Тексты лекции по статистике (1 часть)
для студентов экономических специальностей
заочного факультета
8 лекций (6 тем)

Делидович Л.П., старший преподаватель

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

Гомель 2014

ТЕМА 1. ПРЕДМЕТ И ОСНОВНЫЕ КАТЕГОРИИ СТАТИСТИКИ

1) Краткая история развития статистики.
2) Предмет статистики и основные разделы.
3) Основные категории статистики. Этапы статистического исследования.

1) КРАТКАЯ ИСТОРИЯ РАЗВИТИЯ СТАТИСТИКИ.

Родоначальником статистики как науки является Адольф Кетле (22.2.1796, Гент, - 17.2.1874, Брюссель), бельгийский математик, астроном, метеоролог, социолог. Профессор математики и астрономии в Брюсселе (1819). С 1820 член, с 1834 секретарь Бельгийской АН. С 1832 директор основанной им астрономической и метеорологической обсерватории в Брюсселе. В 1841-74 председатель Центральной бельгийской статистической комиссии, учрежденной по его инициативе; организатор Первого международного статистического конгресса (Брюссель, 1853); председатель первого международного метеорологического совещания - Конференции по морской метеорологии (1855).

С позиций позитивизма Кетле утверждал, что социальная жизнь и физические явления подчиняются законам одного порядка и должны изучаться точными методами математической статистики. Выдвинул понятие "среднего человека" (homme moyen), обладающего средними физическими, моральными и интеллектуальными характеристиками; отдельные индивиды, по Кетле, лишь искаженные выражения среднего типа. Он доказал, что некоторые общественные явления (рождаемость, смертность, преступность и др.) подчиняются определённым закономерностям.

Кетле оказал значительное влияние на развитие количественных методов в социальных исследованиях.

Основными особенностями статистики XX века является использование теории вероятностей и ЭВМ.

В настоящее время в Республике Беларусь работает Министерство Статистики и Анализа, которое является руководящим органом в области статистики.

2) ПРЕДМЕТ СТАТИСТИКИ И ОСНОВНЫЕ РАЗДЕЛЫ.

Предметом статистики является количественная сторона массовых социально-экономических явлений в неразрывной связи с их качественной стороной, конкретных условий, места и времени.

В настоящее время в статистике можно выделить несколько разделов:

1) теория статистики (методы);
2) национальное счетоводство;

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

3) народно-хозяйственный учет (принцип двойной записи) – для определения основных макропоказателей;

4) международная статистика (сравнительная);

5) анализ динамики явлений и построение на его основе различных статистических индикаторов;

6) статистическое моделирование.

3) ОСНОВНЫЕ КАТЕГОРИИ СТАТИСТИКИ. ЭТАПЫ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ.

Основные категории статистики:

1) **Статистическая совокупность** – это множество элементов, объединенных единой качественной основой. Исчезновение одного или нескольких элементов не уничтожает совокупность как таковую.

Элемент этого множества называется **единицей совокупности**.

2) **Признак** – это свойство, характеристика единицы совокупности.

3) **Вариация** – изменчивость признаков у отдельных единиц совокупности.

Значение, которое принимает варьирующий признак, называют **вариантой**.

4) **Статистический показатель** – обобщающая количественная характеристика статистической совокупности или ее части. Статистический показатель может быть рассчитан только на основании значения признака.

5) **Статистическая закономерность**. Одной из важнейших задач статистики является установление статистических закономерностей характеризующих совокупность.

Особенностью статистической закономерности является то, что она не верна для каждой единицы, а проявляется только как тенденция для всей совокупности в целом. Для того чтобы выявить статистическую закономерность проводят статистическое исследование, состоящее из 3-х основных этапов:

1) статистическое наблюдение – сбор первичного материала;

2) статистическая обработка полученных материалов – различные группировки и расчет различных статистических показателей;

3) статистический анализ – позволяет выявить статистические закономерности.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

ТЕМА 2. СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

1) **Сущность статистического наблюдения**
2) **Формы, виды и способы статистического наблюдения**
3) **План статистического наблюдения**
4) **Контроль материалов статистического наблюдения. Ошибки статистического наблюдения**

1) СУЩНОСТЬ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Статистическое наблюдение - первая стадия статистического исследования, представляющего собой научно-организованный по единой программе учет фактов, характеризующих явления и процессы общественной жизни.

Наблюдение является фундаментом статистического исследования. В процессе наблюдения формируются данные, которые на последующих этапах исследования подвергаются обработке и анализу. Организация статистического наблюдения предполагает определение объекта и единицы наблюдения, разработки программы и организационного плана проведения наблюдения.

Объект статистического наблюдения - это совокупность общественных явлений и процессов, которые полежат данному статистическому наблюдению.

Изучение объекта статистического наблюдения предполагает выделение в его составе отдельных единиц.

Единицей статистического наблюдения является составной элемент объекта наблюдения, являющийся носителем признаков, подлежащих регистрации в процессе данного наблюдения. От единицы наблюдения следует отличать отчетную единицу.

Отчетная единица - это источник сведений о единицах наблюдения.

Программа статистического наблюдения представляет собой перечень вопросов, по которым нужно получить сведения в отношении каждой исследуемой единицы в процессе наблюдения.

2) ФОРМЫ, ВИДЫ И СПОСОБЫ НАБЛЮДЕНИЯ

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

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

Особенность отчетности состоит в том, что она обязательна, документально обоснована и юридически подтверждена подписью руководителя.

Примером второй формы наблюдения – специального статистического обследования – является **проведение переписей населения**.

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

В зависимости от задач статистического исследования и характера изучаемого явления учет фактов можно производить:

- систематически, постоянно охватывая факты по мере их возникновения – это будет текущее наблюдение (отчетность);

- регулярно, но не постоянно, а через определенные промежутки времени – это будет периодическое наблюдение (переписи населения).

Рис. 1. Формы, виды и способы статистического наблюдения

С точки зрения полноты охвата фактов статистическое наблюдение может быть сплошным и несплошным. **Сплошное наблюдение** представляет собой полный учет всех единиц изучаемой совокупности. **Несплошное наблюдение** организуют как учет части единиц совокупности, на основе которой можно получить обобщающую характеристику всей совокупности. К видам несплошного наблюдения относятся: способ основного массива, выборочные наблюдения, монографические описания.

При непосредственном учете фактов сведения получают путем личного учета единиц совокупности: пересчета, взвешивания, измерения и т.д.

Документальный способ сбора статистической информации базируется на систематических записях в первичных документах, подтверждающих тот или иной факт.

В ряде случаев для заполнения статистических формуляров прибегают к **опросу населения**, который может быть произведен **экспедиционным, анкетным или корреспондентским способом**.

Существуют различные способы формирования выборочной совокупности. Это, во-первых, индивидуальный отбор, включающий такие

РЕПОЗИТОРИЙ ГГУ

разновидности, как собственно случайный, механический, стратифицированный, и, во-вторых, серийный, или гнездовой, отбор.

3) ПЛАН СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Организационный план статистического наблюдения представляет собой перечень мероприятий, необходимых для успешного выполнения работы по сбору и обработке материалов, с указанием сроков и исполнителей.

Это основной документ, в котором отображаются важнейшие вопросы организации и проведения намеченных мероприятий. В организационном плане указываются: органы наблюдения, время наблюдения, сроки наблюдения, а также подготовительные работы к наблюдению, в том числе порядок комплектования и обучения кадров, необходимых для проведения наблюдения, порядок его проведения, приема и сдачи материалов, получения и представления предварительных и окончательных итогов и др. При организации статистического наблюдения обязательно должен быть решен вопрос о времени проведения наблюдения, включая выбор сезона наблюдения, установление срока (периода) и критического момента наблюдения.

