

Федеральное агентство по образованию
Государственное образовательное учреждение
высшего профессионального образования
Владимирский государственный университет

Е.М. МАРЧЕНКО

СТАТИСТИКА

Учебное пособие

В трех частях

Часть 1. ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Владимир 2009

УДК 658 я 73

ББК 65.053

М30

Рецензенты:

Доктор экономических наук, профессор директор филиала
Всероссийского заочного финансово-экономического института
в г. Владимире
Л. К. Корецкая

Доктор экономических наук, профессор кафедры экономики
Владимирского филиала Российской академии государственных
служащих при Президенте Российской Федерации
О. В. Лантев

Печатается по решению редакционного совета
Владимирского государственного университета

Марченко, Е. М.

М30 Статистика : учеб. пособие. В 3 ч. Ч. 1. Общая теория ста-
тистики / Е. М. Марченко ; Владим. гос. ун-т. – Владимир :
Изд-во Владим. гос. ун-та, 2009. – 80 с.
ISBN 978-5-89368-973-9

Приводятся теоретический материал и примеры расчета по статистиче-
скому наблюдению, сводке и группировке данных, статистическим величинам,
показателям вариации, выборочному наблюдению, статистическому изучению
взаимосвязи, рядам динамики, индексам.

Предназначено для студентов 1 – 3-го курсов всех форм обучения по
специальностям 080502 – экономика и управление на предприятиях городско-
го хозяйства, 080105 - финансы и кредит, 080111 - маркетинг, 080115 - тамо-
женное дело, изучающих дисциплины «Статистика» и «Социально-экономическая
статистика».

Табл. 25. Ил. 6. Библиогр.: 8 назв.

УДК 658 я 73

ББК 65.053

ISBN 978-5-89368-973-9

© Владимирский государственный
университет, 2009

ОГЛАВЛЕНИЕ

Введение.....	5
1. СТАТИСТИКА КАК НАУКА И ЕЁ ИНФОРМАЦИОННАЯ БАЗА.....	6
1.1. Возникновение статистики как науки.....	6
1.2. Предмет и метод статистики. Единая система учета и статистики. Организация статистики в России.....	7
2. СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ.....	10
2.1. Этапы статистического исследования. Характеристика статистического наблюдения.....	10
2.2. Виды и способы статистического наблюдения.....	11
2.3. Организация статистического наблюдения	13
2.4. Требования, предъявляемые к информации.....	15
3. СВОДКА И ГРУППИРОВКА ДАННЫХ	17
3.1. Понятие сводки, группировки, классификации.....	17
3.2. Ряды распределения: виды и основные характеристики.....	19
3.3. Статистические таблицы: элементы и принципы их построения.....	20
3.4. Графическое изображение рядов распределения	22
4. СТАТИСТИЧЕСКИЕ ВЕЛИЧИНЫ.....	24
4.1. Виды статистических величин.....	24
4.2. Относительные величины.....	24
4.3. Средние величины. Средняя арифметическая: свойства и методы расчета.....	26
4.4. Другие виды степенных средних.....	30
4.5. Структурные средние.....	31
5. ПОКАЗАТЕЛИ ВАРИАЦИИ. ИЗМЕНЧИВОСТЬ.....	35
5.1. Виды показателей вариации.....	35
5.2. Свойства и методы расчета дисперсии.....	37
5.3. Дисперсия альтернативного признака.....	38
5.4. Правило сложения дисперсий.....	39
5.5. Нормальное распределение и его характеристики.....	42
5.6. Моменты распределения. Показатели формы распределения.....	43

6. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ.....	46
6.1. Виды и формы взаимосвязей. Понятие корреляционного и регрессионного анализа.....	46
6.2. Парная регрессия.....	47
6.3. Множественная (многофакторная) регрессия.....	49
6.4. Параметрические методы изучения связи.....	50
6.5. Принятие решений на основе уравнения регрессии.....	52
7. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ.....	54
7.1. Характеристика выборочного наблюдения.....	54
7.2. Ошибки выборки.....	55
7.3. Определение необходимой численности выборки.....	58
7.4. Малая выборка.....	59
8. РЯДЫ ДИНАМИКИ.....	61
8.1. Виды рядов динамики и их характеристика.....	61
8.2. Условия сопоставимости рядов динамики.....	62
8.3. Расчет среднего уровня ряда динамики.....	62
8.4. Показатели рядов динамики.....	63
8.5. Смыкание рядов динамики.....	66
9. ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ.....	69
9.1. Понятие и виды экономических индексов.....	69
9.2. Агрегатные индексы.....	70
9.3. Средневзвешенные, или средние из индивидуальных индексов.....	72
9.4. Средние индексы.....	74
9.5. Индексные методы анализа.....	76
Рекомендательный библиографический список.....	78

