

Дисциплина «Статистика» состоит из трех разделов:

1. **Общая теория статистики** - общественная наука, разрабатывающая общие понятия, категории, методы сбора, обработки, обобщения и анализа массовых данных.
2. **Социально-экономическая статистика** - общественная наука, изучающая количественную сторону массовых социальных и экономических явлений и процессов в неразрывной связи с их качественной стороной. Объектом изучения социально-экономической статистики является общество во всем многообразии его форм и проявлений. Все показатели социально-экономической статистики и методология их исчисления рассматриваются в свете теории и практики применения системы национальных счетов в условиях рыночной экономики.
3. **Статистика финансов** - общественная наука, характеризующая количественную сторону массовых явлений и процессов, происходящих в финансово-кредитной сфере, в неразрывной связи с их качественной стороной в целях выявления общих закономерностей и процессов в развитии социальной и экономической сферы общества, страны.

Раздел 1

Общая теория статистики

Тема 1. Предмет, метод и задачи общей теории статистики

Тема 2. Массовые статистические наблюдения

**Тема 3. Сводка и группировка статистических данных.
Статистические таблицы**

Тема 4. Абсолютные и относительные величины

Тема 5. Анализ рядов распределения

Тема 6. Анализ взаимосвязи между явлениями

Тема 7. Выборочное наблюдение

Тема 8. Индексы

Тема 9. Ряды динамики

Тема 1. Предмет, метод и задачи общей теории статистики

- 1.1. Предмет общей теории статистики**
- 1.2. Основные понятия теории статистики**
- 1.3. Стадии и методы статистического исследования**
- 1.4. Задачи статистики**

1.1. Предмет общей теории статистики

Изначально термин «**статистика**»

(происходит от лат. *status* - состояние, положение вещей) употреблялся в значении «политическое состояние» (отсюда итал. *stato* - государство и *statista* - знаток государства).



1623-1687

Уильям Петти (Petty) - английский экономист, родоначальник буржуазной классической политической экономии. Получил медицинское образование в университетах Лейдена, Парижа и Оксфорда. Обладал разносторонними способностями: в 1647 г. изобрел копировальную машину, в 1649 г. получил степень доктора физики, в 1651 г. стал профессором анатомии и музыки. Петти был крупным землевладельцем. В 1652 г. по поручению правительства Кромвеля провел «обзор земель» Ирландии. Петти выступил как идеолог английской буржуазии, усилившей свою власть после Английской буржуазной революции XVII в. Главные труды Петти: «Трактат о налогах и сборах» (1662), «Слово мудрым» (1665), «Политическая анатомия Ирландии» (1672), **«Политическая арифметика»** (1683) и др.

Экономическое развитие общества Петти ставил в зависимость от объективных законов, хотя и отождествлял общественные, экономические законы и законы природы, рассматривая их как вечные и неизменные. Метод, примененный Петти при исследовании экономических явлений, заимствован им из естественных наук и дополнен статистическим анализом. Петти - первый автор теории трудовой стоимости. Различал внутреннюю стоимость, которую называл «естественной ценой», и **рыночную цену**. Петти первым из экономистов поставил вопрос о дифференциальной земельной ренте. Научной была также постановка им вопроса о цене земли.

Взгляды Петти по вопросам экономической политики отражали тенденцию подчинения развития экономики страны интересам промышленного капитала, хотя он и считал закономерным вмешательство государства в регулирование народного хозяйства. Учение Петти в целом описательно, но при анализе ряда экономических явлений Петти приближается к раскрытию их сущности.



1796-1874

Адольф Кетле (Lambert-Adolph-Jacques Quetelet) главным образом известен как отец современной статистики. Посетив Англию, Шотландию, Швейцарию, Италию и Германию, Кетле, в 1832 г., принял заведование построенной по его плану в Брюсселе обсерваторией; в 1834 г. избран в постоянные секретари брюссельской академии наук; преподавал также астрономию и геодезию в Ecole militaire, состоял директором бельгийского статистического бюро и председателем учрежденной по его инициативе бельгийской центральной статистической комиссии.

Труды Кетле в области математики, физики, астрономии и в особенности метеорологии (наблюдения над температурой Земли, работы по электричеству воздуха, наблюдения над так называемыми воздушными волнами; по его мысли состоялся в 1873 г. в Вене первый международный метеорологический конгресс и было положено начало организации систематических наблюдений над метеорологическими явлениями одновременно в разных странах) имеют несомненную ценность, но главной заслугой Кетле являются его работы в области статистики. Всех сочинений Кетле по статистике насчитывается 65.

Весьма существенное значение имела практическая деятельность Кетле как организатора первого международного статистического конгресса. Кетле был проникнут убеждением, что общественные явления могут и должны быть изучаемы только на основании правильно устроенного систематического наблюдения, и во все продолжение своей долгой жизни упорно трудился над осуществлением этой мысли; все организаторы статистических учреждений в Европе с середины 50-х годов были его учениками, и до самого конца своей жизни, на целом ряде статистических конгрессов, с брюссельского (1862) до петербургского (1872) включительно, Кетле поддерживал их своей опытностью. Если в настоящее время сделалась, до известной степени, возможной *сравнительная* статистика, то исключительно благодаря некоторому объединению принятых в различных странах способов и приемов наблюдения над явлениями общественной жизни, к каковому объединению Кетле постоянно стремился.



1711-1765

Первый русский профессор и член Петербургской Академии наук, человек энциклопедических знаний, разносторонних интересов и способностей. В центре внимания экономических воззрений Ломоносова стоял вопрос об обеспечении экономической независимости и самостоятельности России. Возглавляя в 1758 г. Географический департамент АН, Ломоносов выдвинул идею плана экономико-географического атласа России, отличающегося от ранее изданного Академией наук «Атласа Российского» (1745). Для получения необходимых сведений Ломоносов в 1759 г. разработал, в частности, проект статистической анкеты (формы запросов). С академической деятельностью Ломоносова, превратившего Географический департамент в центр статистико-географического изучения хозяйства России, связаны работы по созданию «Экономического лексикона российских продуктов» (1763). В начатой работе над «Лексиконом» Ломоносов пытался составить целостное представление о структуре производства и обращении продуктов, о движении товарных потоков в масштабах всей страны с выделением конкретных видов и региональной характеристикой товаров. Ломоносов придавал большое значение изучению населения. В трактате «О сохранении и размножении Российского народа» (1761) Ломоносов дал глубокую характеристику демографической ситуации в России в середине XVIII в., обосновал необходимость принятия мер по стимулированию рождаемости, снижению смертности, в т. ч. детской, улучшению миграционных процессов. Этот первый русский трактат о политике населения положил начало демографической науке в России. Важную роль в широком статистико-географическом описании страны Ломоносов отводил публикации текущих сведений о производстве и торговле. Он обосновал в 1759 г. необходимость издания на русском языке экономической газеты «Внутренние российские ведомости» и позднее $\frac{3}{4}$ экономического журнала «Экономические и физические сочинения». Ломоносов $\frac{3}{4}$ основоположник названной им русской экономической географии, отличающейся от господствовавшего формально описательного государственоведения. География была у Ломоносова, в сущности, своеобразной системой статистики в понимании XVIII в., которая в основном близка к современной экономической географии.

Предметом статистики является количественная сторона массовых социально-экономических явлений в непосредственной связи с качественным содержанием, конкретными условиями, местом и временем.

Из данного определения следуют **основные черты предмета статистической науки:**

1. Статистика - наука общественная.
2. В отличие от других общественных наук статистика изучает количественную сторону общественных явлений.
3. Статистика изучает массовое явление.
4. Статистика изучает количественную сторону явлений в неразрывной связи с количественной стороной и это находит свое воплощение в существовании системы статистических показателей.
5. Статистика изучает количественную сторону явлений в конкретных условиях места и времени.

Статистика изучает также влияние природных и технических факторов на количественные отношения общественной жизни и влияние жизнедеятельности общества на среду обитания.

1.2. Основные понятия теории статистики

1. Статистическая совокупность – это множество единиц изучаемого явления, объединенных в соответствии с задачей исследования единой качественной основой, общей связью, но отличающихся друг от друга отдельными признаками.

2. Статистический показатель – это количественная оценка свойства изучаемого явления.

Виды статистических показателей

Учетно-оценочные показатели
(размеры, объемы,
уровни изучаемого явления)

Аналитические показатели
(относительные и средние
величины,
показатели вариации и т.д.).

3. Признак – это качественная особенность единицы совокупности.

Показатель:

{Качественная сторона} = {Количественная сторона}



4. Статистическая закономерность - это форма проявления причинной связи, выражающаяся в последовательности, регулярности, повторяемости событий с достаточно высокой степенью вероятности, если причины (условия), порождающие события не изменяются или изменяются незначительно.

Статистические закономерности устанавливаются на основе анализа массовых данных.

1.3. Стадии и методы статистического исследования

Статистическое исследование количественной стороны общественных явлений проходит *три стадии*.

- 1. Сбор первичной статистической информации.** На этой стадии применяются *методы массового наблюдения*, так как изучаемые статистикой закономерности проявляются в достаточно большом массиве данных на основе действия *закона больших чисел*.
- 2. Статистическая сводка и обработка первичной информации.** Важнейшим методом второй стадии является *метод статистических группировок*, позволяющий выделить однородные совокупности, разделить их на группы и подгруппы. На этой стадии переходят от описания отдельных единиц к описанию их групп и объекта в целом посредством подсчета итогов, вычисления обобщающих показателей в виде относительных средних величин.
- 3. Анализ статистической информации**, который позволяет раскрыть причинные связи изучаемых явлений, определить влияние и взаимодействие различных факторов, оценивать эффективность принимаемых управленческих решений, возможные экономические и социальные последствия складывающихся ситуаций. Здесь применяется весь арсенал статистических методов - ряды динамики, индексы, методы математической статистики и т. д. Выводы и анализ излагаются в текстовой форме и сопровождаются таблицами и графиками.

1.4. Задачи статистики

На каждом этапе развития перед статистикой встают специфические задачи, обусловленные характером самого этапа. В условиях **рыночной экономики** социально-экономическая статистика призвана решать новые задачи. Для этого требуются:

- повышение ее качества и оперативности,
- совершенствование отчетности,
- углубление социально-экономического анализа.

Особое внимание должно быть уделено

- совершенствованию методологии анализа важнейших пропорций:
 - между производством и потреблением,
 - потреблением и накоплением,
 - между производством средств производства и производством предметов потребления,
 - между отдельными отраслями;
- изучению структуры экономики и технико-экономических сдвигов, научно-технического прогресса;
- выявлению диспропорций, которые могут возникнуть в экономике; вскрытию и более полному использованию всех возможностей рыночной экономики.

Большое значение имеет также оценка состояния экономики и уровня жизни населения.

Практическое решение этих задач призвана осуществлять система органов государственной статистики. Конкретные задачи определены:

- Гражданским кодексом Российской Федерации (часть первая от 30.11.94 № 51-ФЗ);
- Федеральным законом № 282-ФЗ от 29 ноября 2007 года «Об официальном статистическом учете и системе государственной статистики в Российской Федерации»;
- Федеральным законом №24-ФЗ от 20 февраля 1995 «Об информации, информатизации и защите информации»;
- Федеральным законом № 184-ФЗ от 27 декабря 2002 «О техническом регулировании»;
- Федеральным законом № 129-ФЗ от 08 августа 2001 «О государственной регистрации юридических лиц и индивидуальных предпринимателей»);
- Указом Президента РФ от 9 марта 2004 г. № 314 «О системе и структуре федеральных органов исполнительной власти»;
- Постановлениями Правительства РФ от 7 апреля 2004 г. № 188 «Вопросы Федеральной службы государственной статистики» и от 30 июля 2004 г. № 399 «Об утверждении Положения о Федеральной службе государственной статистики».

В постановлении Правительства РФ от 30 июля 2004 г. отмечается, что **Федеральная служба государственной статистики** (далее – Служба) осуществляет следующие полномочия в установленной сфере деятельности:

- вносит в Правительство РФ проекты федеральных законов, нормативных правовых актов Президента РФ и Правительства РФ и другие документы, по которым требуется решение Правительства РФ, а также проект ежегодного плана работы и прогнозные показатели деятельности Службы;
- на основании и во исполнение Конституции Российской Федерации, федеральных конституционных законов, федеральных законов, актов Президента РФ и Правительства РФ самостоятельно принимает нормативные правовые акты по вопросам в установленной сфере деятельности, за исключением вопросов, правовое регулирование которых в соответствии с Конституцией РФ и федеральными конституционными законами, федеральными законами, актами Президента РФ и Правительства РФ осуществляется исключительно федеральными конституционными законами, федеральными законами, нормативными правовыми актами Президента РФ и Правительства РФ;
- представляет официальную статистическую информацию Президенту РФ, Правительству РФ, Федеральному Собранию Российской Федерации, иным органам государственной власти, органам местного самоуправления, средствам массовой информации, организациям и гражданам, а также международным организациям;
- разрабатывает официальную статистическую методологию для проведения **статистических наблюдений** и формирования **статистических показателей**, в пределах своей компетенции обеспечивает соответствие указанной методологии международным стандартам;
- осуществляет подготовку, проведение и подведение итогов Всероссийской переписи населения, а также ее методологическое обеспечение;
- осуществляет подготовку, проведение и методологическое обеспечение статистических обследований;
- обеспечивает заинтересованных пользователей данными бухгалтерской отчетности **юридических лиц**, осуществляющих свою деятельность на территории Российской Федерации;
- организует деятельность федеральных органов исполнительной власти по формированию государственных информационных ресурсов в области государственной статистики;
- разрабатывает и ведет общероссийские классификаторы технико-экономической и социальной информации;
- проводит конкурсы и заключает государственные контракты на размещение заказов на поставку товаров, выполнение работ и оказание услуг для нужд Службы, а также на проведение научно-исследовательских работ для государственных нужд;
- обобщает практику применения законодательства Российской Федерации в установленной сфере деятельности;
- осуществляет функции главного распорядителя и получателя средств федерального бюджета, предусмотренных на содержание Службы и реализацию возложенных на Службу функций;
- организует прием граждан, обеспечивает своевременное и полное рассмотрение устных и письменных обращений граждан, принятие по ним решений и направление ответов заявителям в предусмотренный законодательством Российской Федерации срок;
- обеспечивает в пределах своей компетенции защиту сведений, составляющих государственную тайну;
- обеспечивает в пределах своей компетенции соответствующий режим хранения и защиты полученной в процессе деятельности Службы информации, составляющей служебную, банковскую, налоговую, коммерческую тайну, и иной конфиденциальной информации;
- обеспечивает мобилизационную подготовку Службы, а также контроль и координацию деятельности подведомственных организаций по их мобилизационной подготовке;
- организует профессиональную подготовку работников Службы, их переподготовку, повышение квалификации и стажировку;
- взаимодействует с соответствующими органами государственной власти иностранных государств и международными организациями;
- ведет в соответствии с законодательством Российской Федерации работу по комплектованию, хранению, учету и использованию архивных документов, образовавшихся в процессе деятельности Службы;
- осуществляет иные полномочия, если такие полномочия предусмотрены федеральными законами, нормативными правовыми актами Президента РФ или Правительства РФ.