Сезон (время года) для наблюдения следует выбрать такой, в котором изучаемый объект пребывает в обычном для него состоянии. Например, перепись населения в нашей стране чаще всего проводится зимой, так как наблюдается наименьшее передвижение населения.

Под *периодом (сроком)* проведения наблюдения понимается время начала и окончания сбора сведений.

Время наблюдения - это время, к которому относятся данные собранной информации. Для предупреждения неполного учета или повторного счета для всех единиц статистической совокупности устанавливается единое время регистрации изучаемых показателей.

Критической называют дату, по состоянию на которую сообщаются сведения. При переписях обычно устанавливаются время начала (дата, а иногда и час) и время окончания регистрации наблюдения фактов. Например, Всесоюзная перепись населения 1989 г. проводилась в течение 8 дней, с 12 по 19 января.

Критическим моментом наблюдения выбирают полночь, момент окончания одних суток и начала других. Так, критическим моментом Всесоюзной переписи населения в 1989 г. было 12 ч ночи с 12 января на 13 января. Все сведения о каждом жителе страны фиксировались такими, какими они были по состоянию на данный момент. Умершие после 12 ч ночи вносились в переписные листы, а родившиеся после 12 ч ночи учету не подлежали и в переписные листы не записывались.

4) КОНТРОЛЬ МАТЕРИАЛОВ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ. ОШИБКИ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Материалы, собранные в результате статистического наблюдения, подвергаются всесторонней проверке и контролю. Как показывает практика,

даже при четко организованном статистическом наблюдении встречаются погрешности, ошибки, требующие исправления. На современном этапе контроль данных в условиях сплошной информатизации приобретает особое значение. Предусматриваются различные методы проверки получаемых данных. Это, прежде всего, счетный (арифметический) и логический контроль.

Счетный (арифметический) контроль сводится к проверке общих и групповых цифровых итогов и их сопоставлению; задача его - обнаружить и исправить неверные итоги числовых показателей.

Логический контроль проводится путем сопоставления полученных данных с другими известными признаками, показателями. Возможно сопоставление за прошлый период по одной и той же единице, или за один и тот же период с данными по другой единице наблюдения. В результате контроля выявляются неправдоподобные случаи, т.е. логический контроль выявляет возможность ошибки.

Ошибки наблюдения по источнику происхождения можно подразделить на следующие:

- преднамеренные (злостные),
- непреднамеренные.

Непреднамеренные ошибки в свою очередь подразделяются наследующие:

- случайные,
- систематические,
- репрезентативности (представительности).

В соответствии с этими особенностями методы контроля данных могут быть разными.

Преднамеренные ошибки завышают или занижают конкретные значения признака, показателя. Они могут грубо искажать действительное положение. Поэтому преднамеренные ошибки требуют сплошного контроля.

Случайные ошибки чаще связаны с невнимательностью регистратора, небрежностью в заполнении документации, неточностью измерительных приборов и т.п. При достаточно большом числе зарегистрированных явлений они могут нейтрализовать друг друга. И все же подчеркнем, что одной из задач наблюдения является исключение таких ошибок. В целях контроля программа может предусматривать использование опознавательных признаков. Они образуют адресную часть и содержат, например, фамилию, имя и отчество, адрес и т.п. В случае обнаружения ошибки эти признаки позволяют обратиться к источнику информации и исправить ошибку.

Возможно включение в программу контрольных вопросов. Они позволяют, например, уточнить возраст по году рождения или выполнить логический контроль данных.

Непреднамеренные систематические ошибки возникают в самых различных ситуациях, например, при округлении признака в большую или меньшую сторону. Или при изучении бюджетов семей выясняется, что

некоторые виды расходов забываются. Такие ошибки требуют корректировки в соответствии с особенностями явлений и процессов. Ошибки представительности, или репрезентативности, свойственны только выборочному наблюдению. Они показывают, в какой степени выборочная совокупность представляет (репрезентует) генеральную совокупность. Эти ошибки возникают вследствие того, что наблюдению подвергается лишь часть единиц изучаемой совокупности, и сведения, полученные в результате этого, не могут абсолютно точно отобразить свойства всей массы явлений совокупности.

Достоверность статистических данных обеспечивается различными способами. К ним относятся: рациональная разработка программы наблюдения и инструкций, разъясняющих содержание показателей: систематический контроль и проверка постановки учета и отчетности на всех предприятиях, в учреждениях и организациях; подбор и подготовка кадров и т.д.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СК

ТЕМА 3. ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ

1) **Сущность группировки и ее основные виды**
2) **Этапы построения группировок**
3) **Ряды распределения**
4) **Пример построения аналитической группировки**

1) СУЩНОСТЬ ГРУППИРОВКИ И ЕЕ ОСНОВНЫЕ ВИДЫ

Обработка статистических данных начинается с построения группировок.

Группировка – распределение единиц совокупности на группы однородные в определенном отношении и характеристика выделенных групп с помощью статистических показателей.

Результаты группировки представляются в виде таблиц.

С помощью группировок решается несколько задач, основными из которых:

1) выделение типов социально-экономических явлений – с помощью типологической группировки (например: группировка предприятий по формам собственности, группировка населения по социальным группам и т.д.);

2) изучение структуры статистической совокупности - с помощью структурных группировок и рядов распределения (например: группировка по полу);

3) изучение взаимосвязи между признаками - с помощью аналитической группировки.

Группировка, в которой группы образованы по одному признаку, называется **простой**. Для характеристики явления бывает недостаточно разбить совокупность на группы по какому-либо однородному признаку. В этом случае строят сложные группировки.

Сложной называется группировка, в которой разделение совокупности на группы производится по двум и более признакам, взятым в сочетании (комбинации).

2) ЭТАПЫ ПОСТРОЕНИЯ ГРУППИРОВОК

1) Выбор группировочного признака (признаков) – выбор производится исходя из целей исследования.

2) Определение числа групп и величины интервала группировки.

Если группировка проводится по атрибутивному признаку, то обычно выделяют столько групп, сколько значений этого признака.

Если значений много, то пользуются классификацией.

Если группировка проводится по дискретному признаку и значений этого признака немного, то выделяют столько групп, сколько значений этого признака.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Если группировка проводится по непрерывному признаку или дискретному со многими значениями, то число групп можно определить по следующей формуле:

n = 1+3,322 lgN,

где N-число единиц совокупности.

Если признак непрерывный или дискретный со многими значениями, необходимо найти интервал группировки.

Интервалы бывают:

- **равные**, когда разность между максимальным и минимальным значениями в каждом из интервалов одинакова;
- **неравные**, когда, например, ширина интервала постепенно увеличивается, а верхний интервал часто не закрывается вовсе;
- **открытые**, когда имеется только либо верхняя, либо нижняя граница;
- **закрытые**, когда имеются и нижняя, и верхняя границы.

В случае равных интервалов величина интервала может быть определена как:

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

или

$$i = \frac{X_{max} - X_{min}}{1 + 3{,}322 \cdot \lg N}.$$

где i - величина интервала;

n – число групп;

Xmax, Xmin – максимальное и минимальное значения признака.

3) Подсчет численности единиц совокупности в каждой группе и по совокупности в целом.

4) Характеристика выделенных групп с помощью различных статистических показателей.

5) Оформление результатов группировки в таблице.

