

Во-вторых, *изменяемость происходит под воздействием определенных условий, в том числе и непредвиденных (форс-мажорных)*. Например, вариация городов по численности населения зависит от географического положения города, развитости городской инфраструктуры и экономического положения в регионе. Спрос на туристические поездки в Таиланд в начале 2005 года значительно снизился после разрушений, вызванных серией цунами, что отразилось

па выручке туристических операторов, организующих отдых в этой стране.

Именно *наличие вариации и изменчивости данных определяет необходимость статистики как науки* и делает неизбежным процесс постоянного отслеживания и анализа информации о том, как развивается общество и экономика, как работает та или иная отрасль, насколько эффективно функционирует конкретное предприятие.

Важную роль в статистике играет *закон больших чисел* - общий принцип, в силу которого количественные закономерности, присущие массовым явлениям, отчетливо проявляются лишь при достаточно большом числе наблюдений. Единичные явления в большей степени подвержены действию случайных и несущественных факторов, чем масса в целом. При большом числе наблюдений случайные отклонения от общей закономерности развития взаимно погашаются. В результате этого обобщающие показатели становятся типичными, отражающими действие только постоянных и существенных факторов.

Статистическое исследование всегда завершается расчетом и анализом различных по виду и форме выражения статистических показателей.

Статистический показатель представляет собой количественную характеристику социально-экономических явлений и процессов в конкретных условиях места и времени.

Статистический показатель непосредственно связан с внутренним содержанием изучаемого явления или процесса, его сущностью. Изучаемые статистикой процессы и явления достаточно сложны и не могут быть описаны одним отдельно взятым показателем. В таких случаях используется **система статистических показателей**, позволяющих учесть различные характеристики изучаемого объекта или явления. Так, например, изучая промышленное предприятие, нельзя ограничиться только величиной выпуска продукции. Для полной экономической характеристики необходимо учесть численность персонала, производительность его труда, стоимость основных производственных фондов, финансовые показатели - затраты, прибыль, инвестиции в развитие производства.

В отличие от признака статистический показатель получается расчетным путем. Это может быть простой подсчет единиц совокупности, суммирование значений признака части или всех единиц совокупности, сравнение двух или нескольких величин, а также более сложные расчеты.

Все показатели можно классифицировать по охвату единиц совокупности, временной и пространственной определенности.

Вопрос 3. Сущность и виды статистического наблюдения

Статистическое наблюдение - это массовый, планомерный, научно организованный сбор данных о социально-экономических явлениях и

процессах, заключающийся в регистрации отобранных признаков у каждой единицы изучаемой совокупности. Например, при переписи населения по каждому жителю страны регистрируются сведения о поле, возрасте, семейном положении, образовании и др.

Статистическое наблюдение, как правило, носит массовый характер - для получения наиболее точных данных и выявления закономерностей и взаимосвязей при проведении наблюдения необходимо получить данные от максимально возможного числа изучаемых единиц совокупности.

Любое исследование проводится по заранее разработанному плану. Структура элементов **плана статистического наблюдения** представлена ниже.

I. ПРОГРАММНО-МЕТОДОЛОГИЧЕСКИЕ ВОПРОСЫ НАБЛЮДЕНИЯ

1. Объект наблюдения — совокупность, о которой должны быть собраны нужные сведения.

2. Единица наблюдения - единица изучаемой совокупности (человек, предприятие, банк), являющаяся носителем признаков, подлежащих регистрации.

3. Программа наблюдения - перечень существенных признаков, подлежащих регистрации, и перечень вопросов для их регистрации.

4. Период наблюдения - время, в течение которого осуществляется регистрация признаков у единиц наблюдения по установленной программе.

5. Критический момент наблюдения - момент времени, по состоянию на который производится регистрация собираемых сведений. На практике критический момент назначается на начало периода наблюдения.

6. Формуляр наблюдения - специальный бланк для записи ответов на вопросы программы. Формуляр включает *титულную часть*, где указывается вид наблюдения; наименование организации, которая проводит наблюдение; кем и когда утвержден формуляр; *программную часть* - перечень вопросов программы, места для записи ответов на них и коды ответов; *адресную часть*, где указывается точный адрес каждой единицы или группы единиц наблюдения и др. сведения.

7. Инструкция по заполнению формуляра наблюдения.

II. ОРГАНИЗАЦИОННЫЕ ВОПРОСЫ НАБЛЮДЕНИЯ

1. Описание особенностей объекта, целей и задач наблюдения, организации (подразделения), осуществляющей(-его) наблюдение.