С целью реализации полномочий в установленной сфере деятельности Федеральная служба государственной статистики имеет право:

- запрашивать и получать сведения, необходимые для принятия решений по вопросам, отнесенным к компетенции Службы;
- заказывать проведение необходимых исследований, испытаний, анализа и оценок, а также научных исследований по вопросам надзора в установленной сфере деятельности;
- давать юридическим и **физическим лицам** разъяснения по вопросам, отнесенным к компетенции Службы;
- осуществлять контроль над деятельностью территориальных органов Службы и подведомственных организаций;
- привлекать для проработки вопросов в установленной сфере деятельности научные и иные организации, ученых и специалистов;
- применять предусмотренные законодательством Российской Федерации меры ограничительного, предупредительного и профилактического характера, направленные на недопущение и (или) пресечение нарушений юридическими лицами и гражданами обязательных требований в установленной сфере деятельности, а также меры по ликвидации последствий указанных нарушений;
- создавать координационные, совещательные и экспертные органы (советы, комиссии, группы, коллегии), в том числе межведомственные;
- учреждать знаки отличия и награждать ими граждан за высокие достижения в установленной сфере деятельности.



содержание



поиск

- ▶ [РАЗДЕЛ С ДОБЫЧА ПОЛЕЗНЫХ ИСКОПАЕМЫХ](#)
- ▶ [РАЗДЕЛ D ОБРАБАТЫВАЮЩИЕ ПРОИЗВОДСТВА](#)
 - [Подраздел DA Производство пищевых продуктов, включая напитки, и табака](#)
 - [Подраздел DB Текстильное и швейное производство](#)
 - [Подраздел DC Производство кожи, изделий из кожи и производство обуви](#)
 - [Подраздел DD Обработка древесины и производство изделий из дерева](#)
 - [Подраздел DE Целлюлозно-бумажное производство](#)
 - [Подраздел DF Производство кокса, нефтепродуктов и ядерных материалов](#)
 - [Подраздел DG Химическое производство](#)
 - [Подраздел DH Производство резиновых и пластмассовых изделий](#)
 - [Подраздел DI Производство прочих неметаллических минеральных продуктов](#)
 - [Подраздел DJ Металлургическое производство и производство готовых металлических изделий](#)
 - [Подраздел DK Производство машин и оборудования](#)
 - [Подраздел DL Производство электрооборудования, электронного и оптического оборудования](#)
 - [Подраздел DM Производство транспортных средств и оборудования](#)
 - [Подраздел DN Прочие производства](#)
- [Раздел E Производство и распределение электроэнергии, газа и воды](#)
- [Раздел F Строительство](#)
- [Раздел G Оптовая и розничная торговля](#)
- [Раздел H Гостиницы и рестораны](#)
- [Раздел I Транспорт и связь](#)
- [Раздел J Финансовая деятельность](#)
- [Раздел K Операции с недвижимым имуществом, аренда и предоставление услуг](#)
- [Раздел L Государственное управление и обеспечение военной безопасности](#)
- [Раздел M Образование](#)
- [Раздел N Здравоохранение и предоставление социальных услуг](#)
- [Раздел O Предоставление прочих коммунальных, социальных и персональных услуг](#)
- [Раздел P Предоставление услуг по ведению домашнего хозяйства](#)

Тема 2. Массовые статистические наблюдения

- 2.1. Основные требования, предъявляемые к массовым статистическим наблюдениям
- 2.2. Основные организационные формы, виды и способы статистического наблюдения
- 2.3. Организационный план статистического наблюдения
- 2.4. Ошибки статистического наблюдения

2.1. Основные требования, предъявляемые к массовым статистическим наблюдениям

1. Точность и достоверность.
2. Полнота в отношении охвата данных:
 - в пространстве,
 - во времени,
 - по важности признаков.
3. Сопоставимость, единообразие.
4. Учет последующей обработки данных.
5. Своевременность представления.

2.2. Основные организационные формы, виды и способы статистического наблюдения

В российской статистике используются *три основные организационные формы* статистического наблюдения:

1. **Статистическая отчетность** (предприятий, организаций, учреждений и т. п.). Основная форма статистического наблюдения, с помощью которой статистические органы в определенные сроки получают от предприятий, организаций и учреждений необходимые данные в виде установленных в законном порядке отчетных документов, скрепленных подписями лиц, ответственных за представление этих документов и достоверность собираемых сведений. Действующая отчетность делится на *типовую и специализированную*. По срокам представления отчетность бывает *ежедневная, недельная, двухнедельная, месячная, квартальная и годовая*.
2. **Специально организованное статистическое наблюдение** (переписи, единовременные учеты и обследования) проводится для получения данных, отсутствующих в отчетности, или для проверки ее данных.
3. **Регистры**. Регистровое наблюдение - форма непрерывного наблюдения за долговременными процессами, имеющими фиксированное начало, стадию развития и фиксированный конец. Оно основано на ведении статистического регистра. Регистр представляет собой систему, постоянно следящую за состоянием единицы наблюдения, которая характеризуется совокупностью показателей. Все показатели хранятся до полного завершения наблюдения за единицей обследуемой совокупности.

В практике статистики различают регистры населения и регистры предприятий.

Регистр населения - поименованный и регулярно актуализируемый перечень жителей страны. Программа наблюдения содержит общие признаки: пол, дата, место рождения, дата вступления в брак, брачное состояние. Регистр населения, как любой регистр, охватывающий наблюдением значительную совокупность единиц, содержит данные по ограниченному числу признаков.

Регистр предприятий охватывает все виды экономической деятельности и содержит значения основных признаков по каждой единице наблюдаемого объекта за определенный период или момент времени. Регистры предприятий включают в себя данные о времени создания (регистрации) предприятия, его название и адрес, телефон, сведения об организационно-правовой форме, структуре, видах экономической деятельности, количестве занятых и др.

Единый государственный регистр предприятий и организаций всех форм собственности (**ЕГРПО**) дает возможность организовать сплошное наблюдение, а по ограниченному кругу статистических показателей предприятий, зарегистрированных на территории России, позволяет получать непрерывные ряды показателей в случае изменения территориальной, отраслевой и других структур совокупности.

Регистр содержит данные о таких показателях, как среднесписочная численность работников, средства, направляемые на потребление, остаточная стоимость основных средств, балансовая прибыль (убыток), уставный фонд.

ЕГРПО позволяет проводить отбор и группировку любой совокупности единиц по одному или нескольким признакам.

Данные о единицах наблюдения собираются в процессе государственной регистрации предприятий и последующего учета. При закрытии предприятия ликвидационная комиссия в десятидневный срок уведомляет об этом службу ведения регистра.

Пользователями регистра могут быть любые **юридические** или **физические лица**, заинтересованные в получении информации.

Виды статистического наблюдения классифицируются по следующим признакам:

■ **времени регистрации фактов:**

- 1) *непрерывное (текущее),*
- 2) *периодическое и*
- 3) *единовременное;*

■ **охвату единиц совокупности:**

- 1) *сплошное и*
- 2) *несплошное.*

Организационные формы несплошного наблюдения (различаются способом отбора единиц наблюдения):

- *выборочное (единицы отбираются в случайном порядке);*
- *метод основного массива (отбираются единицы, составляющие наибольший удельный вес в совокупности);*
- *монографическое (отбирается одна единица совокупности, обладающая наиболее характерными признаками).*

Способами статистического наблюдения являются

- ***непосредственное наблюдение*** (замер, подсчет);
- ***документальный учет фактов*** (необходимые сведения берутся из соответствующих документов);
- ***опрос*** (сведения фиксируются со слов опрашиваемого).

В статистике применяются следующие **виды опросов:**

- ***экспедиционный (устный);***
- ***саморегистрации;***
- ***явочный;***
- ***корреспондентский;***
- ***анкетный.***

2.3. Организационный план статистического наблюдения

Вопросы организационного плана статистического наблюдения

Программно-методологические вопросы	Организационные вопросы
Цель наблюдения	Органы наблюдения
Объект наблюдения - статистическая совокупность, в которой протекают исследуемые социально-экономические явления и процессы. Указываются основные отличительные черты объекта	Время наблюдения – выбор сезона, периода, критического момента наблюдения
Единица наблюдения - составной элемент объекта, являющийся носителем признаков, подлежащих регистрации	Подготовительная работа – составление списка отчетных единиц, подбор и подготовка кадров, подготовка статистического инструментария (размножение бланков)
Отчетная единица (единица совокупности) - субъект, от которого поступают данные о единице наблюдения	
Программа наблюдения - это перечень признаков (или вопросов), подлежащих регистрации в процессе наблюдения. Вопросы в программе могут быть закрытыми и открытыми.	Пропаганда (реклама) проводимых статистических наблюдений
Инструментарий статистического наблюдения – статистические формуляры (индивидуальные или списочные) и инструкции к ним	

2.4. Ошибки статистического наблюдения

Точностью статистического наблюдения называют степень соответствия величины какого-либо показателя, определенной по материалам статистического наблюдения, действительной величине этого показателя.

Расхождение между расчетными и действительными значениями изучаемых величин называется **ошибкой наблюдения**.

В зависимости от причин возникновения различают ошибки регистрации и ошибки репрезентативности.

Ошибки регистрации - это отклонения между значением показателя, полученным в ходе статистического наблюдения, и фактическим, действительным значением показателя.

Ошибка репрезентативности - отклонение значения показателя обследованной совокупности от его величины по исходной совокупности.

Виды ошибок
массового статистического
наблюдения

Ошибки регистрации:

- случайные и
- систематические
(преднамеренные и
непреднамеренные)

Характерны как для
сплошного, так и
несплошного наблюдения

*Ошибки
репрезентативности:*

- случайные (оцениваются с
помощью аппарата
математической статистики) и
- систематические (нельзя
оценить с помощью аппарата
математической статистики)

Характерны для
несплошного наблюдения

После получения статистических формуляров следует провести проверку полноты и качества собранных данных.

Контроль полноты - это проверка того, насколько полно объект охвачен наблюдением, иначе говоря, о всех ли единицах наблюдения собраны сведения.

Контроль качества материала осуществляется с помощью *логического и арифметического* контроля.

Тема 3. Сводка и группировка статистических данных. Статистические таблицы

Программа сводки включает определение:

- групп и подгрупп;
- системы показателей;
- видов таблиц.

С помощью *метода группировок* решаются следующие задачи:

- выделение социально-экономических типов явлений;
- изучение структуры явления и структурных сдвигов, происходящих в нем;
- выявление связи и зависимости между явлениями.

Группировка – это разбиение совокупности на группы, однородные по какому-либо признаку.

Классификация – это узаконенная, общепринятая, нормативная группировка.

Метод группировки основывается на *двух категориях*:

- 1) *группировочный признак*;
- 2) *интервал*.

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.

Для определения число групп с используется формула Стерджесса:

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

где n - число групп;

N - число единиц совокупности.

Длина интервала:

$$L = R_{max} / n = (x_{max} - x_{min}) / n$$

Классификационный признак	Виды группировок	Примечание
Цель исследования	<i>Типологическая</i>	Решает задачу выявления и характеристики социально-экономических типов (частных подсовокупностей).
	<i>Структурная</i>	Дает возможность описать составные части совокупности или строение типов, а также проанализировать структурные сдвиги.
	<i>Аналитическая (факторная)</i>	Позволяет оценивать связи между взаимодействующими признаками.
Число признаков	<i>Простая</i>	Среди простых группировок особо выделяют ряды распределения. <i>Ряд распределения</i> – это группировка, в которой для характеристики групп применяется один показатель – численность групп. Выделяют <i>атрибутивные, вариационные (дискретные и непрерывные)</i> ряды распределения.
	<i>Многомерная</i>	Выделяют три меры сходства: коэффициенты подобия; коэффициенты связи; показатели расстояния. Методы многомерной классификации: метод дендритов, метод шаров, метод корреляционных плеяд, многомерная средняя, кластерный анализ.
Отношение между признаками	<i>Иерархические</i>	Выполняется по двум и более признакам, при этом значения второго признака определяются областью значения первого.
	<i>Неиерархические</i>	Строгой зависимости между признаками не существует.
Очередность обработки информации	<i>Первичные</i>	Составлены на основе первичных данных
	<i>Вторичные</i>	Результат перегруппировки ранее сгруппированного материала
Временной критерий	<i>Статические</i>	Характеристика совокупности на определенный момент или за определенный период
	<i>Динамические</i>	Характеристика динамики совокупности (матрицы перехода, миграционные матрицы или матрицы мобильности).