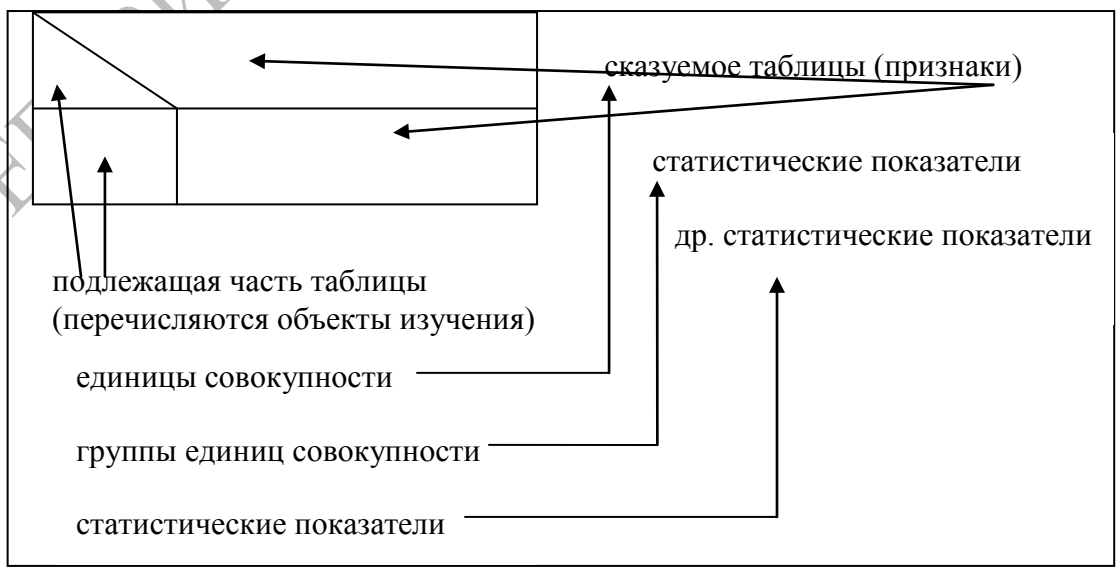

Рис. 2. Схема таблицы

РЕПОЗИТОРИЙ ГГУ ИМЕНИ Ф. СКОРИНЫ

Почти всегда признаки и статистические показатели имеют **единицы измерения**.

После таблицы **обязательно** должны быть сделаны выводы и разъяснения.

3)РЯДЫ РАСПРЕДЕЛЕНИЯ

Ряд распределения – это группировка, с помощью которой устанавливается распределение единиц совокупности по величине или значению некоторого признака.

Ряды, построенные по атрибутивному признаку, называются **атрибутивными рядами распределения**.

Ряды распределения, построенные по количественному признаку, называются **вариационными рядами**.

Примером атрибутивных рядов могут служить распределения населения по полу, занятости, национальности, профессии и т.д.

Примером вариационного ряда распределения может служить распределение населения по возрасту, рабочих – по стажу работы, заработной плате и т.д.

Вариационные ряды в зависимости от характера вариации подразделяются на **дискретные и интервальные**.

Вариационные ряды распределения состоят их двух элементов **вариантов и частот**.

Вариантами называются числовые значения количественного признака в ряду распределения, они могут быть положительными и отрицательными, абсолютными и относительными.

Частоты – это численности отдельных вариантов или каждой группы вариационного ряда. Сумма всех частот называется объемом совокупности и определяет число элементов всей совокупности.

Пример дискретного ряда распределения:

Таблица 1. Распределение рабочих цеха по тарифным разрядам

1)Тарифный разряд	2)Число рабочих	3) В % к итогу	Накопленная частота
2	40	20,0	40
3	64	32,0	104
4	60	30,0	164
5	22	11,0	186
6	14	7,0	200
итого	200	100,0	---

1) Тарифный разряд – признак, по которому распределяют. 2,3,4,5,6 – варианты.
2) Число рабочих: 40, 64, 60, 22, 14 – частоты.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

3) В % к итогу: 20,0; 32,0; 30,0; 11,0; 7,0 – частости.

Выводы: в этом цехе больше половины рабочих имеют 3 и 4 разряды, т.е. средний уровень квалификации.

Графически дискретный ряд распределения может быть представлен полигоном распределения.

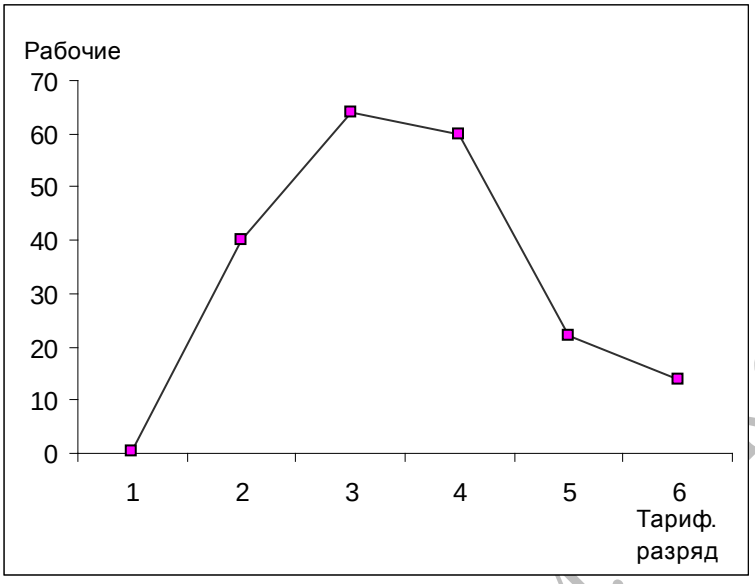

Рис. 3. Распределение рабочих цеха по тарифным разрядам

Пример. По следующим данным построить интервальный ряд распределения, выделив 4 группы с равными интервалами по стоимости основных средств:

Таблица 2. Стоимость основных средств и объем продукции по предприятиям

№ предприятия	Основные средства, млрд. руб	Объем продукции, млрд. руб.
1	1,8	2,2
2	1,2	1,0
3	2,8	2,8
4	2,4	2,9
5	3,1	3,6
6	5,2	6,4
7	3,9	4,3
8	3,5	4,6
9	3,5	4,5
10	4,3	5,6
11	4,5	5,4

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

12	4,0	4,8

$$i = \frac{5.2 - 1.2}{4} = 1\ (млрд.руб.)$$

Таблица 3. Распределение предприятий по стоимости основных средств

Стоимость основных средств, млрд. руб.	Число предприятий	В % к итогу	Накопленная частота
1,2-2,2 2,2-3,2 3,2-4,2 4,2-5,2	2 3 4 3	16,7 25,0 33,3 25,0	2 5 9 13
Итого	12	100,0	--

Здесь значения признаков (варианты) представлены интервалами.

В % к итогу – частости.

Вывод: Чаще всего предприятия имеют основные средства стоимостью 3,2-4,2 млрд. руб.

Графически интервальный ряд распределения может быть представлен гистограммой распределения:

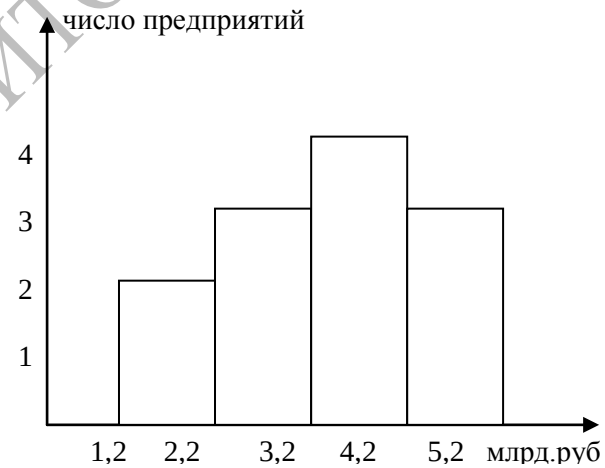

Рис.4. Гистограмма распределения предприятий по стоимости основных средств

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

4) ПРИМЕР ПОСТРОЕНИЯ АНАЛИТИЧЕСКОЙ ГРУППИРОВКИ

Аналитическая группировка проводится, чтобы определить **наличие** и **направление** связей между факторным и результативным признаком.