Введение

Пособие «Статистика» включает три части: «Общая теория статистики», «Статистические методы прогнозирования», «Социально-экономическая статистика».

Во второй части рассматриваются методы выявления тенденций в рядах динамики, построения прогнозов, выделения сезонной составляющей, учета автокорреляции, применения адаптивных методов в прогнозировании. Третья часть пособия посвящена системе национальных счетов, демографической статистике, статистике уровня жизни, трудовых ресурсов, внешнеэкономической и инновационной деятельности, моральной статистике.

Любой экономист, менеджер, занимающийся планированием, анализом, регулированием, мотивацией, в том или ином виде экономической деятельности в любом секторе национальной экономики, обязательно будет работать с цифровым материалом. Следовательно, чтобы цифры «заговорили» и их можно было правильно интерпретировать, на основе имеющихся данных получить недостающие показатели, необходимо владеть определенным инструментарием. Поэтому методам сбора, обработки и анализа информации посвящена первая часть курса «Общая теория статистики».

Освоение данного материала, позволит студентам проводить статистические наблюдения, группировать данные, исследовать взаимосвязи между признаками, определять надежность полученных характеристик, рассчитывать динамику показателей и влияние отдельных факторов на общий результат, то есть понимать и разговаривать на языке цифр.

1. СТАТИСТИКА КАК НАУКА И ЕЁ ИНФОРМАЦИОННАЯ БАЗА

- *Возникновение статистики как науки* • *Предмет и метод статистики. Единая система учета и статистики. Организация статистики в России*

1.1. Возникновение статистики как науки

Статистику можно рассматривать как:

- 1) совокупность сведений о состоянии какой-либо социально-экономической сферы (например статистика уровня жизни, смертности, брака);
- 2) сферу деятельности людей (проведение переписи, работа комитетов по статистике, заполнение форм статистической отчетности в организациях);
- 3) науку, поскольку статистика имеет свой предмет и методы исследования.

Первые статистические данные имеют тысячелетнюю историю. Цензы на имущество в древнем Китае, Египте, учет родившихся, земельные кадастры в античном мире – все эти документы свидетельствуют о необходимости сбора и обработки информации на всех этапах развития государства.

Возникновение статистики как науки относят к середине XVII в., когда с развитием капитализма в Европе наблюдаются быстрый рост и процветание страховых компаний. В указанный период страховые компании собирают большой массив информации, для обработки которого требуется специальный инструментарий.

В это же время в Германии в крупных университетах начинают читать курс государственоведения, который нуждается в системе показателей для оценки состояния и развития государства. Для решения этих проблем и возникает наука статистика. Пред-

полагается, что термин «статистика» произошел от латинских слов *stato* (государство) и *status* (политическое состояние, положение вещей). Отсюда следует, что статист - это человек, который описывает состояние государства.

Статистика сформировалась как наука на основе двух направлений:

1) политической арифметики в Англии (основатель Уильям Петти (1623 - 1687));

2) описательной школы в Германии (основатели Герман Конринг (1606 - 1681) и Готфрид Ахенваль (1719 - 1772)).¹

1.2. Предмет и метод статистики. Единая система учета и статистики. Организация статистики в России

Предметом статистики является выявление закономерностей развития социально-экономических явлений общества и их взаимосвязи.

Объектом изучения статистики являются **совокупности** данных, то есть **множество единиц** варьирующего признака. Так, в переписи 2002 г. единицей совокупности являлось домохозяйство. Каждая единица совокупности отличается определенным набором признаков. Статистику интересует только варьирующий признак, то есть признак, имеющий разные значения у отдельных единиц совокупности. Статистика, абстрагируясь от случайного единичного, выявляет типичное, общее, что присуще всем единицам совокупности. В соответствии с законом больших чисел при обработке большого массива информации происходит взаимное погашение случайных значений, единичное проявляется как типичное, и закономерности, присущие массовым явлениям, проявляются как закономерности средних величин.