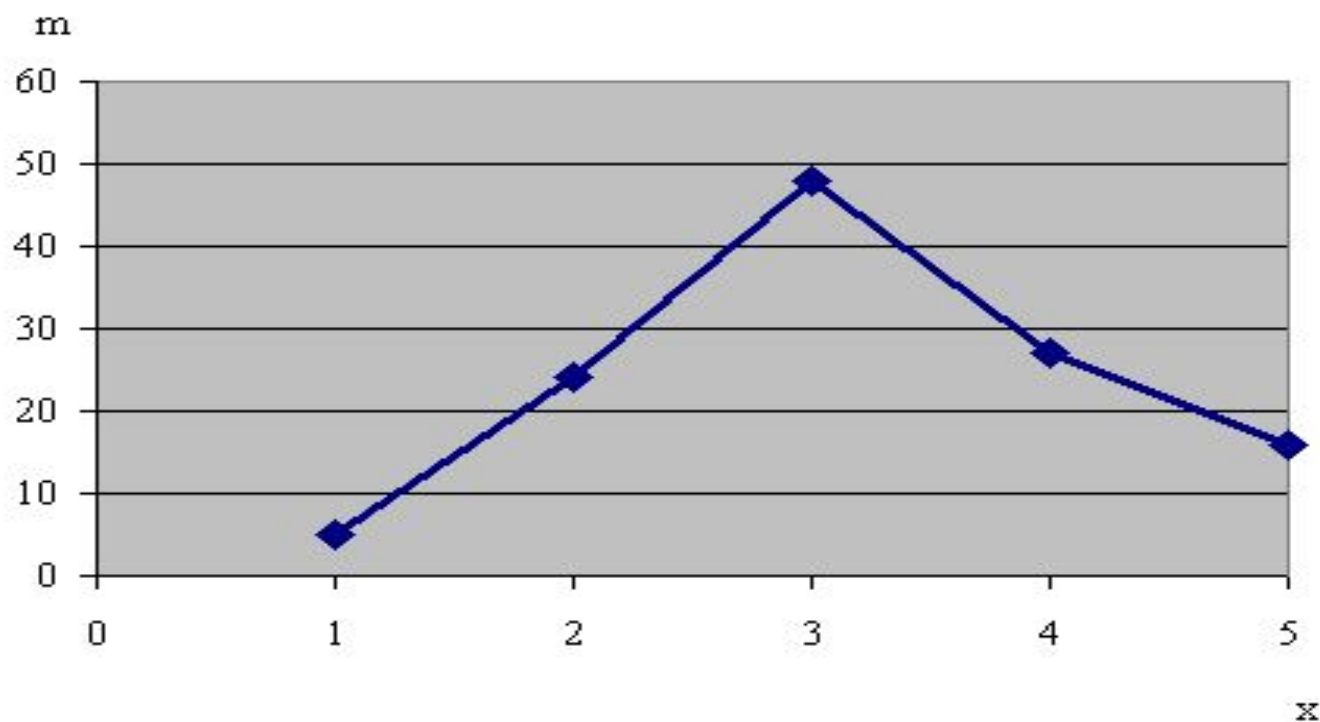
Дискретный ряд

Типологическая группировка

Структурная группировка

Разряд, x	Число рабочих, m	Разряд, x	Число рабочих (абсолютная частота), m	Относительная частота, f
1	5	1	5	4%
2	24	2	24	20%
3	48	3	48	40%
4	27	4	27	23%
5	16	5	16	13%
Итого	120	Итого	120	100%

Полигон - Распределение рабочих по разрядам



Вариационный ряд в равных интервалах

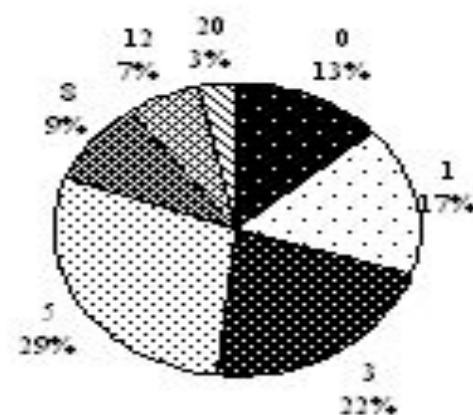
Структурная группировка

Интервалы по x		Число рабочих, m	Относительная частота, f
Начало интервала	Конец интервала		
90	100	4	3%
100	110	13	11%
110	120	21	18%
120	130	36	30%
130	140	28	23%
140	150	11	9%
150	160	3	3%
160	170	4	3%
Итого		120	100%



Вариационный ряд в неравных интервалах

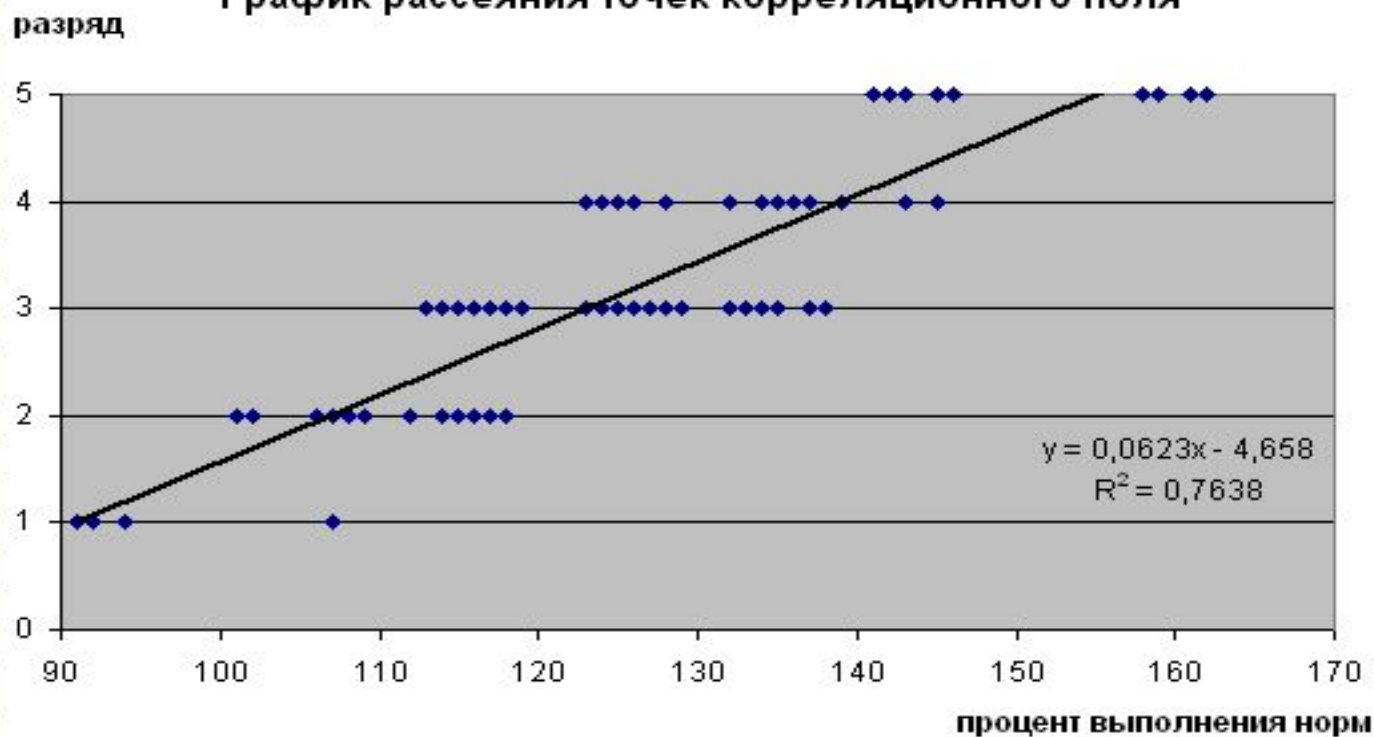
Интервалы по x		Число рабочих, m	Относительная частота, f
Начало интервала	Конец интервала		
0	1	16	13%
1	3	20	17%
3	5	26	22%
5	8	35	29%
8	12	11	9%
12	20	8	7%
20	40	4	3%
Итого		120	100%



Аналитическая (факторная группировка)

Процент выполнения норм, x		Разряд, y					Итого
Начало интервала	Конец интервала	1	2	3	4	5	
90	100	4					4
100	110	1	12				13
110	120		12	9			21
120	130		26	10			36
130	140			13	15		28
140	150				2	9	11
150	160					3	3
160	170					4	4
Итого		5	50	32	17	16	120

График рассеяния точек корреляционного поля



Методы многомерной классификации. Кластерный анализ

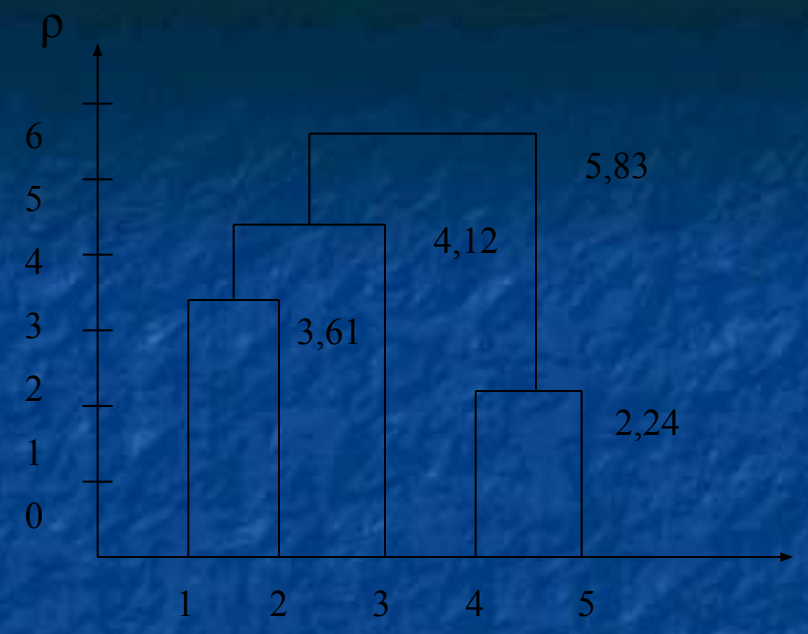
Cluster (англ.) – группа элементов,
характеризуемых каким-либо общим свойством

Пример

Провести классификацию $n=5$ семей по двум показателям: уровень расходов (млн руб.) за летние месяцы на культурные нужды, спорт и отдых – x^1 и питание x^2 .

С8 $f_8 = ((B2-B3)^2 + (C2-C3)^2) * 0,5$

	A	B	C	D	E	F
1	N семьи (i)	x_i^1	x_i^2			
2	1	2	10			
3	2	4	7			
4	3	8	6			
5	4	12	11			
6	5	13	9			
7		1	2	3	4	5
8	1	0,00	3,61	7,21	10,05	11,05
9	2	3,61	0,00	4,12	8,94	9,22
10	3	7,21	4,12	0,00	6,40	5,83
11	4	10,05	8,94	6,40	0,00	2,24
12	5	11,05	9,22	5,83	2,24	0,00
13		1	2	3	45	
14	1	0,00	3,61	7,21	10,05	
15	2	3,61	0,00	4,12	8,94	
16	3	7,21	4,12	0,00	5,83	
17	45	10,05	8,94	5,83	0,00	
18			12	3	45	
19		12	0,00	4,12	8,94	
20		3	4,12	0,00	5,83	
21		45	8,94	5,83	0,00	
22				12 3	45	
23			12 3	0,00	5,83	
24			45	5,83	0,00	