Аналитическую группировку всегда проводят по факторному признаку. Затем в каждой выделенной группе рассчитывают среднее значение результативного признака.

Если с увеличением факторного признака средние значения результативного признака возрастают, то между факторным и результативным признаком существует **прямая связь**.

Если с увеличением факторного признака средние значения результативного признака уменьшаются, то между факторным и результативным признаком существует **обратная связь**.

Если с увеличением факторного признака средние значения результативного признака меняются беспорядочно, то связи между этими признаками нет.

Построим аналитическую группировку по данным таблицы 2.

Таблица 4. Вспомогательная таблица.

Основные средства, млрд. руб.	Число предприятий	Основные средства, млрд. руб.	Объем продукции, млрд. руб
А	1	2	3
1,2-2,2	1 2	1,8 1,2	2,2 1,0
Итого	--	3,0	3,2
2,2-3,2	3 4 5	2,8 2,4 3,1	2,8 2,9 3,6
Итого	--	8,3	9,3
3,2-4,2	7 8 9 12	3,9 3,5 3,5 4,0	4,3 4,6 4,5 4,8
Итого	--	14,9	18,2
4,2-5,2	10 11	4,3 4,5	5,6 5,4

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

	6	5,2	6,4
Итого	--	14,0	17,4
Всего	--	40,2	48,1

Название таблицы аналитической группировки: "Зависимость результативного признака от факторного признака"

Таблица 5. Зависимость объема продукции от основных средств.

Основные средства, млрд. руб.	Число предприятий	В среднем на одно предприятие	
		Основные средства, млрд.руб.	Объем продукции, млрд.. руб
1,2-2,2	2	1,5	1,6
2,2-3,2	3	2,8	3,1
3,2-4,2	4	3,7	4,5
4,2-5,2	3	4,7	5,8
Итого	12	3,3	4,0

Т.к. с увеличением основных средств происходит увеличение объемов продукции в среднем на 1 предприятие, то между этими двумя признаками прямая связь.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Рис.5. Зависимость объема продукции от основных средств.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

ТЕМА 4. АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

1) **Абсолютные величины.**
2) **Относительные величины.**

1) АБСОЛЮТНЫЕ ВЕЛИЧИНЫ

Изучая массовые общественные явления, статистика в своих выводах опирается на числовые данные, полученные в конкретных условиях места и времени. Результаты статистического наблюдения регистрируются, прежде всего, в форме первичных абсолютных величин.

Абсолютная величина – это простейший показатель характеризущий либо численность единиц совокупности (группы) либо сумму значений признака по группе.

Абсолютные величины являются результатом суммирования других величин. Иногда абсолютная величина является разностью двух других абсолютных величин.

Абсолютная величина всегда имеет единицу измерения.

В статистике все абсолютные величины являются именованными, измеряются в конкретных единицах и, в отличие от математического понятия абсолютной величины, могут быть как положительными, так и отрицательными (убытки, убыль, потери и т.п.).

Натуральные единицы измерения могут быть простыми (тонны, штуки, метры, литры) и сложными, являющимися комбинацией нескольких разноименных величин (грузооборот железнодорожного транспорта выражается в тонно-километрах, производство электроэнергии – в киловатт-часах).

В статистике применяют и абсолютные показатели, выраженные в **условно-натуральных единицах измерения** (например, различные виды топлива пересчитываются в условное топливо). В этом случае выбирают т.н. продукт-представитель и помощью коэффициентов пересчета переводят объемы всех других видов продукции в объем продукта-представителя. Продуктом-представителем выбирается продукт с наибольшим объемом производства или традиционный продукт.

Коэффициенты пересчета рассчитываются на основе определенной важной характеристики продукции:

Таблица 6. Пересчет моющих веществ в условное мыло 40% жирности

Виды моющих средств	Выпуск, тыс.т.	Содержание жира, %	Коэффициент пересчета	Выпуск условного мыла, тыс.т.
Мыло хозяйств.	53	40,0	1,0	53
Мыло туалет.	20	70,0	1,75	35
Порошок	54	20,0	0,5	27
Паста моющая	68	50,0	1,25	85
Итого	--	--	--	200

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Стоимостные единицы измерения используются, например, для выражения объема разнородной продукции в стоимостной (денежной) форме – рублях. При использовании стоимостных измерителей принимают во внимание изменения цен с течением времени. Этот недостаток стоимостных измерителей преодолевают применением "неизменных" или "сопоставимых" цен одного и того же периода.

В трудовых единицах измерения (человеко-днях, человеко-часах) учитываются общие затраты труда на предприятии, трудоемкость отдельных операций.

Человеко-день = Численность людей * Время

2) ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

Относительная величина в статистике – это обобщающий показатель, который дает числовую меру соотношения двух сопоставляемых абсолютных величин. Так как многие абсолютные величины взаимосвязаны, то и относительные величины одного типа в ряде случаев могут определяться через относительные величины другого типа.

Основное условие правильного расчета относительной величины – сопоставимость сравниваемых показателей и наличие реальных связей между изучаемыми явлениями. Таким образом, по способу получения относительные показатели – всегда величины производные, определяемые в форме коэффициентов, процентов, промилле, продецимилле и т.п. Однако нужно помнить, что этим безразмерным по форме показателям может быть, в сущности, приписана конкретная, и иногда довольно сложная, единица измерения. Так, например, относительные показатели естественного движения населения, такие как коэффициенты рождаемости или смертности, исчисляемые в промилле (‰), показывают число родившихся или умерших за год в расчете на 1 000 человек среднегодовой численности; относительная величина эффективности использования рабочего времени – это количество продукции в расчете на один отработанный человеко-час и т.д.

Относительная величина представляет дробь. Знаменатель дроби называется **базой сравнения.**

Если при расчете относительной величины база сравнения принимается за 1, то относительная величина выражается коэффициентом.

Если при расчете относительной величины база сравнения принимается за 100, то относительная величина в процентах.

В экономических расчетах чаще всего используются 4 относительные величины;

1) относительная величина структуры:

$$\frac{\textit{Число единиц совокупности, обладающих значением признака}}{\textit{Численность совокупности}}$$ – *доля* (*удельный вес*,%)

2) относительная величина планового задания:

$$\frac{\text{плановая величина}}{\text{базисная величина}}$$

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Пример: На предприятии планируют поднять заработную плату в этом квартале на 10%

$$\frac{110\%}{100\%}=1{,}1$$

3) относительная величина выполнения плана:

$$\frac{\text{Фактич. величина}}{\text{Выполнеие плана}}$$

План по росту заработной платы недовыполнен на 6%

$$\frac{100\%-6\%}{100\%}=0{,}94(94\%)$$

4) относительная величина динамики (темп роста):

$$\frac{\text{Фактич. величина}}{\text{Базисная величина}}$$

Взаимосвязь: $\frac{П}{Б}*\frac{Ф}{П}=\frac{Ф}{Б}$

Как фактически изменилась заработная плата на этом предприятии:
1,1*0,94=1,034 или 103,4%,
т.е. заработная плата увеличилась на 3,4%

Для более глубокого анализа можно рассчитать следующие относительные величины:

А) Относительные величины координации - характеризуют отношение частей данной совокупности к одной из них, принятой за базу сравнения. ОВК показывают, во сколько раз одна часть совокупности больше другой либо сколько единиц одной части приходится на 1, 10, 100, 1000, ... единиц другой части. Относительные величины координации могут рассчитываться и по абсолютным показателям, и по показателям структуры.

Б) Относительные величины сравнения - характеризуют сравнительные размеры одноименных абсолютных величин, относящихся к одному и тому же периоду либо моменту времени, но к различным объектам или территориям. Посредством этих показателей сопоставляются мощности различных видов оборудования, производительность труда отдельных рабочих, производство продукции данного вида разными предприятиями, районами, странами.