2. Место и даты проведения наблюдения.

3. Подготовительные работы к наблюдению: подбор и обучение кадров, составление списков единиц совокупности, планирование мероприятий рекламной компании проводимого наблюдения и т.д.

4. Порядок приема и сдачи материалов наблюдения и представления предварительных и окончательных итогов наблюдения.

5. Финансирование и материально-техническое обеспечение работ. Основная цель статистического наблюдения - это сбор статистической информации о социально-экономических явлениях и процессах для получения обобщающих характеристик. Различают три основные формы статистического наблюдения (рис. 3).

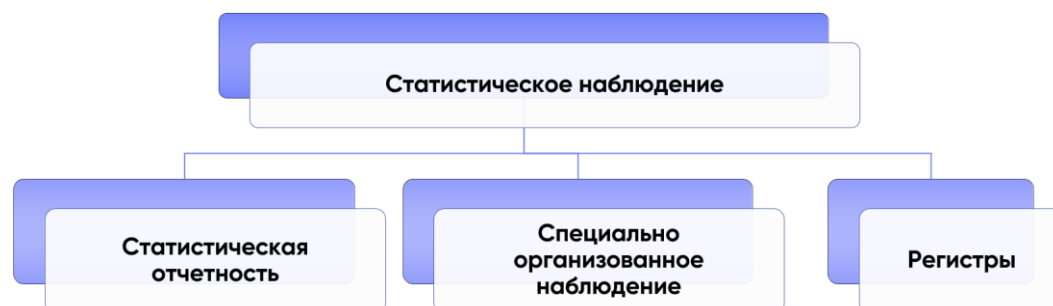


Рис. 3. Формы статистического наблюдения

Отчётность – это способ получения статистической информации от юридических лиц. Отчетность представляет собой специально разработанные формы, включающие в себя те признаки, которые подлежат регистрации. Формы статистической отчётности разрабатываются и утверждаются органами государственной статистики РФ. Любое юридическое лицо, являющееся субъектом экономики РФ, обязано предоставлять отчётность органам государственной статистики по месту своей регистрации по установленным отчётным формам и в установленные сроки.

Специально организованные статистические наблюдения проводятся для получения данных, не содержащихся в предоставляемой отчетности, или необходимых для проверки или уточнения данных, содержащихся в отчетах. Особым его видом является перепись.

Перепись - это специально проводимые широкомасштабные работы по сбору необходимой статистической информации об изучаемых объектах в границах отрасли, региона или страны в целом.

Перепись населения - это организация сбора, обработки и публикации демографических, экономических и социальных данных обо всем населении, проживающем в определенный момент времени в стране.

Принципы переписи населения: всеобщность охвата населения переписью; непосредственное получение сведений путем опроса конкретных людей; самоопределение людей при ответах на вопросы (без предъявления подтверждающих сведения документов); конфиденциальность сообщаемых населением сведений (результаты переписи публикуются только в сводном виде по стране, краю, области).

В 2006 году по состоянию на 1 июля проводилась Всероссийская сельскохозяйственная перепись - массовый сбор данных от

юридических и физических лиц, которые имеют земельные участки, предназначенные или используемые для производства сельскохозяйственной продукции, либо имеют сельскохозяйственных животных. Результаты этой переписи будут иметь важное значение для разработки эффективной агропромышленной политики и формирования объективной информации о состоянии продовольственного комплекса, который влияет на продовольственную и экономическую безопасность страны.

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

Кроме переписей, к специально организованному наблюдению относятся другие единовременные работы по сбору информации, в частности в рамках социологических или маркетинговых исследований.

Регистровое наблюдение представляет собой постоянный мониторинг состояния и развития наблюдаемых единиц, заключающийся в размещении и своевременной актуализации информации о них в базе данных. В статистической практике ряда стран применяют регистры населения, постоянно актуализируемые списки жителей страны с указанием их основных социально-демографических признаков, и регистры предприятий.

Регистры предприятий содержат данные о времени создания (регистрации) предприятия, его названии и адресе, об организационно-правовой форме, структуре, виде экономической деятельности, количестве занятых, основных экономических показателях из данных бухгалтерской отчетности и др. В России такой регистр называется *Статистический регистр хозяйствующих субъектов* - база данных об организациях, созданных на территории РФ, их местных единицах, индивидуальных предпринимателях и других типах хозяйствующих субъектов. Статистический регистр ведет Федеральная служба государственной статистики.