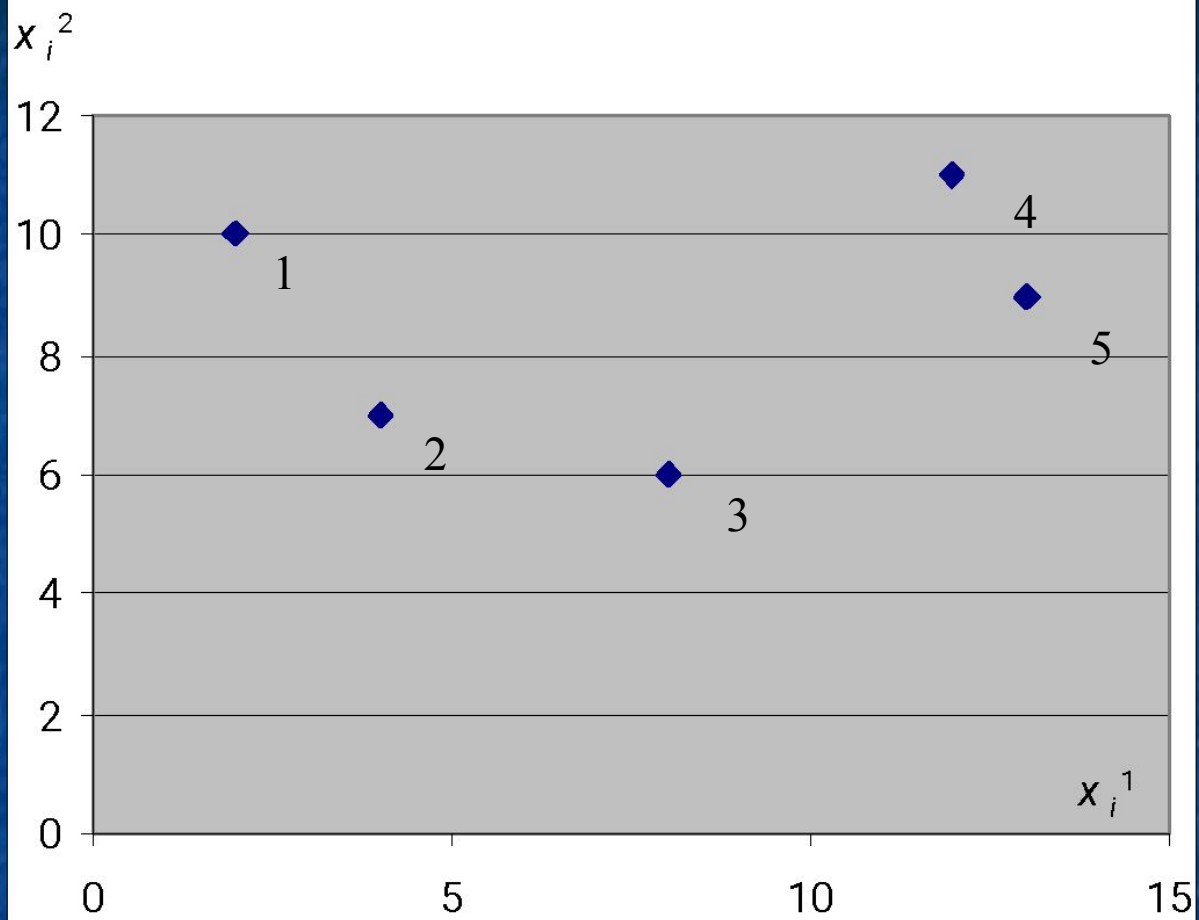
Расстояние между наблюдениями 1 и 2 (евклидово расстояние):

$$\rho_{1,2} = \sqrt{\sum_{j=1}^p (x_1^{(j)} - x_2^{(j)})^2} = \sqrt{(2-4)^2 + (10-7)^2} = 3,61$$

Расстояние между кластерами по принципу "ближайшего соседа":

$$\rho_{1,(4,5)} = \rho(S_1, S_{(4,5)}) = \frac{1}{2} \rho_{1,4} + \frac{1}{2} \rho_{1,5} - \frac{1}{2} |\rho_{1,4} - \rho_{1,5}| = \frac{1}{2} (10,05 + 11,05) - \frac{1}{2} |10,05 - 11,05| = 10,05$$

Исходные данные для классификации



Тема 4. Абсолютные и относительные величины

А. Абсолютные величины отражают физические размеры изучаемых статистикой процессов и явлений, а именно их массу, площадь, объем, протяженность, временные характеристики, а также могут представлять объем совокупности, т. е. число составляющих ее единиц. Абсолютные статистические показатели всегда являются именованными числами и могут выражаться в натуральных, стоимостных или трудовых единицах измерения.

Б. Относительные величины представляют собой результат деления абсолютного показателя на другой и выражают соотношение между **количественными характеристиками** социально-экономических процессов и явлений.

Относительные статистические величины бывают следующих видов:

1. динамики;
2. расчетного задания;
3. выполнения расчетного задания;
4. структуры;
5. координации;
6. интенсивности;
7. сравнения.

1. Относительная величина динамики

(ОВД) - отношение уровня исследуемого процесса или явления за данный период времени (по состоянию на данный момент времени) к уровню этого же процесса или явления в прошлом:

$$\text{ОВД} = \frac{\text{Текущий показатель}}{\text{Предшествующий или базисный показатель}}$$

Пример 1

Имеются следующие данные по предприятию:

(млн руб.)

Показатель	На 1 января 2003 г.	На 1 января 2004 г.
Производственный капитал,	43 750	46 600
в том числе		
оборотные средства,	5 680	5 130
из них:		
собственные средства	3 135	2 565
заемные средства	2 545	2 565

Рассчитаем относительные величины динамики:

а) по производственному капиталу:

$$\frac{46\,600}{43\,750} \cdot 100\% = 106,51\%;$$

б) по оборотным средствам:

$$\frac{5\,130}{5\,680} \cdot 100\% = 90,32\%;$$

в) по собственным средствам:

$$\frac{2\,565}{3\,135} \cdot 100\% = 81,82\%;$$

г) по заемным средствам:

$$\frac{2\,565}{2\,545} \cdot 100\% = 100,79\%.$$

Рассчитанные относительные величины динамики показывают, что на 1 января 2004 г. по сравнению с 1 января 2003 г. производственный капитал предприятия вырос на 6,51%; в то же время оборотные средства уменьшились на 9,68%, а собственные — на 18,18%; заемные средства увеличились на 0,79%.

2. Относительная величина расчетного задания (ОВР_з) -
отношение величины расчетного задания на период к достигнутой величине прошлого периода:

$$\text{ОВР}_{з} = \frac{\text{Показатель, рассчитываемый на } (i + 1) \text{ - й период}}{\text{Показатель, достигнутый в } i \text{ - ом периоде}}$$

3. Относительная величина выполнения расчетного задания (ОВР_{в.з}) - отношение величины, достигнутой в отчетном периоде, к величине расчетного задания:

$$\text{ОВР}_{вз} = \frac{\text{Показатель, достигнутый в } (i + 1) \text{ - й период}}{\text{Показатель, рассчитанный на } (i + 1) \text{ - й период}}$$

Пример 2

В 2002 г. оборот торговой фирмы составил 2500 млн руб. В 2003 г. фирма рассчитывала увеличить оборот до 3200 млн руб. Фактический оборот фирмы в 2003 г. составил 2800 млн руб.

Рассчитаем относительные величины расчетного задания, выполнения расчетного задания и динамики и определим их взаимосвязь:

$$\text{ОВР}_з = \frac{3200}{2500} = 1,28, \text{ или } 1,28 \cdot 100\% = 128\%;$$

$$\text{ОВР}_{вз} = \frac{2800}{3200} = 0,875, \text{ или } 0,875 \cdot 100\% = 87,5\%.$$

Между относительными величинами расчетного задания, выполнения расчетного задания и динамики существует следующая взаимосвязь:

$$\text{ОВР}_з \cdot \text{ОВР}_{вз} = \text{ОВД}.$$

В нашем примере:

$$1,28 \cdot 0,875 = 1,12, \text{ или } 1,12 \cdot 100\% = 112\%;$$

или

$$\text{ОВД} = \frac{2800}{2500} = 1,12, \text{ или } 1,12 \cdot 100\% = 112\%.$$

$$\text{ОВР}_з \cdot \text{ОВР}_{вз} = \text{ОВД}$$

4. Относительная величина структуры (ОВС) - соотношение структурных частей изучаемого объекта и их целого:

$$\text{ОВС} = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель по совокупности в целом}}$$

Пример 3

На основании данных примера 1
рассчитать относительные
величины структуры

Показатель	На 1 января 2003 г.	На 1 января 2004 г.
Производственный капитал, в том числе	43 750	46 600
оборотные средства,	5 680	5 130
из них:		
собственные средства	3 135	2 565
заемные средства	2 545	2 565

1. Доля оборотных средств в общей стоимости производственного капитала предприятия составит:

а) на 1 января 2003 г.:

$$\frac{5680}{43750} \cdot 100\% = 12,98\%;$$

б) на 1 января 2004 г.:

$$\frac{5130}{46600} \cdot 100\% = 11,0\%;$$

2. Доля собственных средств в общей стоимости оборотных средств составит:

а) на 1 января 2003 г.:

$$\frac{3135}{5680} \cdot 100\% = 55,19\%;$$

б) на 1 января 2004 г.:

$$\frac{2565}{5130} \cdot 100\% = 50,0\%.$$

Из полученных показателей видно, что на предприятии доля оборотных средств в общей стоимости производственного капитала снизилась на 1,98% и на 1 января 2004 г. составила 11%.

Доля собственных средств в общей стоимости оборотного капитала на 1 января 2003 г. составила 55,19%, а на 1 января 2004 г. снизилась и составила 50%, т. е. только на 50% необходимые предприятию запасы и затраты обеспечиваются собственными средствами.

5. Относительная величина координации (ОВК) - отношение одной части совокупности к другой части этой же совокупности:

$$\text{ОВК} = \frac{\text{Показатель, характеризующий } i \text{ - ю часть совокупности}}{\text{Показатель, характеризующий } i \text{ - ю часть совокупности, выбранную в качестве базы сравнения}}$$

Пример 4

На основании данных примера 1
рассчитать относительные
величины координации

Показатель	На 1 января 2003 г.	На 1 января 2004 г.
Производственный капитал,	43 750	46 600
в том числе		
оборотные средства,	5 680	5 130
из них:		
собственные средства	3 135	2 565
заемные средства	2 545	2 565

1. Соотношение заемных и собственных средств предприятия составит:

а) на 1 января 2003 г.:

$$\frac{2545}{3135} = 0,8118;$$

б) на 1 января 2004 г.:

$$\frac{2565}{2565} = 1.$$

Этот показатель свидетельствует о том, что на каждый рубль собственных средств предприятия на 1 января 2003 г. приходилось 81,18 коп. заемных средств, а на 1 января 2004 г. — 1 руб.

2. Соотношение собственных и заемных средств составит:

а) на 1 января 2003 г.:

$$\frac{3135}{2545} = 1,2318;$$

б) на 1 января 2004 г.:

$$\frac{2565}{2565} = 1.$$

Этот показатель характеризует финансовую устойчивость или платежеспособность предприятия, т. е. его способность рассчитываться по своим обязательствам.

6. Относительная величина интенсивности (ОВИ) характеризует степень распространения изучаемого процесса или явления и представляет собой отношение исследуемого показателя к показателю присущей ему среды:

$$\text{ОВИ} = \frac{\text{Показатель, характеризующий явление } A}{\text{Показатель, характеризующий среду распространения явления } A}$$

Разновидностью относительной величины интенсивности является относительная величина уровня экономического развития, характеризующая производство продукции в расчете на душу населения и играющая важную роль в оценке развития экономики государства.

Пример 5

Рассчитаем относительный показатель уровня экономического развития на основании следующих данных: в 2002 г. валовой внутренний продукт Российской Федерации составил 10 863,3 млрд руб., среднегодовая численность населения — 145,18 млн человек. Отсюда искомый показатель равен:

$$\frac{10\,863,3 \text{ млрд руб.}}{145,18 \text{ млн человек}} = 74\,826(\text{руб.})$$

Следовательно, в 2002 г. ВВП на душу населения составил 74 826 руб.

7. Относительная величина сравнения (ОВС_р) -
соотношение одного и того же абсолютного
показателя, характеризующего разные объекты:

$$\text{ОВС}_p = \frac{\text{Показатель, характеризующий объект } A}{\text{Показатель, характеризующий объект } B}$$

Пример 6

Рассчитаем относительную величину сравнения на основании следующих данных. В 2002 г. инвестиции в экономику Российской Федерации составили:

Кипра — 2,327 млн дол. США;

США — 1,133 млн дол. США.

$$\frac{2,327}{1,133} = 2,0538, \text{ или } 205,38\%.$$

Инвестиции Кипра в экономику Российской Федерации в 2 раза превышают инвестиции США.

Тема 5. Анализ рядов распределения

- 5.1. Средние величины
- 5.2. Анализ вариации
- 5.3. Анализ формы кривой
распределения

Основные стадии анализа рядов распределения:

1. Оценка массового уровня признака с помощью расчета средних показателей.
2. Оценка колеблемости признака.
3. Оценка формы кривой распределения.

5.1. Средние величины

- 5.1.1. Принципы применения средних величин
- 5.1.2. Классификация средних величин
- 5.1.3. Степенные средние
- 5.1.4. Структурные средние

5.1.1. Принципы применения средних величин

1. При определении средней в каждом конкретном случае нужно исходить из качественного содержания осредняемого признака, учитывать взаимосвязь изучаемых признаков, а также имеющиеся для расчета данные.
2. Средняя величина должна прежде всего рассчитываться по однородной совокупности.
3. Общие средние должны подкрепляться групповыми средними.
4. Необходим обоснованный выбор единицы совокупности, для которой рассчитывается средняя.

5.1.2. Классификация средних величин

Виды средних величин

Степенные средние

Структурные средние

Простые

(для
несгруппированных
данных)

Взвешенные

(для
сгруппированных
данных)

Мода

Медиана (1/2)
Квартиль (1/4)
Дециль (1/10)
*Перцентиль
(1/100)*

5.1.3. Степенные средние

Простая средняя
считается по
несгруппированным
данным:

$$\bar{x} = \sqrt[z]{\frac{\sum_{i=1}^n x_i^z}{n}}$$

Взвешенная средняя
считается по
сгруппированным
данным:

$$\bar{x} = \sqrt[z]{\frac{\sum_{i=1}^k x_i^z \cdot m_i}{\sum_{i=1}^k m_i}}$$

← =n

Вид степенной средней	Показатель степени, z	Формула расчета	
		Простая	Взвешенная
Гармоническая	-1	$\bar{x} = \frac{n}{\sum_{i=1}^n 1/x_i}$	$\bar{x} = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n w_i/x_i}, \text{ где } w_i = m_i \cdot x_i$
Геометрическая	0	$\bar{x} = \sqrt[n]{\prod_{i=1}^n x_i}$	$\bar{x} = \sqrt[\sum_{i=1}^k m_i]{\prod_{i=1}^k (x_i)^{m_i}}$
Арифметическая	1	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i}$
Квадратическая	2	$\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum_{i=1}^k x_i^2 \cdot m_i}{\sum_{i=1}^k m_i}}$

Вид степенной средней	Показатель степени, z	Формула расчета	
		Простая	Взвешенная
Кубическая	3	$\bar{x} = \sqrt[3]{\frac{\sum_{i=1}^n x_i^3}{n}}$	$\bar{x} = \sqrt[3]{\frac{\sum_{i=1}^k x_i^3 \cdot m_i}{\sum_{i=1}^k m_i}}$

Если рассчитать все виды средних для одних и тех же исходных данных, то значения их окажутся не одинаковыми. Здесь действует **правило мажорантности средних**:

с увеличением показателя степени z увеличивается и соответствующая средняя величина:

$$\bar{x}_{\text{гарм}} \leq \bar{x}_{\text{геом}} \leq \bar{x}_{\text{ариф}} \leq \bar{x}_{\text{квадр}} \leq \bar{x}_{\text{куб}}$$

В отдельных случаях веса могут быть представлены не абсолютными величинами, а относительными (в % или долях единицы). Тогда используют формулу средней:

*Доля группы в
совокупности*

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i} = \sum_{i=1}^k \left(x_i \cdot \frac{m_i}{\sum_{i=1}^k m_i} \right) = \sum_{i=1}^k (x_i \cdot d_i).$$

В интервальном вариационном ряду для расчета средней арифметической взвешенной определяются и используются значения середины интервалов.