В) Относительные величины интенсивности - характеризуют степень распределения или развития данного явления в той или иной среде. Представляют собой отношение абсолютного уровня одного показателя, свойственного изучаемой среде, к другому абсолютному показателю, также присущему данной среде и, как правило, являющемуся для первого показателя факторным признаком. Так, при изучении демографических

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

процессов рассчитываются показатели рождаемости, смертности, естественного прироста и т.д. как отношение числа родившихся (умерших) или величины прироста населения за год к среднегодовой численности населения данной территории в расчете на 1000 чел. Если получаемые значения очень малы, то делают расчет на 10 000 человек.

Относительными величинами интенсивности выступают, например, показатели выработки продукции в единицу рабочего времени, затрат на единицу продукции, трудоемкости, эффективности использования производственных фондов и т.д., поскольку их получают сопоставлением разноименных величин, относящихся к одному и тому же явлению и одинаковому периоду или моменту времени. Метод расчета относительных величин интенсивности применяется при определении средних уровней (среднего уровня выработки, средних затрат труда, средней себестоимости изделий, средней цены и т.д.). Поэтому распространено мнение, что относительные величины интенсивности – это один из способов выражения средних величин.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

ТЕМА 5. СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ

1) **Сущность средних величин и их основные виды.**
2) **Свойства средней арифметической.**
3) **Центры распределения. Структурные средние.**
4) **Система показателей вариации.**
5) **Правило сложения дисперсий.**
6) **Дисперсия альтернативного признака.**

1) СУЩНОСТЬ СРЕДНИХ ВЕЛИЧИН И ИХ ОСНОВНЫЕ ВИДЫ.

Средняя величина – это обобщающий показатель, характеризующий типический уровень явления. Он выражает величину признака, отнесенную к единице совокупности.

Средняя всегда обобщает количественную вариацию признака, т.е. в средних величинах погашаются индивидуальные различия единиц совокупности, обусловленные случайными обстоятельствами. В отличие от средней абсолютная величина, характеризующая уровень признака отдельной единицы совокупности, не позволяет сравнивать значения признака у единиц, относящихся к разным совокупностям. Так, если нужно сопоставить уровни оплаты труда работников на двух предприятиях, то нельзя сравнивать по данному признаку двух работников разных предприятий. Оплата труда выбранных для сравнения работников может быть не типичной для этих предприятий. Если же сравнивать размеры фондов оплаты труда на рассматриваемых предприятиях, то не учитывается численность работающих и, следовательно, нельзя определить, где уровень оплаты труда выше. В конечном итоге сравнить можно лишь средние показатели, т.е. сколько в среднем получает один работник на каждом предприятии. Таким образом, возникает необходимость расчета средней величины как обобщающей характеристики совокупности.

Вычисление среднего – один из распространенных приемов обобщения; средний показатель выделяет то общее, что характерно (типично) для всех единиц изучаемой совокупности, в то же время он игнорирует различия отдельных единиц. В каждом явлении и его развитии имеет место сочетание случайности и необходимости. При исчислении средних в силу действия закона больших чисел случайности взаимопогашаются, уравновешиваются, поэтому можно абстрагироваться от несущественных особенностей явления, от количественных значений признака в каждом конкретном случае. В способности абстрагироваться от случайности отдельных значений, колебаний и заключена научная ценность средних как обобщающих характеристик совокупностей.

Виды средних и способы их вычисления

Рассмотрим теперь виды средних величин, особенности их исчисления и области применения. Средние величины делятся на два больших класса: степенные средние, структурные средние.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

К **степенным средним** относятся такие наиболее известные и часто применяемые виды, как средняя геометрическая, средняя арифметическая и средняя квадратическая.

В качестве **структурных средних** рассматриваются мода и медиана.

Степенные средние в зависимости от представления исходных данных могут быть **простыми** и **взвешенными**. ***Простая средняя*** считается по не сгруппированным данным и имеет следующий общий вид:

$$\overline{X} = \sqrt[m]{\frac{\sum X_i^m}{n}},$$

где X_i – варианта (значение) осредняемого признака;

m – показатель степени средней;

n – число вариант.

В зависимости от степени m получают различные виды средних величин.

Если же данные сгруппированы, то используется формулы **средних взвешенных**, где весами выступают частоты f (повторяемость варианты).

Взвешенная средняя считается по сгруппированным данным и имеет общий вид

$$\overline{X} = \sqrt[m]{\frac{\sum X_i^m f_i}{\sum f_i}}$$

где X_i – варианта (значение) усредняемого признака или серединное значение интервала, в котором измеряется варианта;

m – показатель степени средней;

f_i – частота, показывающая, сколько раз встречается i-е значение усредняемого признака.

Таблица 7. Виды степенных средних

Вид степенной средней	Показатель степени (m)	Формула расчета	
		Простая	Взвешенная
Гармоническая	-1	$\overline{X} = \frac{n}{\sum \frac{1}{x}}$	$\overline{X} = \frac{\sum m}{\sum \frac{m}{x}}$ $m = xf$
Геометрическая	0	$\overline{X} = \sqrt[n]{\Pi x} = \sqrt[n]{x_1 x_2 .}$	$\overline{X} = \sqrt[\sum f]{\Pi x^f} = \sqrt[\sum]{x_1^{f_1} x_2^{f_2} ..}$
Арифметическая	1	$\overline{X} = \frac{\sum x}{n}$	$\overline{X} = \frac{\sum xf}{\sum f}$

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

Квадратическая	2	$\overline{X} = \sqrt{\frac{\sum}{n}}$	$\overline{X} = \sqrt{\frac{\sum x^2 f}{\sum f}}$
Кубическая	3	$\overline{X} = \sqrt[3]{\frac{\sum}{n}}$	$\overline{X} = \sqrt[3]{\frac{\sum x^3 f}{\sum f}}$

Пример. Рассчитать среднюю заработную плату по двум бригадам:

Таблица 8. Оплата труда по бригадам

№ бригады	Средняя заработная плата, тыс.руб.	Численность бригады, чел.
1	150	25
2	180	15
Итого	161,25	40

$$\overline{X} = \frac{150*25+180*15}{25+15} = 161{,}25 \text{ (тыс.руб.)}$$

Среднее гармоническое взвешенное: $\overline{X} = \frac{\sum m}{\sum \frac{m}{x}}$

$m = xf$

Таблица 9. Оплата труда по бригадам

№ бригады	Средняя заработная плата, тыс.руб	Фонд заработной платы, тыс.руб.....
1	150	3750
2	180	2700
Итого	161,25	6450

$$\overline{X} = \frac{3750+2700}{\frac{3750}{150} + \frac{2700}{180}} = 161{,}25 \text{ (тыс.руб.)}$$

Прежде, чем выбрать формулу для расчетов средней величины нужно словами записать логическую сущность усредняемого признака.

Средняя заработная плата = Фонд заработной платы / численность работников

Средняя урожайность = Валовой сбор / Посевная площадь

Средняя производительность труда = Объем продукции / Численность (Время)

Правило: Если в представленной информации есть данные о числителе логической формулы, то для расчета средней величины используется среднее гармоническое.

Если представлены данные о знаменателе логической формулы, то для расчета средней величины используется среднее арифметическое.