Классификация статистических наблюдений:

1) По *охвату единиц совокупности* наблюдение бывает двух видов: сплошное и несплошное.

При **сплошном наблюдении** обследованию подвергаются все единицы изучаемой совокупности. Примером сплошного наблюдения могут служить переписи населения.

При **несплошном наблюдении** обследованию подвергается только часть единиц изучаемой совокупности. Различают следующие виды несплошного наблюдения: выборочное, метод основного массива, монографическое обследование.

Выборочным называют наблюдение, основанное на принципе случайного отбора тех единиц изучаемой совокупности, которые должны быть подвергнуты наблюдению. Выборочное наблюдение, при правильной его организации и проведении, дает достаточно

достоверные данные для характеристики изучаемой совокупности в целом. Во многих случаях им вполне можно заменить сплошной учет. При этом обеспечивается значительная экономия средств, затрачиваемых на сбор и обработку данных.

Монографическое обследование представляет собой детальное, глубокое изучение и описание отдельных единиц совокупности. Такое исследование проводится с целью выявления имеющихся резервов, оценки результатов экономических экспериментов.

Метод основного массива заключается в том, что обследованию подвергаются наиболее крупные единицы, которые имеют преобладающий удельный вес в совокупности по основному для данного исследования признаку. Например, в ряде отраслей подавляющий объем выпуска продукции приходится на крупные и средние предприятия, поэтому результаты деятельности малых предприятий в этих отраслях практически не отражаются на обобщающих статистических показателях.

2) По *срокам регистрации* наблюдение может быть непрерывным (текущим) и прерывным.

Непрерывным называют наблюдение, которое ведется постоянно, и регистрация фактов производится по мере их свершения. Так, например, осуществляется регистрация рождений, заключенных браков, разводов и др. в органах ЗАГС.

Прерывное наблюдение повторяется через определенные, равные промежутки времени (например, ежегодное предоставление отчетности в органы государственной статистики) или по мере необходимости, без соблюдения строгой периодичности (как, например, перепись многолетних насаждений, проведенная один раз в прошлом веке).

3) По источнику сведений *различают непосредственное наблюдение, документальное наблюдение и опрос.*

Непосредственным называют такое наблюдение, при котором сами регистраторы путем непосредственного замера, взвешивания или подсчета устанавливают значение признака и производят запись в формуляре наблюдения. Например, инвентаризация основных средств на предприятиях.

Документальное наблюдение предполагает запись ответов на вопросы формуляра на основании соответствующих документов. Например, сбор данных об успеваемости студентов вуза на основе зачетно-экзаменационных ведомостей, заполнение форм статистической отчетности на основании данных бухгалтерского учета 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.

4) В статистике применяются следующие **способы сбора сведений**: *отчетный, экспедиционный, самоисчисление, анкетный.*

Отчетный способ заключается в обязательном представлении хозяйствующими субъектами статистических отчетов о своей деятельности в установленной форме и в установленные сроки.

Экспедиционный способ наблюдения заключается в том, что специально привлеченные и обученные работники посещают каждую единицу наблюдения и сами заполняют формуляр наблюдения. Этим способом собираются сведения при переписях населения.

При способе **самоисчисления** формуляры заполняют сами опрашиваемые. Обязанность специально привлеченных для получения информации сотрудников состоит в раздаче формуляров, инструктаже опрашиваемых, сборе и проверке правильности заполнения формуляров.

Анкетный способ — это сбор данных с помощью специальных вопросников, рассылаемых определенному кругу лиц или публикуемых в периодической печати. Как правило, этим способом получения информации пользуются при проведении социологических опросов. Также многие крупные производители бытовой техники, мебели и других предметов потребления вкладывают анкеты в упаковку товара с просьбой заполнить и вернуть их производителю по указанному адресу.

Точность статистического наблюдения

Под точностью в статистике понимают степень соответствия данных наблюдения реальным их значениям. Возникающие расхождения называются ошибками. Ошибки определяются как разность или как отношение между этими значениями. Как правило, ошибки возникают при регистрации сведений или при измерении.

Ошибки регистрации возникают вследствие неправильного установления фактов в процессе наблюдения, или ошибочной их записи, или того и другого вместе. *Случайные ошибки* возникают в результате опечаток, описок, оговорок. Например, при регистрации даты регистратор вместо 15 июня указал 15 июля. При большом числе наблюдений благодаря действию закона больших чисел эти ошибки более или менее взаимно погашаются.