Пример 1

N предприят ия	Январь		Февраль	
	Средняя зароботная плата, руб.	Численность работников, чел.	Средняя зароботная плата, руб.	Фонд оплаты труда, тыс. руб.
1	4900	450	5700	2565
2	5400	600	5800	3480
Итого	5186	1050	5757	6045

Средняя арифметическая
взвешенная

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i}$$

Средняя
гармоническая
взвешенная

$$\bar{x} = \frac{\sum_{i=1}^n w_i}{\sum_{i=1}^n w_i / x_i}, \text{ где } w_i = m_i \cdot x_i$$

Пример 2

Разряд, x	Число рабочих, t	xt			
1	5	5			
2	24	48			
3	48	144			
4	27	108			
5	16	80			
Итого	120	385			

Средний разряд по сгруппированным данным

средняя арифметическая	3,2
------------------------	-----

Средний разряд по несгруппированным данным

средняя гармоническая	2,8	=CPГAPM(\$C\$2:\$C\$121)
-----------------------	-----	--------------------------

средняя геометрическая	3,0	=CPГEOM(\$C\$2:\$C\$121)
------------------------	-----	--------------------------

средняя арифметическая	3,2	=CP3HA Ч(\$C\$2:\$C\$121)
------------------------	-----	---------------------------



Пример 3

Интервалы по x		Число рабочих, m	$x_{ц}$	$x_{ц}m$
Начало интервала	Конец интервала			
90	100	4	95	380
100	110	13	105	1365
110	120	21	115	2415
120	130	36	125	4500
130	140	28	135	3780
140	150	11	145	1595
150	160	3	155	465
160	170	4	165	660
Итого		120		15160

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i}$$

Средний процент выполнения норм по сгруппированным данным

126,33

Средний процент выполнения норм по несгруппированным данным

126,37

Пример 4

Финансирование инвестиций за счет собственных средств по предприятиям АО за отчетный период характеризуются следующими данными.

Необходимо определить средний удельный вес собственных средств в общем объеме инвестиций по АО.

Осредняемый признак

Номер предприятия	Общая сумма инвестиций, млн руб.	Удельный вес собственных средств в общем объеме инвестиций, %		
1	15,5	69,3	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i}$	
2	40	57,4		
3	32	62,8		
Итого	87,5			
Вес группы	—			
	x	=	61,5	%

Пример 5

Портфель инвестора состоит из акций трех компаний. Их доходность равна соответственно 15, 18 и 20%, а доля в портфеле – 20, 45 и 35%. Какова средняя доходность портфеля?

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot m_i}{\sum_{i=1}^k m_i} = \sum_{i=1}^k \left(x_i \cdot \frac{m_i}{\sum_{i=1}^k m_i} \right) = \sum_{i=1}^k (x_i \cdot d_i) = 15 \cdot 0,2 + 18 \cdot 0,45 + 20 \cdot 0,35 = 0,181 \text{ (18,1\%).}$$

Пример 6

Номер предприятия	Величина активов, млн руб.	Рентабельность активов, %	Прибыль, млн руб.
1	5000	8	400
2	8000	9,5	760
Итого	13000	8,9	1160

Пример 7

Определить среднеквартальный коэффициент роста выпуска продукции.

Показатель	Условное обозначение	Квартал года					
		1	2	3	4		
Выпуск продукции, млн руб.	$y_i, i=0\div 3$	100	120	125	122		
Коэффициент роста выпуска продукции (цепной)	$k_i=y_i/y_{i-1}$		1,20	1,04	0,98	$\bar{x} = \sqrt[n]{\prod_{i=1}^n x_i}$	

$$k_{3/0} = \frac{y_3}{y_0} = k_1 \cdot k_2 \cdot k_3 = \bar{k} \cdot \bar{k} \cdot \bar{k} = \bar{k}^3$$

$$\bar{k} = \sqrt[3]{k_1 \cdot k_2 \cdot k_3} = \sqrt[3]{1,2 \cdot 1,04 \cdot 0,976} = \sqrt[3]{\frac{122}{100}} = 1,068$$

Простая геометрическая средняя

Пример 8

Определить среднегодовой темп роста душевого дохода в США в рассматриваемом периоде.

Годы			Среднегодовые темпы роста душевого дохода в США, %
Начало периода	Окончание периода	Длина периода	
1913	1950	37	101,6
1950	1973	23	102,4
1973	1992	19	101,4
Итого		79	101,8

$$\bar{k} = \sqrt[37+23+19]{1,016^{37} \cdot 1,024^{23} \cdot 1,014^{19}} = 101,8\%.$$

$$\bar{x} = \sqrt[\sum_{i=1}^k m_i]{\prod_{i=1}^k (x_i)^{m_i}}$$

Взвешенная геометрическая средняя

5.1.4. Структурные средние

1. Мода (M_o) - наиболее часто встречающееся значение признака в совокупности.

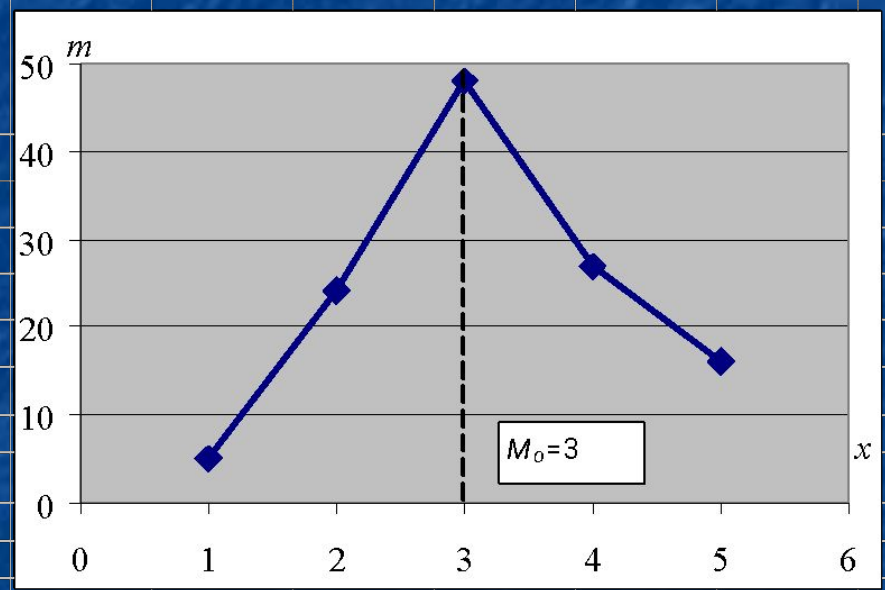
Несгруппированные данные - рост студентов (см)						
165	162	167	154	170	175	162

Для *дискретного ряда* мода определяется по таблице или графику.

Разряд, x	Число рабочих, m
1	5
2	24
3	48
4	27
5	16
Итого	120

max

$M_0=3$



Для *интервального ряда* мода определяется в два этапа:

- определяется модальный интервал;
- уточняется значение моды внутри модального интервала по формуле:

$$M_0 = x_0 + i \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2},$$

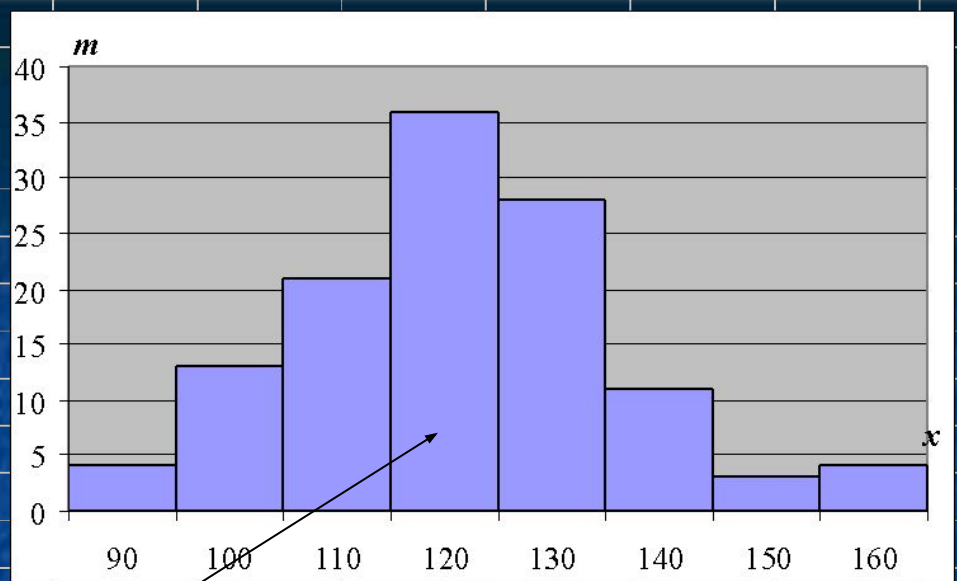
x_0 – начало модального интервала; i – длина интервала;

Δ_1 – разность частот модальной и домодальной;

Δ_2 – разность частот модальной и замодальной.

Формула используется и для рядов с неравными интервалами, только в этом случае необходимо сначала рассчитать плотность распределения, затем по плотности определить модальный интервал и провести расчет по формуле.

Интервалы по x		Число рабочих, m
Начало интервала	Конец интервала	
90	100	4
100	110	13
110	120	21
120	130	36
130	140	28
140	150	11
150	160	3
160	170	4
Итого		120



Модальный интервал

$M_0 = 126,522$

$$M_0 = x_0 + i \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2} = 120 + 10 \cdot \frac{36 - 21}{(36 - 21) + (36 - 28)} = 126,52\%$$

2. Медиана (Me) – значение признака в середине ранжированного ряда.

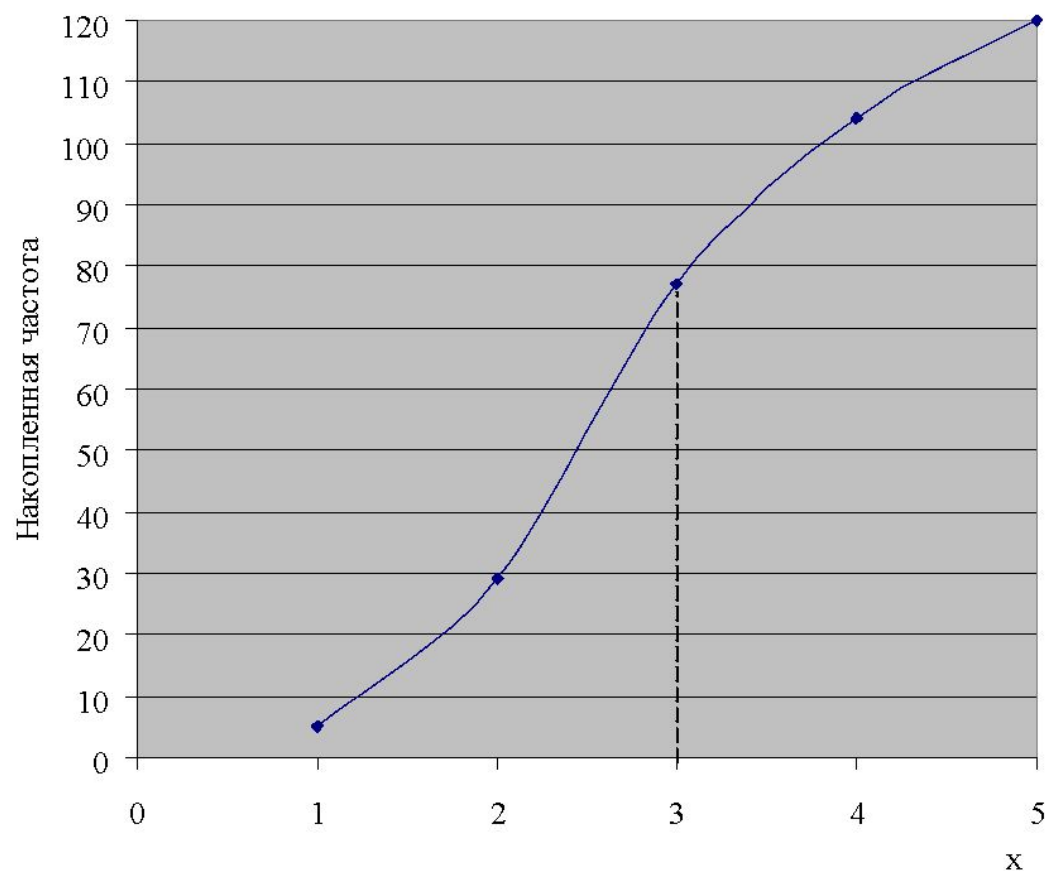
Определение медианы для *несгруппированных данных*