2. СВОЙСТВА СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ

1) Среднее арифметическое постоянной величины равна этой постоянной:

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

$$A - const, A = \overline{A}$$

2) Алгебраическая сумма линейных отклонений варианты от средней арифметической равна 0:

$$\sum(x-\overline{x})=0 - \text{ для несгруппированных,}$$

$$\sum(x-\overline{x})f=0 - \text{ для сгруппированных;}$$

3) Минимальное: Сумма квадратов отклонений варианты от средней арифметической есть число минимальное:

$$\sum(x-\overline{x})^2 - \mathbf{min} \text{ (для несгруппированных),}$$

$$\sum(x-\overline{x})^2 f - \mathbf{min} \text{ (для сгруппированных);}$$

Расчетные свойства:

4) Если каждую варианту X уменьшить или увеличить на определенное число, то средняя величина уменьшается или увеличивается на это число.

5) Если каждую варианту X уменьшить или увеличить в одно и тоже число раз, то средняя величина уменьшается или увеличивается в это число раз.

6) Если каждую частоту f уменьшить или увеличить в одно и тоже число раз, то средняя величина не изменится.

$$\overline{x}=\frac{\sum xf}{\sum f}=\frac{\sum x\frac{f}{\sum f}}{\sum\frac{f}{\sum f}}=\sum xd$$

$$\frac{f}{\sum f}=\sum xd$$

Т.о., средняя величина зависит от варианты X и от структуры совокупности, которая характеризуется долями d.

7) Средняя суммы равна сумме средней:

$$\overline{X+Y}=\overline{X}+\overline{Y}$$

3. ЦЕНТРЫ РАСПРЕДЕЛЕНИЯ. СТРУКТУРНЫЕ СРЕДНИЕ.

Ряд распределения имеет 3 центра: 1) средняя арифметическая; 2) мода; 3) медиана.

1) Средняя арифметическая:

Рассчитаем среднюю арифметическую для дискретного ряда распределения представленного в таблице 1:

$$x=\frac{\sum xf}{\sum f}==\frac{2*40+3*64+...+6*14}{200}=3{,}53$$

При расчете средней величины по интервальному ряду распределения в качестве варианты X берется середина интервала. Если интервал открытый то при расчете средней величины его условно закрывают, принимая равным соседнему закрытому интервалу.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Рассчитаем среднюю величину основных средств по таблице 3:

$$x = \frac{\sum xf}{\sum f} == \frac{1{,}7*2 + 2{,}7*3 + 3{,}7*4 + 4{,}7*3}{12} = 3{,}4 \text{ млрд.руб.}$$

Расчет средней величины по "способу моментов": $\bar{x} = im_1 + A$, где m_1 - первый условный момент

Условный момент представляет среднюю арифметическую взвешенную из условных вариант $\frac{x - A}{i}$;

$$m_1 = \frac{\sum\left(\frac{x - A}{i}\right)f}{\sum f};$$

$$A = \frac{x_{\min} + x_{\max}}{2}.$$

В качестве *i* надо брать величину интервалов, если число интервалов нечетное и половину интервала, если число интервалов четное.

Рассчитаем среднюю величину основных средств способом моментов:

Таблица 10. К расчету средней величины способом моментов

x		-A	$\frac{x}{}$	$\frac{x - A}{i}$*f
1,7 2,7 3,7 4,7		1,5 0,5 ,5 ,5	-3 -1 13	-6 -3 4 9
Итого	2		-	4

$A = \frac{1{,}7 + 4{,}7}{2} = 3{,}2$,

i=0,5

m_1=4/12=0.33

$\bar{x}$ =0,5*0,33+3,2=3,4 (млрд.руб)

2) Мода

Мода – наиболее часто встречающаяся варианта.

Определим моду тарифного разряда по таблице 1:

$M_o = 3$

Для интервальных рядов распределения сначала находится модальный интервал, затем мода находится по формуле:

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

$$M_o = X_{M_o} + i\frac{f_{M_o} - f_{-1}}{(f_{M_o} - f_{-1}) + (f_{M_o} - f_{+1})}$$

X_{M_o} - нижняя граница модального интервала;

i- величина модального интервала;

f - частота модального интервала;

f_{-1} - частота интервала предшествующего модальному интервалу;

f_{+1}-- частота интервала следующего за модальным интервалом.

$$M_o = 3{,}2 + 1 * \frac{4-3}{(4-3)+(4-3)} = 3{,}7 \text{ млрд.руб.}$$

3) Медиана

Медиана- варианта, стоящая в середине ряда распределения.

Номер медианы:

№ Me=$\frac{n}{2}$ - если число единиц в совокупности четное;

№ Me=$\frac{n+1}{2}$ - если число единиц в совокупности нечетное.

Найдем медиану тарифного разряда по таблице 1:

$$№ \text{ Me}=\frac{200}{2}=100$$

Me=3

Следовательно, половина рабочих цеха имеет разряд не выше 3-го.

Прежде чем найти медиану для интервального ряда распределения ищут интервал, в который входит срединная варианта, затем внутри этого интервала определяют медиану по формуле:

$$Me = X_{Me} + i\frac{\frac{n}{2} - S_{-1}}{fMe},$$

где x_{Me}- нижняя граница медианного интервала;

i- величина медианного интервала;

n- число единиц совокупности;

S_{-1}- накопленная частота интервала предшествующего медианному;

fMe- частота медианного интервала

Найдем медиану основных средств по таблице 3:

$$№ \text{ Me}=\frac{12}{2}=6$$

$$Me = 3{,}2 + 1 * \frac{\frac{12}{2} - 5}{4} = 3{,}45 \text{ млрд.руб}$$

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

4) СИСТЕМА ПОКАЗАТЕЛЕЙ ВАРИАЦИИ

Для характеристики совокупности особенно в том случае если значение признака существенно колеблется, дополнительно к расчету средней величины определяют ряд показателей вариации.

Для измерения вариации используют абсолютные и относительные показатели.

1) **Размах вариации: $R=X_{max} - X_{min}$** – диапазон изменения признака.

2) **Среднее линейное отклонение** – показывает среднее отклонение варианты от средней величины:

$$l = \frac{\sum |x - \bar{x}|}{n}$$ - для не сгруппированных данных;

$$l = \frac{\sum |x - \bar{x}| f}{\sum f}$$ - для сгруппированных данных;

3) **Среднее квадратическое отклонение** - показывает среднее отклонение

от средней величины:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$ - для не сгруппированных данных;

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}}$$ - для сгруппированных данных;

Все 3 показателя имеют те же единицы измерения, что и признак.

4) **Дисперсия** – квадрат среднего квадратического отклонения:

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$ или $$\sigma^2 = \frac{\Sigma (x - \bar{x}_i)^2 f}{\Sigma f_i}$$

Не имеет единиц измерения.

Свойства дисперсии:

1) D(const)=0
2) Если каждую варианту X уменьшить или увеличить на одно и тоже число раз, то дисперсия не изменится;
3) Если каждую варианту X уменьшить или увеличить в одно и тоже число раз i, то дисперсия уменьшится или увеличится в i^2 раз.