Систематические ошибки наиболее опасны, поскольку приводят к сильному искажению данных. Наиболее показательными такими ошибками являются: занижение или округление населением своего возраста на цифры, оканчивающиеся на 5 или 0; сокрытие экономическими субъектами реальных размеров финансовых результатов их деятельности, стремление респондентов указать заниженное значение своего дохода и т.п.

При несплошном наблюдении возникают **ошибки репрезентативности** (или ошибки представительности). Они заключаются в том, что значения признаков по отобранной выборочной совокупности не отражают реально существующей картины в совокупности в целом.

С целью выявления ошибок проводится контроль полученных материалов. После проведения наблюдения весь собранный материал проверяют на полноту охвата единиц. Если выявлены упущенные единицы наблюдения, дальнейшие действия зависят от того, представляется возможным восполнение пробелов или нет.

Для проверки качества заполнения формуляров и других документов наблюдения используют логический и арифметический контроль.

Логический контроль состоит в сопоставлении между собой ответов на вопросы формуляра наблюдения и выяснении их логической совместимости. При обнаружении несовместимых ответов пытаются путем дальнейших сопоставлений установить, какой из ответов является верным.

Арифметический контроль состоит в проверке различных расчетов, результаты которых проведены в формуляре наблюдения, в частности, итогов, вычисления процентов, расчетов средних величин и т.п.

Вопрос 4. Статистическая сводка и группировка данных

Собранные данные необходимо систематизировать и подготовить к последующей обработке. Для упорядочивания исходных данных статистика применяет два основных метода - метод сводки и метод группировки.

Сводка - это научная обработка первичных данных с целью получения обобщенных характеристик изучаемого социально-экономического явления по ряду существенных для него признаков (рис. 4).

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

Данные можно обрабатывать централизованно - в одной организации или децентрализованно - последовательно передавая результаты сводки от территориальных органов в федеральные (как при обработке отчетности).



Рис. 4. Виды статистической сводки

Группировкой называется разбиение общей совокупности единиц объекта наблюдения по одному или нескольким существенным признакам на более однородные группы. Группы могут различаться между собой по числу объектов и в качественном отношении. В соответствии с познавательными задачами различают три основных вида группировок (рис. 5).

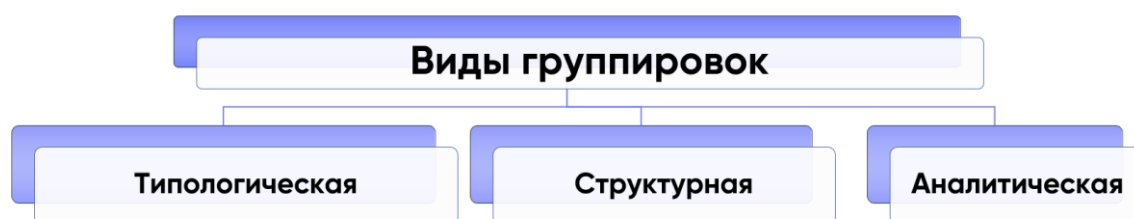


Рис. 5. Виды группировок

Типологическая группировка - это разбиение разнородной совокупности единиц наблюдения на отдельные качественно однородные группы и выявление на их основе устойчивых социально-экономических типов явлений. Например, группировка предприятий и организаций по формам собственности, группировка торговых предприятий по их принадлежности к сферам промышленных и продовольственных товаров.

Структурная группировка предназначена для изучения состава совокупности по какому-либо признаку. Например, изучение возрастной структуры безработных в Российской Федерации.

Аналитическая группировка выявляет взаимосвязи между изучаемыми признаками. В статистике выделяют факторные и результативные признаки. **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.

При построении аналитической группировки, как правило, единицы совокупности группируются по факторному признаку; каждая выделенная группа характеризуется средними величинами результативного признака.

По способу построения группировки бывают простые и комбинационные. **Простой** называется группировка, в которой группы образованы только по одному признаку.

Комбинационной называется группировка, в которой образование групп производится по двум и более признакам, взятым в сочетании (комбинации). Рекомендуются сначала группировать единицы по качественным признакам, а затем - по количественным. Например, в группировке водителей автопарка по уровню квалификации (классу) и производительности труда вначале все водители делятся на две группы

по классу, а затем внутри каждого класса производится деление по проценту выполнения плана

Принципы построения статистических группировок

Построение статистических группировок включает следующие этапы:

1. Определение группировочного признака.
2. Определение числа групп.
3. Расчет ширины интервала группировки.
4. Выбор признаков, которые будут характеризовать выделенные группы.