¹ Более подробно с историей развития науки статистики в странах Европы и в России можно ознакомиться: Б.Г. Плошко, И.И. Елисеева. История статистики. - М.: Финансы и статистика, 1990. – 295 с.; Теория статистики: Учебник / Под ред. проф. Г.Л. Громыко. - 2-е изд., перераб. и доп. - М.: ИНФРА-М, 2005. - 476 с. (Классический университетский учебник).

Признаки могут быть как количественными (объемы продукции), так и качественными (категории работников: рабочие, служащие, специалисты, руководители; сельское и городское население).

Показатель представляет количественное выражение признака, например прибыль, себестоимость, средняя заработная плата.

Наука «Статистика» использует набор присущих ей методов:

- сбора информации (статистическое наблюдение);
- обработки информации (сводка и группировка);
- анализа (ряды динамики, индексы, дисперсионный анализ, корреляционно-регрессионный анализ).

В процессе развития науки «Статистика» сформировались следующие самостоятельные научные направления:

1. Общая теория статистики.
2. Экономическая статистика (статистики национальных счетов, статистики строительства, промышленности, торговли и т. д.).
3. Социальная статистика (уровень жизни, демографическая и др.).

Первое направление имеет методологический характер, так как определяет принципы и методы работы с информацией, получаемой в ходе сбора, обработки и анализа данных.

Второе и третье направления связаны с разработкой показателей оценки социально-экономических явлений на макро-, мезо- и микроуровне, например валового внутреннего продукта, уровня занятости, уровня безработицы, миграционного оборота, выручки, прибыли и других.

В мае 2004 г. организована Федеральная служба государственной статистики, которую возглавляет Росстат РФ. Наряду с государственной статистикой существует ведомственная статистика, которую ведут все министерства и ведомства на основе информации, поступающей к ним от подведомственных предприятий и организаций.

Государственная статистика основана на единой системе учета и статистики, которая включает:

1) оперативно-технический учет: первичный, детальный, низовой, в основном использует натурально-вещественные показатели;

2) бухгалтерский учет: непрерывный сплошной учет хозяйственных операций предприятия, опирается на денежную оценку;

3) статистический учет: агрегированный учет, использует результаты вышеперечисленных учетов и собственных методов получения информации, например хронометража, анкетирования, переписи).

Все виды учета взаимосвязаны, а их результаты не должны противоречить друг другу.

Постановлением ВС РФ от 23.10.92 № 3708-1 утверждена Государственная программа перехода РФ на принятую в международной практике систему учета и статистики в соответствии с требованиями развития рыночной экономики.² С этой целью были разработаны и последовательно внедрены Единый государственный регистр предприятий и организаций всех форм собственности (ЕГРПО), Классификатор видов экономической деятельности (ОКВЭД). Органы статистики РФ активно сотрудничают с международными статистическими организациями: Евростатом, Статистической комиссией ООН, МВФ, Мировым банком.

Контрольные вопросы к теме 1

- 1. Что является предметом изучения науки «Статистика»?*
- 2. Назовите объект изучения науки «Статистика».*
- 3. Какие Вы знаете методы науки «Статистика»?*
- 4. Какие два научных направления положили начало формированию науки «Статистика»?*

² Национальное счетоводство : Учеб. / Под ред. Г.Д. Кулагиной. – М. : Финансы и статистика, 1997. - 448 с.

2. СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

- *Этапы статистического исследования. Характеристика статистического наблюдения*
- *Виды и способы статистического наблюдения*
- *Организация статистического наблюдения*
- *Требования, предъявляемые к информации*

2.1. Этапы статистического исследования. Характеристика статистического наблюдения

Любое статистическое исследование включает в себя следующие этапы:

- 1) формулировка целей;
- 2) сбор информации (статистическое наблюдение);
- 3) обработка информации (сводка и группировка);
- 4) анализ информации (индексный, дисперсионный, корреляционно-регрессионный анализ);
- 5) выводы.