Способы расчета дисперсии:

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

1) исходя из определения: $\sigma^2 = \frac{\Sigma (x - \overline{x}_i)^2 f}{\sum f_i}$

2) исходя из средней из квадратов вариант:

$$\sigma^2 = \overline{X^2} - \left(\overline{X}\right)^2; \ \overline{X^2} = \frac{\sum X_i^2 f_i}{\sum f_i};$$

3) по способу моментов: $\sigma^2 = i^2 (m_2 - m_1^2)$;

m_1 - первый условный момент;

m_2 - второй условный момент;

$$m_1 = \frac{\sum \left(\frac{x - A}{i}\right) f}{\sum f}; m_2 = \frac{\sum \left(\frac{x - A}{i}\right)^2 f}{\sum f}$$

Рассчитаем дисперсию тарифного разряда по данным таблицы 1 тремя способами:

1) $\sigma^2 = \frac{\Sigma (x - \overline{x}_i)^2 f}{f_i} = \frac{(2-3{,}53)^2 * 40 + (3-3{,}53)^2 * 64 + ... + (6-3{,}53)^2 * 14}{200} = 1{,}29$

2) $\sigma^2 = \overline{X^2} - \left(\overline{X}\right)^2$=13,75-3,53=1,29

$\overline{X^2} = \frac{\sum X_i^2 f_i}{\sum f_i} = \frac{2^2 * 40 + 3^2 * 64 + ... + 6^2 * 14}{200} 13{,}75$

3) **Таблица 10. К расчету средней величины способом моментов**

x		$x - A$	$\left(\frac{x - A}{i}\right.$	$\frac{x - A}{i} f$	$\left(\frac{x - A}{i}\right)^2 f$
2 3 4 5 6	0 4 0 2 4	-2 -1 0 1 2	4 1 0 1 4	-80 -64 0 22 28	160 64 0 22 56
Итого	00	-	-	-94	302

$A = \frac{2+6}{2} = 4$, i=1

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

$$m_1 = \frac{\sum\left(\frac{x-A}{i}\right)f}{\sum f} = \frac{-94}{200}; \quad m_2 = \frac{\sum\left(\frac{x-A}{i}\right)^2 f}{\sum f} = \frac{302}{200};$$

$$\sigma^2 = i^2(m_2 - m_1^2) = 1{,}29$$

Показатели относительного рассеивания.

Для характеристики меры колеблемости изучаемого признака исчисляются показатели колеблемости в относительных величинах. Они позволяют сравнивать характер рассеивания в различных распределениях (различные единицы наблюдения одного и того же признака в двух совокупностях, при различных значениях средних, при сравнении разноименных совокупностей). Расчет показателей меры относительного рассеивания осуществляют как отношение абсолютного показателя рассеивания к средней арифметической, умножаемое на 100%.

1. **Коэффициентом осцилляции** показывает относительную колеблемость крайних значений признака относительно средней.

$$K_0 = \frac{R}{\bar{x}} \cdot 100\%$$

2. **Относительное линейное отклонение** характеризует относительное усредненное значение абсолютных отклонений от средней величины.

$$Ke = \frac{l}{\bar{x}} * 100$$

3. **Коэффициент вариации** является наиболее распространенным показателем колеблемости, используемым для оценки типичности средних величин.

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

В статистике совокупности, имеющие коэффициент вариации больше 30–35 %, принято считать неоднородными.

Для анализа колеблемости признаков также используют показатели дифференциации.

1. По несгруппированным первичным данным можно рассчитать **коэффициент фондовой дифференциации**:

$$K_{ф.д.} = \frac{\overline{X}_{наибольшее}}{\overline{X}_{наименьшее}},$$

где $\overline{X}_{наибольшее}$ - средняя величина, рассчитанная для 10% самых больших значений признака.

$\overline{X}_{наименьшее}$ - средняя величина, рассчитанная для 10% самых маленьких значений признака.

2. Если данные сгруппированы, то рассчитывают **коэффициент децильной дифференциации**:

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

$K_{д.д.} = \frac{D_9}{D_1}$, где D_1 и D_9 - соответственно 1 и 9 децили.

Дециль - значение признака, которому в ряду распределения соответствует 10-я доля совокупности.

Процедура нахождения децилей аналогична процедуре нахождения медианы для интервального ряда распределения:

1) определяют № децили: для 1-й децили: № $D_1 = \frac{n}{10}$;

для 9-й децили: № $D_9 = \frac{9}{10}n$;

2) находят интервалы, в которые входят эти децили и внутри этих интервалов находят децили по формулам:

$$D_1 = X_{D_1} + i\frac{\frac{n}{10} - S_{-1}}{fD_1}; \qquad D_9 = X_{D9} + i\frac{\frac{9}{10}n - S_{-1}}{fD_9},$$

где X_{D_1} и X_{D9} - соответственно нижние границы интервалов, в которые входят 1 и 9 децили;

i - величина интервалов, в которые входят 1 и 9 децили;

fD_1 и fD_9 - соответственно частоты интервалов, в которые входят 1 и 9 децили;

S_{-1} - накопленная частота интервала, предшествующая децильному (в первой формуле для 1-й децили, во второй формуле для 2-й децили).

Таблица 11. Распределение населения района по среднедушевому доходу

Месячный среднедушевой доход, тыс.руб	Численность		Накопленные частоты
	тыс.чел.	в % к итогу	
20-40	9,2	6,2	9,2 (S_{-1})
D_1 - 40-60	25,2	17,0	34,4 (D_1)
60-100	32,9	22,2	67,3
100-150	30,0	20,2	97,3
150-200	27,4	18,5	124,7 (S_{-1})
D_9 - 200-300	15,5	10,5	140,2 (D_9)
300-500	4,9	3,3	145,1
500 и выше	3,1	2,1	148,2
Итого	148,2	100	-

№ $D_1 = \frac{n}{10} = \frac{148,2}{10} 14,8$

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

№ $D_9 = \frac{9}{10} n = 133{,}4$

$D_1 = 40 + 20\frac{14{,}8 - 9{,}2}{25{,}2} = 44{,}4 тыс.руб$ - max доход для 10% самого бедного населения;

$D_9 = 200 + 100\frac{133{,}4 - 124{,}7}{15{,}5} = 256{,}1 тыс.руб$ - min доход для 10% самого богатого населения.

$K_{д.д.} = \frac{D_9}{D_1} = \frac{256{,}1}{44{,}4} = 5{,}8$ (min доход самого богатого населения больше max дохода самого бедного населения в 5,8 раз).

5) ПРАВИЛО СЛОЖЕНИЯ ДИСПЕРСИЙ

На колеблемость результативного признака оказывает влияние множество факторных признаков.

Одной из задач статистики является определение влияния какого-либо факторного признака на колеблемость результативного признака. Всю колеблемость результативного признака измеряют т.н. общей дисперсией результативного признака.

$$\sigma^2 = \frac{\Sigma(x - \overline{x})^2}{n},$$

X – варианта результативного признака;

$\overline{x}$ - средняя величина результативного признака, рассчитанная по всей совокупности;

n – число единиц совокупности.

Для того чтобы установить влияние какого-то одного факторного признака на колеблемость результативного признака проводят аналитическую группировку по этому факторному признаку.

Результативный признак колеблется внутри каждой выделенной группы под влиянием других факторных признаков, которые не положены в основу аналитической группировки. Эту колеблемость измеряют с помощью внутригрупповой дисперсии:

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

i-номер группы;

$\overline{x}_i$ - внутригрупповая средняя результативного признака;

f_i - число единиц в группе.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Для того чтобы определить колеблемость результативного признака под влиянием колеблемости других признаков по всей совокупности рассчитывают среднюю из этих внутригрупповых дисперсий.

$$\overline{\sigma}_i^2 = \frac{\Sigma \sigma_i^2 f_i}{\Sigma f_i}.$$

Групповые средние колеблются вокруг общей средней. Эту колеблемость измеряют с помощью межгрупповой дисперсии:

$$\delta^2 = \frac{\Sigma (\overline{x}_i - \overline{x})^2 f}{\Sigma f}.$$

Межгрупповая дисперсия показывает колеблемость результативного признака под влиянием колеблемости факторного признака, по которому проводили аналитическую группировку. Т.о., всю колеблемость результативного признака разделили на 2 части:

1) колеблемость результативного признака под влиянием изучаемого факторного признака (меряется межгрупповой дисперсией);

2) колеблемость результативного признака под влиянием всех других факторных признаков (меряется средней из внутригрупповых дисперсий):

$\delta^2 + \overline{\sigma}_i^2 = \sigma^2$ - **Правило сложения дисперсий**

Если $\delta^2 = 0$, то связей между результативным и факторным признаками - нет.