Группировочным (или *основанием* группировки) называется количественный или качественный признак, по которому проводится разбиение единиц совокупности на группы.

Число выделяемых групп зависит от вида группировочного показателя, степени его вариации и объема изучаемой совокупности.

Если группировочный признак качественный, то число групп будет равно числу вариантов этого признака. Например, группируя сотрудников по полу, мы получим 2 группы: сотрудники-мужчины и сотрудники-женщины.

При определении числа групп необходимо учесть **размах вариации** группировочного признака (**R**), который определяется по формуле:

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

Где X_{max} - максимальное значение группировочного признака;

X_{min} - минимальное значение группировочного признака.

Чем больше размах вариации группировочного признака, тем большее число групп может быть образовано. Однако при слишком большом их числе возникает проблема «пустых» групп, т.е. не содержащих ни одного объекта.

Число групп можно определить математически или экспертным путем. Математический способ предполагает использование **формулы Стерджесса**, которая дает хорошие результаты при большом объеме совокупности:

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

где m - число выделяемых групп;

n - общее число единиц совокупности.

Интервал - это значения признака, лежащие в определенных границах.

Нижней границей называется наименьшее, а **верхней** - наибольшее значение признака в интервале. **Ширина** интервала - это разность между его верхней и нижней границами. По ширине интервалы группировки бывают **равные** (одинаковые) и **неравные**

Ширина равного интервала h определяется по следующей формуле:

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

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

m - число выделяемых в совокупности групп.

Существуют следующие *правила округления ширины интервала h* :

1. Если h имеет один знак до запятой (например: 0,67; 1,487; 3,82), полученные значения округляют до десятых (0,7; 1,5; 3,8.).
2. Если h имеет две значащие цифры до запятой (например: 14,876), это значение округляют до целого числа (15).
3. В случае, когда h является трех-, четырех- или еще более значимым числом, его величину следует округлить до ближайшего числа, кратного 100 или 50. Например, 652 следует округлить до 650 или до 700.

Интервалы группировки бывают также **закрытыми и открытыми**

Если максимальные или минимальные значения сильно отличаются от других значений группировочного признака, то для определения ширины интервала используют значения, несколько превышающие минимум, и несколько меньше, чем максимум. Полученную по формуле величину округляют и используют в качестве ширины интервала, а первый и/или последний интервалы группировки открывают по верхней или нижней границе. Это делается для того, чтобы учесть в открытых интервалах единицы, имеющие аномально большие или малые значения группировочного признака.

Если значение признака у какой-то единицы совпадает с верхней границей интервала (например, в табл. 3.8 40 млн. руб. - это верхняя граница 1 интервала (20- 40) и нижняя граница 2 (40-60)), такая единица, как правило, относится к следующему интервалу (ко 2-му, 40-60).

Если в основании группировки лежит дискретный признак, то нижняя граница каждого интервала, начиная со второго, берется равной верхней границе предыдущего интервала, увеличенной на 1. Например, группируя страховые кампании числу занятого персонала, можно выделить следующие группы (чел.): 100-150, 151-200, 201-300, 301-400.

Ряд распределения представляет собой простейшую группировку, в которой каждая выделенная группа характеризуется только количеством входящих в нее единиц совокупности. Различают атрибутивные и вариационные ряды распределения.

Атрибутивными называют ряды распределения, построенные по качественным признакам. **Вариационными** называют ряды распределения, построенные по количественному признаку. Вариационный ряд состоит из двух элементов: вариантов и частот.

Вариантами называются отдельные значения признака. Варианты признака обозначаются как X_i . **Частотами** называются численности отдельных вариантов или каждой группы вариационного ряда. Частоты показывают, как часто встречаются те или иные значения признака в

изучаемой совокупности и обозначаются через F_i . Сумма всех частот определяет объем совокупности. **Частотами** называются частоты, выраженные в долях единицы или в процентах к итогу. Соответственно сумма частостей равна 1 или 100%.

Дискретный вариационный ряд характеризуют распределение единиц совокупности по дискретному признаку, принимающему только целые значения. Например, группы студентов по баллу в сессию.

Интервальный вариационный ряд распределения - это ряд распределения, в котором группировочный признак задан интервалами значений. Например, распределение консалтинговых фирм по величине прибыли.

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

Полигон используется для изображения дискретных вариационных рядов. По оси абсцисс (X) в одинаковом масштабе откладываются значения признака, а по оси ординат (Y) - частоты. Полученные на пересечении осей X и Y точки соединяются прямыми линиями, в результате чего получают ломаную линию, называемую полигоном частот (рис. 6).