Статистическое наблюдение – один из основных этапов исследования, обеспечивающий качество исходной информации. От организации статистического наблюдения зависят другие этапы статистического исследования, так как любая ошибка в статистическом наблюдении может привести к ошибочным выводам.

Цель проведения статистического исследования определяет целесообразность проведения статистического наблюдения, масштабы, сроки, методы проведения наблюдения, объем требуемых ресурсов.

Статистическое наблюдение представляет сбор и регистрацию данных по заранее установленным наиболее существенным признакам. Оно основано на изучении большого массива данных и организовано как планомерный и систематический процесс.

2.2. Виды и способы статистического наблюдения

Статистическое наблюдение можно классифицировать по ряду признаков:

- **по степени охвата** единиц совокупности – сплошное и несплошное (рис. 1);

- **времени** – непрерывное, периодическое, единовременное;

- **источникам информации** – непосредственное наблюдение, документальное наблюдение, опрос;

- **способу проведения наблюдения (сбора данных)** – отчетное, экспедиционное, саморегистрация, анкетное наблюдение.

Сплошное наблюдение представляет собой обследование всех изучаемых единиц совокупности по какому-либо признаку (например сплошная перепись населения).

Несплошное наблюдение представлено монографическим наблюдением, методом основного массива, выборочным методом.

Монографическое наблюдение используется, когда очень мало единиц обследования (например при изучении и распространении передового опыта). Тщательно изучаются и описываются единицы наблюдения, а выводы распространяются на всю изучаемую совокупность.

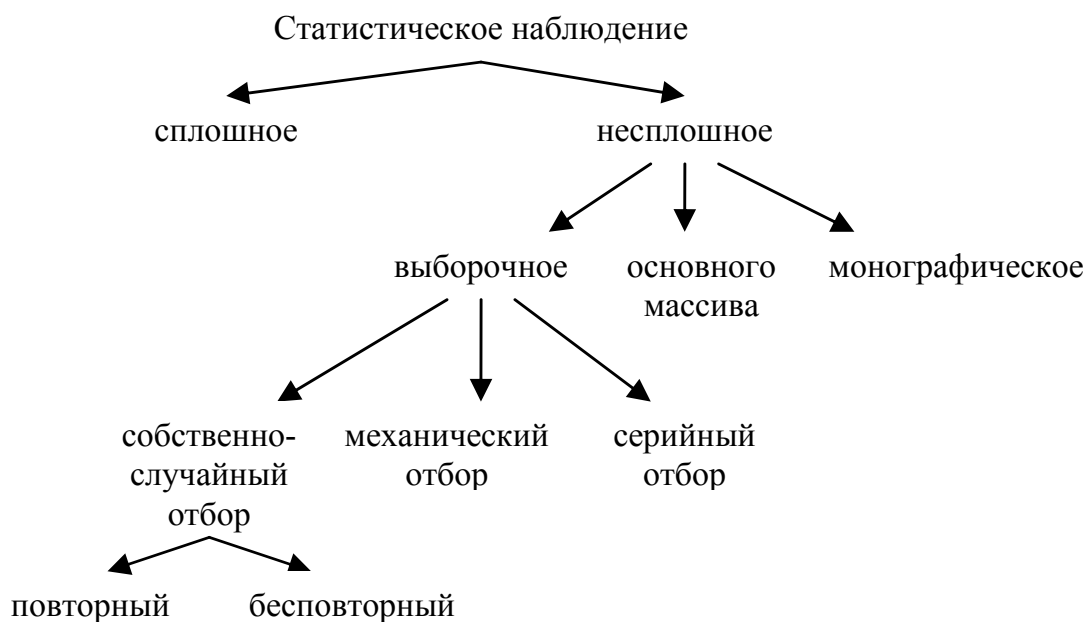


Рис. 1. Виды статистического наблюдения по степени охвата единиц совокупности

При проведении наблюдения методом **основного массива** обследованию подвергаются наиболее важные объекты, имеющие значительный удельный вес в изучаемой совокупности. Границы объекта определяются с помощью введения ценза. Например, при обследовании городского населения рассматриваются города с численностью населения более 300 тыс. жителей, поскольку основная часть городского населения сосредоточена в таких городах.