Если $\sigma^2 = 0$, то между результативным и факторным признаками функциональная связь.

На основе правила сложения дисперсий построено 2 показателя тесноты связи между результативным и факторным признаками:

1) **Коэффициент детерминации**:

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

Коэффициент детерминации характеризует долю колеблемости результативного признака под влиянием изучаемого факторного признака.

$$0 \le \eta^2 \le 1$$

2) **Эмпирическое корреляционное отношение**:

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

$$-1 \le \eta \le 1.$$

Чем ближе этот коэффициент к -1 или 1, тем теснее связь между результативным и факторным признаками. Знак эмпирического корреляционного отношения ставится, исходя из сущности связи между результативным и факторным признаками.

РЕПОЗИТОРИЙ ГГУ ИМ. Ф. СКОРИНЫ

Пример: Рассчитать показатели тесноты связи между объемом продукции и стоимостью основных средств по данным таблиц 2,4,5.

$$\sigma^2 = \delta^2 + \overline{\sigma}_i^2,\ \overline{x} = 4$$

$$\sigma^2 = \frac{\Sigma(x-\overline{x})^2}{n} = \frac{(2{,}2-4)^2 + (1{,}0-4)^2 + \ldots + (4{,}8-4)^2}{12} = 2{,}22$$

$$\delta^2 = \frac{\Sigma(\overline{x}_i - \overline{x})^2 f}{\Sigma f} = \frac{(1{,}6-4)^2 * 2 + (3{,}1-4)^2 * 3 + \ldots + (5{,}8-4)^2 * 3}{4*12} = 2{,}06$$

$$\overline{\sigma}_i^2 = \frac{\Sigma \sigma_i^2 f_i}{\Sigma f_i}$$

$$\sigma_i^2 = \frac{\Sigma(x-\overline{x}_i)^2}{f_i},\ \sigma_1^2 = \frac{(2{,}2-1{,}6)^2 + (1{,}0-1{,}6)^2}{2} = 0{,}36$$

Аналогично рассчитываются 2,3 и 4 дисперсии.

$$\sigma_2^2 = 0{,}1267,\ \sigma_3^2 = 0{,}035,\ \sigma_4^2 = 0{,}1867$$

$$\overline{\sigma}_i^2 = \frac{0{,}36*2 + 0{,}1267*3 + 0{,}035*4 + 0{,}1867*3}{12} = 0{,}15$$

2,22=2,06+0,15 – проверка расчетов

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{2{,}06}{2{,}22} = 0{,}928 \quad (92{,}8\%)$$

Колеблемость объема продукции на 92,8% объясняется колеблемостью основных средств.

$$\eta = \sqrt{\eta^2} = \sqrt{0{,}928} = 0{,}963$$

Т.к. значение 0,963 близко к единице, то связь между объемом продукции и величиной основных средств высокая.

6) ДИСПЕРСИЯ АЛЬТЕРНАТИВНОГО ПРИЗНАКА

Распределение альтернативного признака имеет вид:

x	d
1	p
0	(1-p)
Итого	1

$$\overline{X} = \frac{\Sigma Xf}{\Sigma f} = \Sigma xd$$

$$\overline{x} = p$$

$$\sigma^2 = \frac{\Sigma(x-\overline{x})^2 f}{\Sigma f} = \Sigma(x-\overline{x})^2 d$$

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

Максимальное значение дисперсии альтернативного признака равно 0,25.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

ТЕМА 6. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

1) **Сущность выборочного наблюдения**
2) **Виды и способы выборочного наблюдения**
3) **Оценивание генеральной совокупности и генеральной доли**
4) **Определение численности выборки**

1) СУЩНОСТЬ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ

Выборочное наблюдение представляет собой такой вид статистического наблюдения, при котором обследованию подвергается некоторая часть единиц изучаемой совокупности, отобранная в определенном строго научном порядке, с целью последующей характеристики всей совокупности.

Выборочное наблюдение организуют т.о., чтобы отобранная часть единиц совокупности представляла всю совокупность. Отобранную часть единиц совокупности называют **выборочной совокупностью** или **выборкой**. Всю статистическую совокупность называют **генеральной**.

Показатели, которые рассчитанные на основе выборки, могут быть применены для оценки аналогичных показателей в генеральной совокупности.

Выборочная совокупность будет **представительной** или **репрезентативной** только в том случае, если соблюдаются принципы отбора выборки:

Основные принципы:

1) каждая единица генеральной совокупности должна иметь равную возможность попасть в выборку, т.е. должен соблюдаться принцип случайности отбора;

2) отобранная выборочная совокупность должна в основных чертах отражать структуру генеральной совокупности.

При проведении любого статистического наблюдения возникает ошибка регистрации. При проведении выборочного наблюдения кроме ошибки регистрации возникает ошибка выборки (ошибка репрезентативности), которая представляет собой разницу между величиной статистического показателя, рассчитанной по выборке, и величиной того же показателя, который был бы получен при проведении сплошного наблюдения.

Ошибки репрезентативности могут быть случайными и систематическими.

Случайная ошибка возникает всегда, ее величина зависит от следующих факторов:

1) способ формирования выборочной совокупности;

2) объем выборки;

3) степень колеблемости изучаемого признака генеральной совокупности генеральной совокупности: чем выше степень колеблемости, тем больше ошибка.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ

Систематические ошибки возникают тогда, когда нарушаются принципы отбора в выборочную совокупность.

Причины проведения выборочного наблюдения:

1) экономия средств и времени при проведении наблюдения;
2) невозможность сплошного наблюдения;
3) т.к. выборочное наблюдение обычно небольшое, то его проводят более тщательно, чем сплошное наблюдение, что приводит к резкому снижению ошибки регистрации. Суммарная ошибка регистрации и репрезентативности выборочного наблюдения достаточно часто будет меньше, чем ошибка регистрации сплошного наблюдения.

Сферы применения выборочного наблюдения:

1) статистический контроль качества продукции;
2) изучение общественного мнения;
3) контроль за использованием рабочего времени;
4) расширение программы населения.

2) ВИДЫ И СПОСОБЫ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ

1. Простая случайная выборка. При проведении такой выборки отбор единиц совокупности в выборочную совокупность проводится случайным образом без предварительного разделения генеральной совокупности на группы. В зависимости от способа отбора единиц совокупности различают повторную и бесповторную случайные выборки.

Повторная выборка осуществляется по схеме возвращенного шара.

Бесповторная выборка – по схеме невозвращенного шара.

2. Типическая (районированная) выборка. Этот тип выборки используется в том случае, если заранее известно, что генеральная совокупность неоднородна, и в ней представлены какие-то типы явлений. В этом случае генеральная совокупность разбивается на группы, и отбор производится из каждой группы в случайном порядке.

Число единиц совокупности, отобранных из каждой группы обычно пропорционально удельному весу численности групп в генеральной совокупности. Районированная выборка опирается на территориальное деление. Ошибка типической выборки всегда меньше, чем ошибка простой случайной выборки, т.к. группировка снижает колеблемость признака.

3. Серийная выборка. При проведении такой выборки в случайном порядке выбираются не единицы совокупности, а их группы, которые называются **сериями**. Ошибка серийной выборки больше, чем ошибка простой, т.к. для проверки отбирают малое число серий.

4. Механическая выборка. При проведении такой выборки все единицы совокупности располагаются в определенном порядке, а затем, в зависимости от заданного объема выборки отбирается, например каждая сотая единица совокупности, если проводится однопроцентная выборка.

Механическую выборка является бесповторной случайной выборкой. Если генеральная совокупность бесконечна, то ее, с точки зрения ошибки выборки, рассматривают как повторную простую выборку.

РЕПОЗИТОРИЙ ГГУ ИМ.Ф.СКОРИНЫ