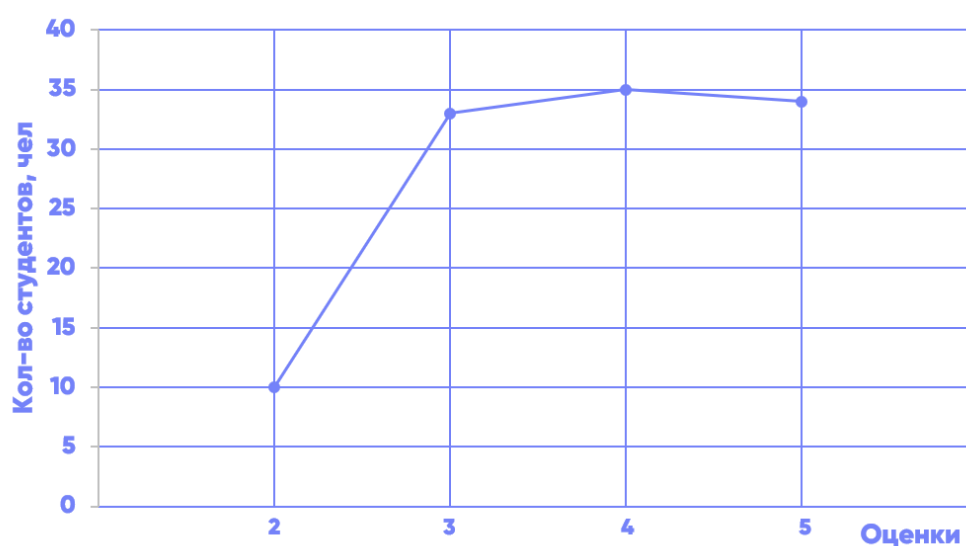


Рис. 6. Полигон распределения студентов по баллам в сессию

Гистограмма применяется для изображения интервального вариационного ряда. При ее построении на оси X откладываются величины интервалов, а на оси Y - частоты, которые изображаются прямоугольниками, построенными на соответствующих интервалах. Гистограмма может быть преобразована в полигон распределения, если середины верхних сторон прямоугольников соединить прямыми линиями (рис. 7).

При построении гистограммы распределения вариационного ряда с неравными интервалами по оси Y наносят не частоты, а плотность распределения признака в соответствующих интервалах. **Плотность распределения** - это частота, рассчитанная на единицу ширины интервала.

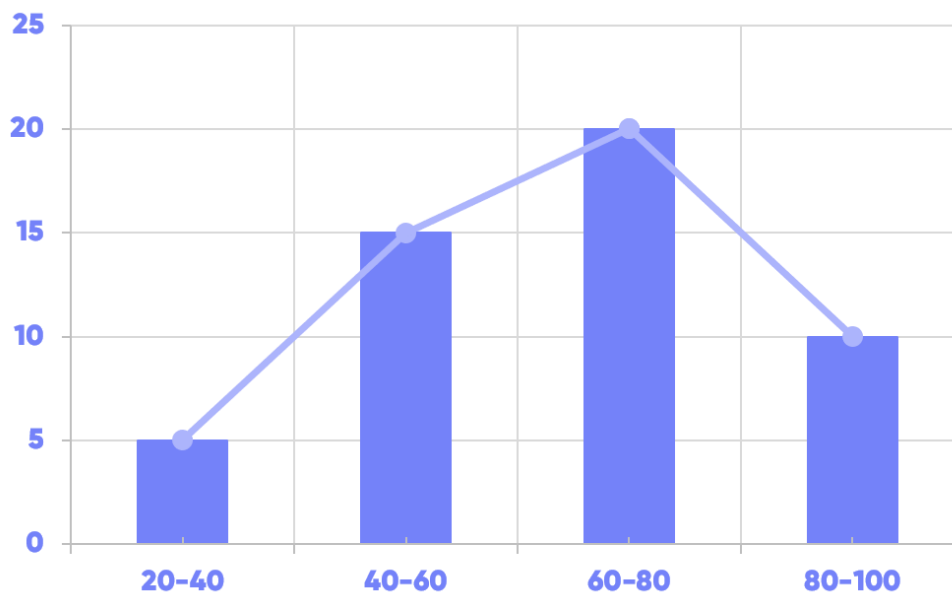


Рис. 7. Гистограмма и полигон распределения консалтинговых фирм по величине прибыли за 2019 г.