Выборочный метод предоставляет равные шансы всем единицам наблюдения попасть в выборку (быть обследованными). Он используется при контроле качества продукции, социологических опросах.

Если изучаемая совокупность неоднородна по своему составу, то проводится **типологический** (качественный) отбор, и только после этого обобщаются результаты с учетом процентного соотношения данных групп в общей совокупности.

При проведении выборочного наблюдения применяют механический, собственно-случайный, серийный отбор.

Механический отбор предусматривает ранжирование совокупности по признаку, отличному от группировочного, и последующий отбор каждой 5, 10, 20-й единицы в зависимости от требуемого объема выборки (20, 10, 5%-ный). Например, в переписи 2002 г. одновременно со сплошной переписью проводилась 25%-ная перепись населения, когда обследованию подвергалось каждое четвертое жилое помещение. Жильцы этих квартир отвечали на дополнительные семь вопросов.

Собственно-случайный отбор (жеребьевка может осуществляться повторно или бесповторно. В первом случае у каждой единицы совокупности имеются равные возможности выбора, а во втором они снижаются, и на долю последней приходится то, что осталось. Так, если в лототрон возвращаются выигрышные шары при розыгрыше лотереи, то это пример повторного отбора, а если они не участвуют в дальнейшем розыгрыше, то – бесповторного отбора.

Серийный отбор предусматривает предварительный выбор серий методом собственно-случайного или механического отбора, и затем сплошное наблюдение внутри серии (например при проверке качества продукции, уложенной в коробки).

Непрерывное наблюдение используется при проверке качества продукции, учете выпуска продукции.

С помощью **периодических** наблюдений регулярно проверяется успеваемость учащихся, оценивается ситуация на предприятии путем анализа отчетных документов.

Единовременное наблюдение проводится по мере необходимости (социологический опрос, разработка нового транспортного маршрута).

Непосредственное наблюдение осуществляется лично регистратором (счетчиком), например хронометраж рабочего времени, подсчет пассажиров.

Документальное наблюдение предусматривает использование документов. При опросе осуществляется запись информации со слов опрашиваемого.

Отчетный – сбор отчетности (документов), эффективный способ, поскольку закреплен правовыми документами. **Экспедиционный** способ доставки предполагает выезд регистраторов на места и сбор ими информации (перепись). **Саморегистрация** – заполнение опрашиваемым специального бланка (фотография рабочего дня – самохронометраж). **Анкетный** – заполнение анкеты самим обследуемым, это менее эффективный способ, но более дешевый по сравнению с экспедиционным., хотя использование сети Интернет значительно повышает его эффективность.

2.3. Организация статистического наблюдения

Для проведения наблюдения определяется объект наблюдения, разрабатывается программа, готовятся формуляры и проводится подготовка ответственных за его проведение. Определение объекта наблюдения предполагает установление единицы наблюдения, территории и времени наблюдения.

При этом под **единицей наблюдения** понимается такой элемент совокупности, который является источником информации, а под **единицей совокупности** – элемент, который является носителем признака. Например, при обследовании доходов работников конкретной области на основе статистической отчетности источником информации будут предприятия области, а носителем признака – работники предприятия.

Если обследование проводится в форме анкетирования работников, то источником информации и носителем признака будет работник. Совокупность единиц наблюдения **составляет объект наблюдения**. Введение **ценза** – ограничения – позволяет установить границы объекта. **Территория** проведения наблюдения охватывает все места нахождения единиц наблюдения. Время наблюдения делится на объективное и субъективное. **Объективное время** – время, к которому относятся собираемые данные. **Субъективное время** – это время сбора информации. Например, если годовую отчетность необходимо представить до 15.03.2009 г., то объективное время – год, субъективное – 2,5 мес.

Для проведения статистического наблюдения разрабатывается программа, готовятся формуляры.

Программа наблюдения должна включать перечень признаков, подлежащих регистрации. Не рекомендуется включать лишние вопросы и вопросы, которые могут трактоваться неоднозначно или могут быть подозрительными для опрашиваемых.

Собираемая информация записывается в **формуляры**, т.е. статистические формы, которые бывают бланковыми или списочными. Бланковый формуляр заполняется на одну единицу совокупности, списочный – на ряд единиц. Формуляры должны сопровождаться инструкцией. Перед проведением наблюдения проводится инструктаж. Хорошо организованное наблюдение позволяет избежать ошибок или снизить их до минимума.