I - <i>нечетное</i> количество данных									
Несгруппированные данные - рост студентов (см)									
165	162	167	154	170	175	162		$M_0=$	162 см
Ранжированные данные									
154	162	162	165	167	170	175		$X_{ср. ариф.} =$	165 см
			Me					$R=$	21 см
II - <i>четное</i> количество данных									
Несгруппированные данные - рост студентов (см)									
165	162	167	154	170	175	162	210	$M_0=$	162 см
Ранжированные данные									
154	162	162	165	167	170	175	210	$X_{ср. ариф.} =$	170,6 см
			$Me = (165+167)/2 = 166$					$R=$	56 см

Для *дискретного* ряда медиана определяют путем накапливания частот с начала ряда до варианта, в котором сумма накопленных частот будет равна или превысит половину объема статистической совокупности.

x	m	Σm	
1	5	5	
2	24	29	
3	48	77	$77 > 120/2$
4	27	104	
5	16	120	
Итого	120		

$Me=3$



Для *интервального ряда* медиана определяется в два этапа:

- определяется медианный интервал;
- уточняется значение медианы внутри медианного интервала по формуле:

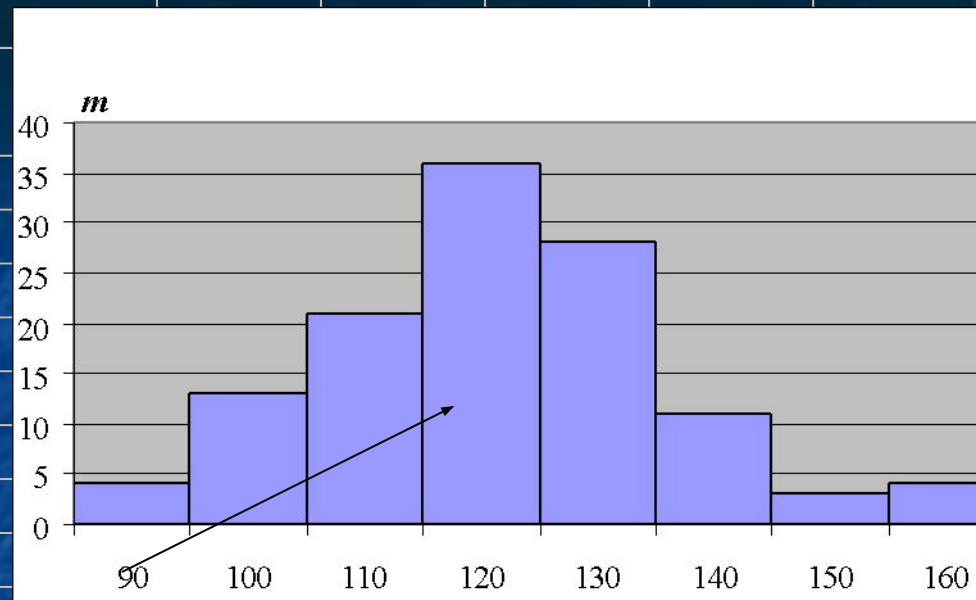
$$M_e = x_e + i \cdot \frac{\frac{n}{2} - \sum m_{e-1}}{m_e},$$

- x_e – начало медианного интервала;
- i – длина интервала;
- $\sum m_{e-1}$ – накопленная частота интервала, предшествующего медианному;
- m_e – частота медианного интервала.

Формула используется и для рядов с неравными интервалами.

Интервалы по x		m	Σm
Начало интервала	Конец интервала		
90	100	4	4
100	110	13	17
110	120	21	38
120	130	36	74
130	140	28	102
140	150	11	113
150	160	3	116
160	170	4	120
Итого		120	

$74 > 120/2$



Медианный интервал

$Me = 126,11$

$$M_e = x_e + i \cdot \frac{\frac{n}{2} - \sum m_{e-1}}{m_e} = 120 + 10 \cdot \frac{\frac{120}{2} - 38}{36} = 126,11.$$

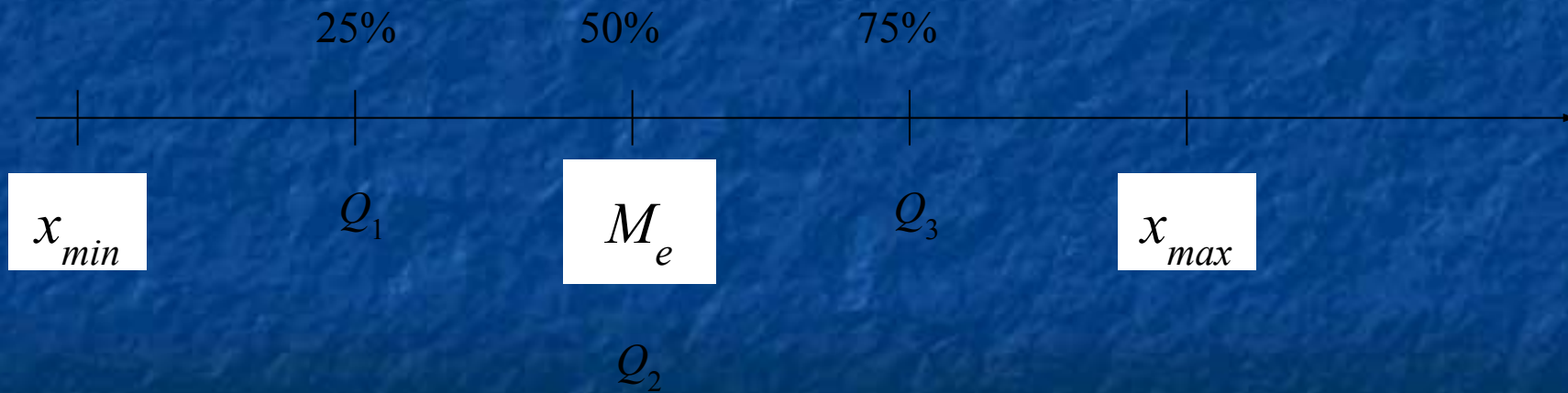
Определение моды и медианы для вариационного ряда с неравными интервалами

Интервалы по x		Число рабочих, m	Относительная частота, f	Длина интервала в условных единицах, i	Плотность в условных единицах $\rho = m/i$	$\Sigma \rho$
Начало интервала	Конец интервала					
0	1	16	13%	0,5	32	32
1	3	20	17%	1	20	52
3	5	26	22%	1	26	78
5	8	35	29%	1,5	23,3	101,3
8	12	11	9%	2	5,5	106,8
12	20	8	7%	4	2	108,8
20	40	4	3%	10	0,4	109,2
Итого		120	100%		109,2	
Условная единица - длина второго интервала =				2	Мо=	0,73
					Ме=	3,62

$$M_0 = x_0 + i \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2} = 0 + 1 \cdot \frac{32 - 0}{(32 - 0) + (32 - 20)} = 0,73 \text{ год.}$$

$$M_e = x_e + i \cdot \frac{\frac{n}{2} - \sum_{m_{e-1}}}{m_e} = 3 + 2 \cdot \frac{60 - 52}{26} = 3,62 \text{ год.}$$

Медиана	—	2 части
Квартиль	—	4 части
Дециль	—	10 частей
Перцентиль	—	100 частей



5.2. Анализ вариации

Колеблемость отдельных значений характеризуют показатели вариации.

Термин «вариация» произошел от латинского *variatio* - изменение, колеблемость, различие.

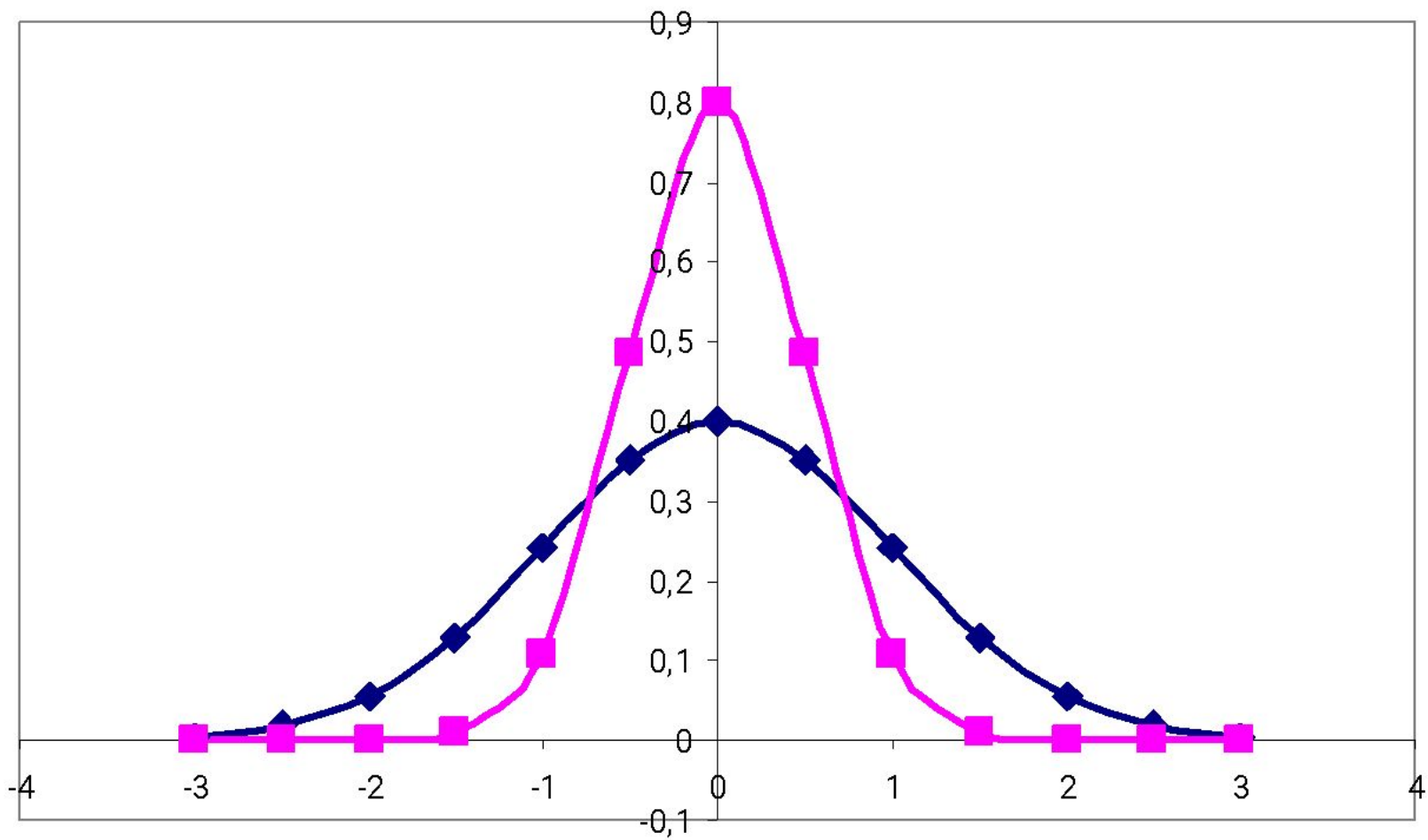
Однако не всякие различия принято называть вариацией.

Под **вариацией** в статистике понимают такие количественные изменения величины исследуемого признака в пределах однородной совокупности, которые обусловлены перекрещивающимся влиянием различных факторов.

Различают *случайную и систематическую вариации* признака.

Анализ систематической вариации позволяет оценить степень зависимости изменений в изучаемом признаке от определяющих его факторов.

Вариация измеряется рядом абсолютных, средних и относительных показателей.



А. Абсолютные, средние
показатели

Б. Относительные
показатели

1. Размах вариации

1. Коэффициент осцилляции

2. Среднее линейное
отклонение

2. Относительное линейное
отклонение

3. Дисперсия

3. Коэффициент вариации

4. Среднеквадратическое
отклонение

A. Абсолютные, средние показатели вариации

показатель	простая форма	взвешенная форма
Размах вариации	$R = x_{\max} - x_{\min}$	
Среднее линейное отклонение	$\bar{d} = \frac{\sum_{i=1}^n x_i - \bar{x} }{n}$	$\bar{d} = \frac{\sum_{i=1}^k x_i - \bar{x} \cdot m_i}{\sum_{i=1}^k m_i}$
Дисперсия	$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$	$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 \cdot m_i}{\sum_{i=1}^k m_i}$
Среднеквадратическое отклонение	$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$	$\sigma = \sqrt{\frac{\sum_{i=1}^k (x_i - \bar{x})^2 \cdot m_i}{\sum_{i=1}^k m_i}}$

Простейшие свойства σ

1. Если все частоты умножить или разделить на какое-либо постоянное число, то σ не измениться.
2. Если ко всем вариантам признака прибавить или отнять какое-либо постоянное число, σ не измениться.
3. Если все варианты признака умножить или разделить на какое-то постоянное число, то σ умножится или разделится на модуль этого числа.
4. Дисперсия равна осредненному квадрату минус квадрат самой средней:

$$\sigma^2 = \frac{\sum_{i=1}^k x_i^2 \cdot m_i}{\sum_{i=1}^k m_i} - (\bar{x})^2.$$

Б. Относительные показатели

Коэффициент осцилляции	$V_R = \frac{R}{x} \cdot 100\%$
Относительное линейное отклонение	$V_{\bar{d}} = \frac{\bar{d}}{x} \cdot 100\%$
Коэффициент вариации	$V_{\sigma} = \frac{\sigma}{x} \cdot 100\%$

Чем меньше V_{σ} и σ , тем надежнее среднее.

Условная граница однородности $V_{\sigma} = 33\%$.

Пример 1

Имеются следующие данные об объемах товарооборота предприятий:

Группы предприятий по объему товарооборота, млн руб. (x_i)	Число предприятий, шт. (f_i)
90–100	28
100–110	48
110–120	20
120–130	4
<i>Итого</i>	<i>100</i>

Определяем показатель размаха вариации:

$$R = 130 - 90 = 40 \text{ (млн руб.)}$$

Пример 2

Имеются следующие данные о производительности рабочих:

Табельный номер рабочего	x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $
1	2	-8	8
2	3	-7	7
3	12	2	2
4	15	5	5
5	18	8	8
Итого	50	0	30

Рассчитаем среднее линейное отклонение:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{30}{5} = 6.$$

$$\bar{x} = 50/5 = 10$$

Пример 3

x_i	m_i	xm_i	$ x_i - x_{cp} \cdot m_i$	$(x_i - x_{cp})^2 \cdot m_i$
1	5	5	11,042	24,384
2	24	48	29,000	35,042
3	48	144	10,000	2,083
4	27	108	21,375	16,922
5	16	80	28,667	51,361
Итого	120	385	100,083	129,792
	$x_{cp} =$	3,208		
	$R =$	4	$v_{\sigma} =$	32%
	$d_{cp} =$	0,834	$v_d =$	26%
	$\sigma^2 =$	1,0816	$v_R =$	125%
	$\sigma =$	1,04		

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

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

$$V_R = \frac{R}{x} \cdot 100\%$$

$$\bar{d} = \frac{\sum_{i=1}^k |x_i - \bar{x}| \cdot m_i}{\sum_{i=1}^k m_i}$$

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 \cdot m_i}{\sum_{i=1}^k m_i}$$

Правило сложения дисперсий

Если некоторая совокупность единиц делится на группы, то наряду с общей дисперсией могут быть также найдены дисперсии для каждой отдельной группы – *групповые (внутригрупповые) дисперсии* σ_i^2 , а также их средняя величина:

$$\overline{\sigma^2} = \frac{\sum_{i=1}^k \sigma_i^2 \cdot m_i}{\sum_{i=1}^k m_i}.$$

Кроме того, может быть вычислена также *межгрупповая дисперсия* δ^2 , характеризующая колеблемость групповых средних около общей средней \bar{x} :

$$\delta^2 = \frac{\sum_{i=1}^k (\bar{x}_i - \bar{x})^2 \cdot m_i}{\sum_{i=1}^k m_i}$$

В математической статистике доказывается, что *общая дисперсия* равна:

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

Средняя из групповых дисперсий

Межгрупповая дисперсия

Пример 4

1 группа - рабочие 2-го разряда			2 группа - рабочие 3-го разряда			
n/n	Дневная выработка (м ³) x ₁	x ₁ ²	n/n	Дневная выработка (м ³) x ₂	x ₂ ²	
	1	3,2	10,24	1	3,9	15,21
	2	3,5	12,25	2	4,2	17,64
	3	4,5	20,25	3	4,8	23,04
	4	4,8	23,04	4	5,1	26,01
				5	5,4	29,16
				6	6,6	43,56
Итого по группе	16	65,78	Итого по группе	30	154,62	

$$\bar{x}_1 = \frac{16}{4} = 4; \bar{x}_2 = \frac{30}{6} = 5; \bar{x} = \frac{16+30}{10} = 4,6;$$

$$\sigma_1^2 = \bar{x}_1^2 - \bar{x}_1^2 = \frac{65,78}{4} - 4^2 = 0,445; \sigma_2^2 = \frac{154,62}{6} - 5^2 = 0,77.$$

Общая дисперсия составит:

$$\sigma^2 = \frac{65,78+154,62}{10} - (4,6)^2 = 0,88.$$

Средняя из групповых равна:

$$\bar{\sigma}^2 = \frac{0,445 \times 4 + 0,77 \times 6}{10} = 0,64.$$

Следовательно, межгрупповая дисперсия должна составить:

$$\delta^2 = \sigma^2 - \bar{\sigma}^2 = 0,88 - 0,64 = 0,24$$

или

$$\delta^2 = \frac{(4-4,6)^2 \times 4 + (5-4,6)^2 \times 6}{10} = 0,24.$$

5.3. Анализ формы кривой распределения

При построения кривой распределения выделяют два подхода:

1. *Прямой*, который заключается в постепенном уменьшении величины интервалов и одновременном, но не столь быстром, увеличении числа наблюдений.
2. *Косвенный*, который заключается в укрупнении интервалов и математическом выравнивании.

Математическое выравнивание сводится к отысканию кривой распределения, которая отражает закономерность изменения плотности в чистом виде.

Математическое выравнивание состоит из трех этапов:

- 1) выбор и обоснование вида кривой;
- 2) выравнивание эмпирических рядов по выбранной кривой;
- 3) оценка соответствия выбранного типа кривой эмпирическому ряду.

Нормальное распределение

Условие его возникновения:

Если интересующий нас признак формируется под влиянием суммарного действия взаимонезависимых факторов, ни один из которых не имеет преобладающего влияния по сравнению с остальными и не отличается исключительно большой дисперсией, то при большом числе таких факторов закон распределения признака становится близким к нормальному.

Нормальное распределение Гаусса-Ляпунова выражается следующей формулой:

$$p_x = \frac{1}{\sqrt{2\pi\sigma}} \cdot e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2}$$

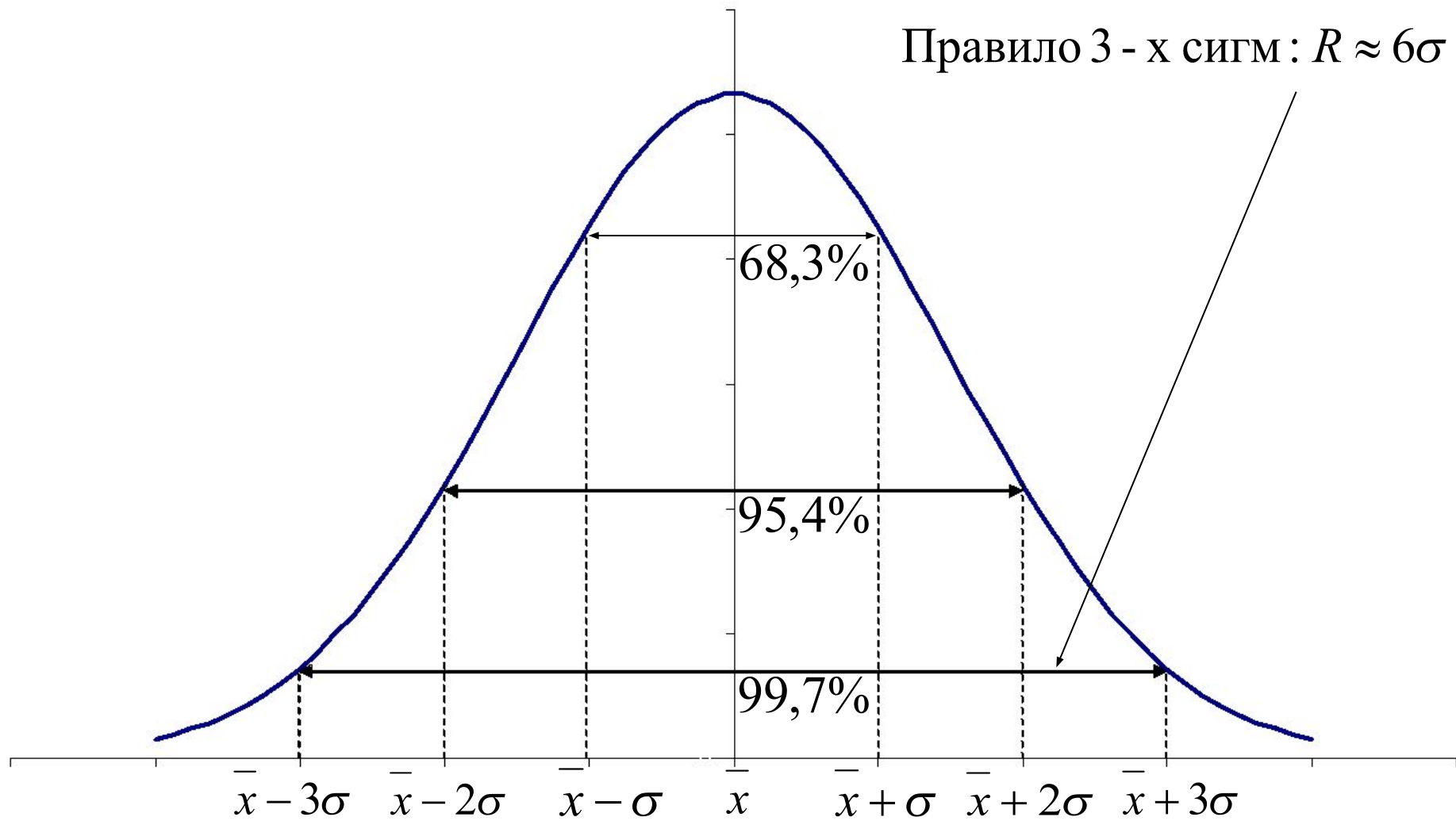
Два основных параметра нормального распределения :

- 1) \bar{x} ;
- 2) σ .

Заменяя через $t = \frac{x - \bar{x}}{\sigma}$, переходим к стандартной системе единиц, где начало отсчета \bar{x} , а единица измерения σ .

$$p_t = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2}$$

Правило 3 - х сигм : $R \approx 6\sigma$



Условие симметрии : $\bar{x} = M_0 = M_e$

Выравнивание эмпирического ряда распределения по нормальной кривой осуществляется в следующей последовательности.

1. Определить центры интервалов в новой системе координат по формуле:

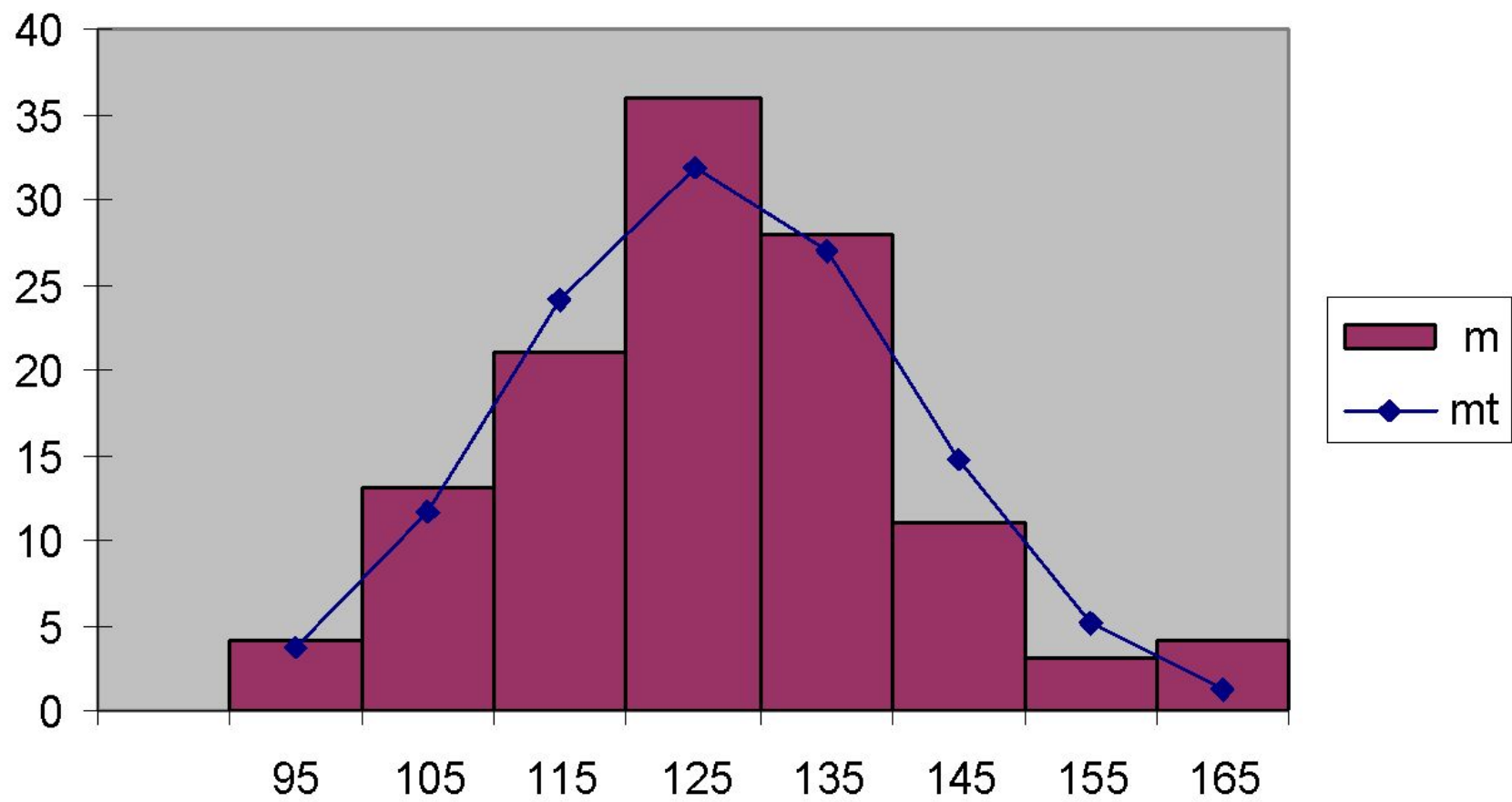
$$t_{\text{ц}} = \frac{x_{\text{ц}} - \bar{X}}{\sigma}$$

2. Найти ординаты нормальной кривой распределения, представляющие относительные плотности распределения в расчете на единицу измерения, по формуле:

$$P_{t_{\text{ц}}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t_{\text{ц}}^2}{2}}$$

3. Для каждого значения $t_{\text{ц}}$ рассчитать значения теоретической частоты (плотности распределения в расчете на интервал):

$$m_t = P_{t_{\text{ц}}} \frac{n \cdot i}{\sigma}$$



Критерий Колмогорова

Несовпадения между эмпирическим и теоретическим (по нормальному закону) рядами распределения обусловлены двумя причинами:

- 1) расхождения чисто случайные;
- 2) несоответствие изучаемого распределения нормальному по своей природе.

Критерий Колмогорова (критерий λ) основан на сопоставлении сумм накопленных эмпирических и теоретических частот и определяется по формуле:

$$\lambda = \frac{|D|_{\max}}{\sqrt{n}},$$

где $|D|_{\max}$ - максимальное значение разности между накопленными эмпирическими и теоретическими частотами,
 n – сумма эмпирических частот.

λ	0,3	2,31
$P(\lambda)$	1	0

Асимметрия распределения

Два варианта оценки степени асимметрии:

1. Коэффициент асимметрии

$$k_a = \frac{\bar{x} - M_0}{\sigma}$$

При левосторонней (отрицательной) асимметрии:

$$M_0 > M_e > \bar{x}; \quad k_a < 0$$

При правосторонней (положительной) асимметрии:

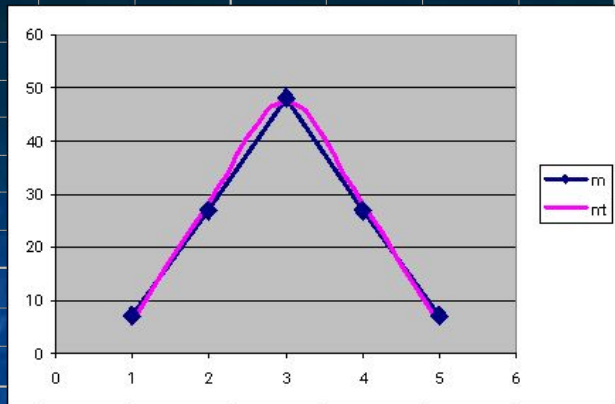
$$M_0 < M_e < \bar{x}; \quad k_a > 0$$

Симметрия

x	m	xm	Σm	$(x-x_{cp})^2 m$	p_x	m_t
1	7	7	7	28	0,0497	6
2	27	54	34	27	0,2418	28
3	48	144	82	0	0,4097	48
4	27	108	109	27	0,2418	28
5	7	35	116	28	0,0497	6
Итого	116	348		110	0,9927	115

$x_{cp} = 3$
 $Me = 3$
 $M_0 = 3$
 $\sigma = 0,9738$

$M_0 = Me = \bar{x}; \quad k_a = 0$
 $k_a = 0$

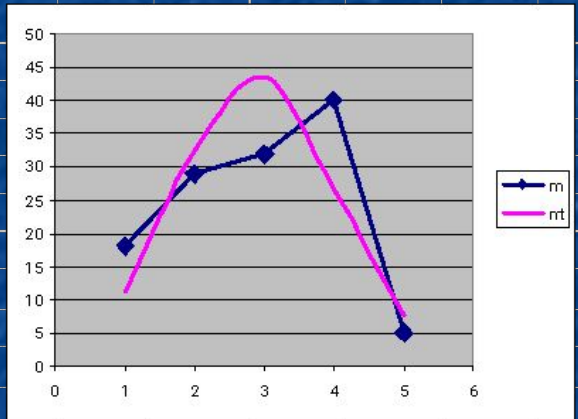


Левосторонняя асимметрия

x	m	xm	Σm	$(x-x_{cp})^2 m$	p_x	m_t
1	18	18	18	63,55372	0,089	11
2	29	58	47	22,408234	0,2606	32
3	32	96	79	0,4682622	0,3501	43
4	40	160	119	50,262747	0,2158	27
5	5	25	124	22,492521	0,0611	8
Итого	124	357		159,18548	0,9766	121

$x_{cp} = 2,879$
 $Me = 3$
 $M_0 = 4$
 $\sigma = 1,133$

$M_0 > Me > \bar{x}; \quad k_a < 0$
 $k_a = -0,989$



Правосторонняя асимметрия

x	m	xm	Σm	$(x-x_{cp})^2 m$	p_x	m_t
1	5	5	5	23,62949	0,0603	7
2	35	70	40	48,232514	0,2061	24
3	30	90	70	0,9073724	0,3382	39
4	25	100	95	17,060491	0,2661	31
5	20	100	115	66,691871	0,1005	12
Итого	115	365		156,52174	0,9711	112

$x_{cp} = 3,1739$
 $Me = 3$
 $M_0 = 2$
 $\sigma = 1,1666$

$M_0 < Me < \bar{x}; \quad k_a > 0$
 $k_a = 1,0062$

$$2. \quad k_a = \frac{M_3}{\sigma^3}, \quad M_3 = \frac{\sum_{i=1}^k (x_i - \bar{x})^3 m_i}{\sum_{i=1}^k m_i}.$$

СКОС

[См. также](#)

Возвращает асимметрию распределения. Асимметрия характеризует степень несимметричности распределения относительно его среднего. Положительная асимметрия указывает на отклонение распределения в сторону положительных значений. Отрицательная асимметрия указывает на отклонение распределения в сторону отрицательных значений.

Синтаксис

СКОС(число1;число2; ...)

Число1, число2, ... — от 1 до 30 аргументов, для которых вычисляется асимметрия. Можно использовать один массив или одну ссылку на массив вместо аргументов, разделяемых точкой с запятой.

Заметки

- Аргументы должны быть числами или именами, массивами или ссылками, содержащими числа.
- Если аргумент, который является массивом или ссылкой, содержит текст, логические значения или пустые ячейки, эти значения игнорируются; ячейки, содержащие нулевые значения, учитываются.
- Если имеется менее трех точек данных, или стандартное отклонение равно нулю, то функция СКОС возвращает значение ошибки #ДЕЛ/0!.
- Уравнение для асимметрии определяется следующим образом:

$$\frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s} \right)^3$$

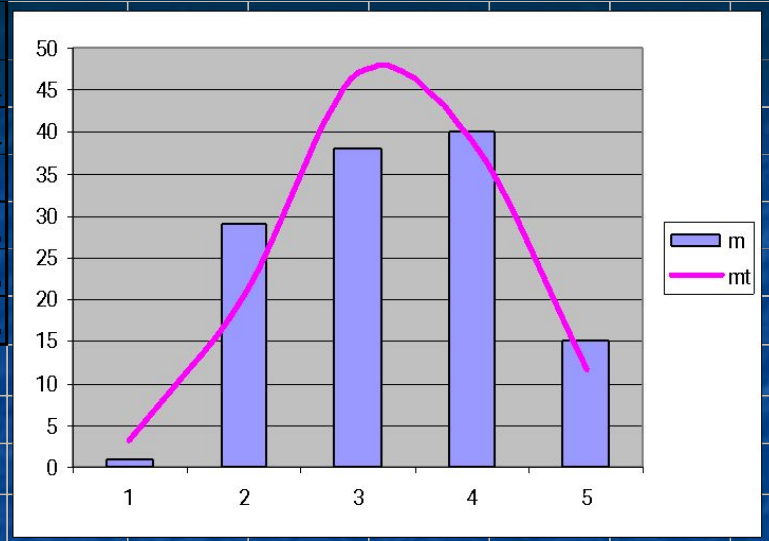
Пример

x	m	xm	Σm	$(x-x_{cp})^2 m$	p_x	m_t	$(x-x_{cp})^3 m$
1	1	1	1	5,3688281	0,0261	3	-12,44
2	29	58	30	50,30577	0,1664	20	-66,2564
3	38	114	68	3,820345	0,3827	47	-1,21133
4	40	160	108	18,655562	0,3176	39	12,74038
5	15	75	123	42,483641	0,0951	12	71,49686
Итого	123	408		120,63415	0,9878	122	4,329566
	$x_{cp} =$	3,3171					
	$M_e =$	3					
	$M_o =$	4					
	$\sigma =$	0,9903					

$$M_o > M_e > \bar{x}; \quad k_a < 0$$

$$1) k_a = -0,69$$

$$2) k_a = 0,0362$$



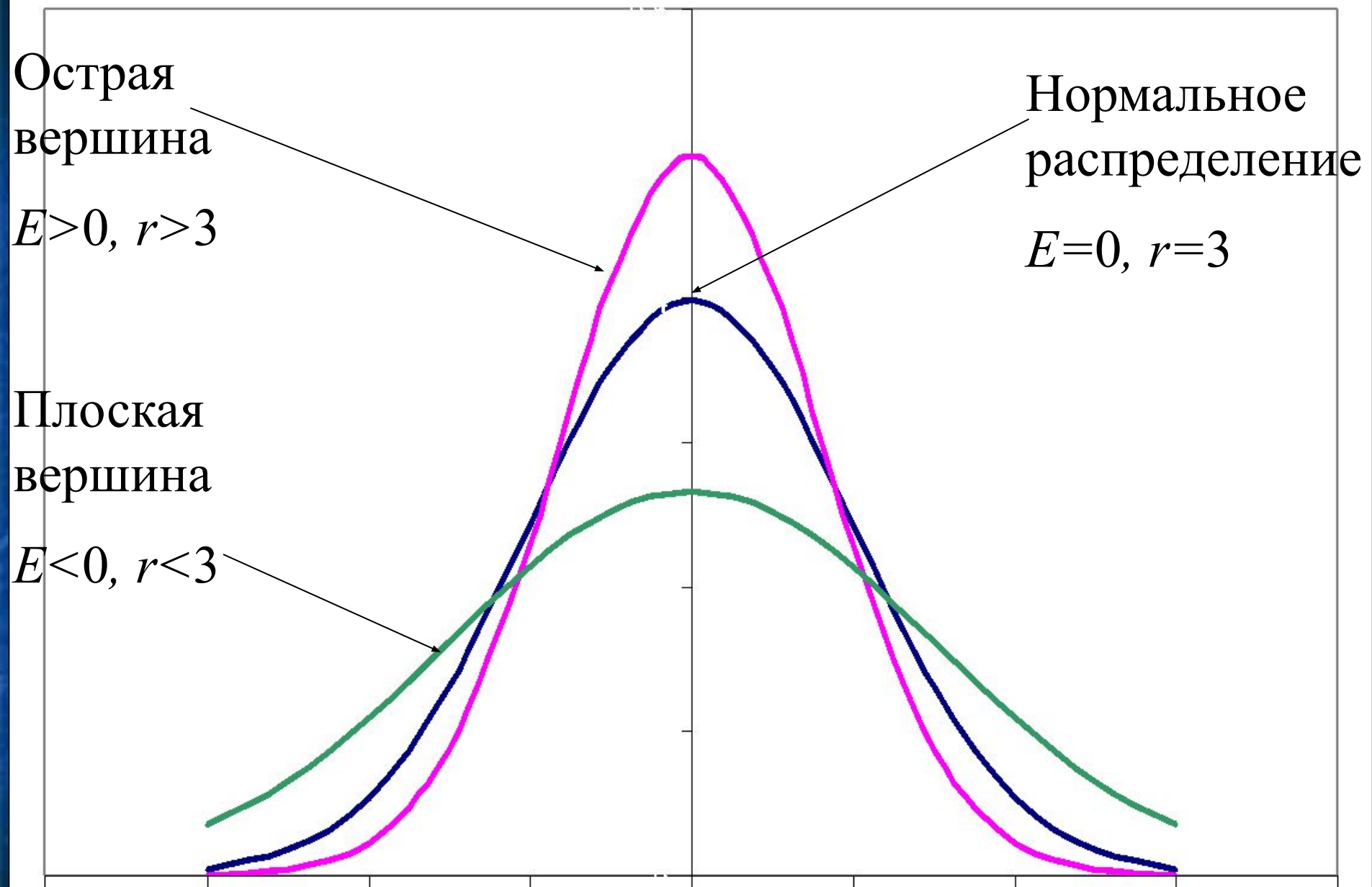
Эксцесс

Используется для характеристики островершинности кривой распределения.

Определяется по формуле:

$$E = \frac{M_4}{\sigma^4} - 3, \text{ где } M_4 = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 m_i}{\sum_{i=1}^k m_i}.$$

$$r = \frac{M_4}{\sigma^4} - \text{мера крутости.}$$



Интервалы по x		$x_{ц}$	m	$x_{ц}m$	$(x_{ц} - x_{cp})^2 m$	$(x_{ц} - x_{cp})^4 m$
Начало интервала	Конец интервала					
90	100	95	4	380	3927,11	3855550,42
100	110	105	13	1365	5916,44	2692639,605
110	120	115	21	2415	2697,33	346457,4815
120	130	125	36	4500	64,00	113,7777778
130	140	135	28	3780	2103,11	157967,0123
140	150	145	11	1595	3832,89	1335548,84
150	160	155	3	465	2465,33	2025956,148
160	170	165	4	660	5980,44	8941428,938
Итого			120	15160	26986,67	19355662,22
$E = \frac{M_4}{\sigma^4} - 3$, где $M_4 = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 m_i}{\sum_{i=1}^k m_i}$.		$x_{cp} =$	126,33			
		$\sigma^2 =$	224,89			
$r = \frac{M_4}{\sigma^4}$ – мера крутости.		$\sigma =$	15,00			
		$M_4 =$	161297,185	(=19355662,22/120)		
		$r =$	3,19	(=161297,2/15 ⁴)		
		$E =$	0,19	(=3,19-3)		
		$E =$	0,35	=ЭКСПЕСС(B2:B121)		