Особой формой организации сбора данных является статистическая отчетность. Государственная отчетность определяет сбор данных в установленные сроки по установленному кругу организаций, утвержденным формам и перечню показателей.

С мая 1992 г. введен в действие закон РФ «Об ответственности за нарушение порядка предоставления государственной статистической отчетности», который устанавливает правовую ответственность организаций за нарушение порядка представления отчетности.

2.4. Требования, предъявляемые к информации

В ходе статистического наблюдения могут возникнуть ошибки. Их можно разделить **по характеру возникновения** на случайные и систематические; **по стадии возникновения** – на ошибки регистрации и ошибки компьютерной обработки; **по причинам возникновения** – на ошибки измерения, выборки, преднамеренные и случайные ошибки.

Случайные ошибки при большом массиве информации не имеют направленности, взаимопогашаются и не искажают результаты анализа.

Систематические ошибки имеют определенную направленность, приводят к искажению результатов и, как правило, являются преднамеренными. Они могут быть связаны с округлением возраста, занижением доходов, завышением затрат и другими причинами.

Ошибки репрезентативности – ошибки выборочного наблюдения: чем больше объем выборки, тем меньше ошибка. Это связано с тем, что структура выборки может не совпадать со структурой генеральной совокупности, откуда производилась выборка. При сплошном наблюдении ошибка отсутствует.

Величину ошибки можно определить и скорректировать результат на её величину.

Требования, предъявляемые к информации:

1. **Достоверность:** информация должна быть точно определена и должна правдиво отражать явления.

2. **Своевременность.**

3. **Сопоставимость.**

Для получения сопоставимой информации необходима со-

поставимость по времени, единицам наблюдения, структуре, границам, методике и другим факторам.

Для исключения ошибок при сборе и обработке информации используют методы логического и математического контроля. Логический контроль предполагает проверку с помощью взаимосвязанных вопросов, а математический контроль осуществляется сопоставлением суммы по строкам и столбцам.

Контрольные вопросы к теме 2

- 1. Приведите определение статистического наблюдения.*
- 2. Какие виды сплошного наблюдения Вы знаете?*
- 3. Что характерно для сплошного наблюдения?*
- 4. Приведите основные принципы проведения выборочного наблюдения.*
- 5. Как организовать механический отбор?*
- 6. Приведите примеры собственно-случайного повторного и бесповторного отбора.*
- 7. Чем субъективное время отличается от объективного времени проведения статистического наблюдения?*
- 8. Какие требования предъявляются к информации?*
- 9. Какие методы контроля информации Вы знаете?*

3. СВОДКА И ГРУППИРОВКА ДАННЫХ

- *Понятие сводки, группировки, классификации* • *Ряды распределения: виды и основные характеристики* • *Статистические таблицы: элементы и принципы их построения* • *Графическое изображение рядов распределения*

3.1. Понятие сводки, группировки, классификации

Сводка – систематизация и подведение итогов: метеосводка, сводка с полей. Сводка не позволяет детально проанализировать информацию. Любая сводка должна опираться на группировку данных, т.е. сначала группировка, а потом сводка данных.

Группировка – разделение совокупностей на ряд групп по наиболее существенным признакам.

Различают качественную и количественную группировку. **Качественная** – атрибутивная, **количественная** – вариационная. В свою очередь, вариационная делится на структурную и аналитическую. **Структурная** группировка предполагает расчет удельного веса каждой группы. Пример: на предприятии 80 % - рабочие, 20 % - служащие, из них 5 % - руководители, 12 % - специалисты. Цель **аналитической** группировки – выявить взаимосвязь между признаками: стажем работы и средним заработком, стажем и выработкой и др.

При проведении группировки необходимо:

- провести всесторонний анализ природы изучаемого явления;
- выявить группировочный признак (один или несколько);
- установить границы групп таким образом, чтобы группы существенно отличались друг от друга и в каждой группе объединялись однородные элементы.

По степени сложности группировки могут быть простые и комбинационные (по признакам).

По исходной информации различают первичную и вторичную группировки, **первичная** осуществляется на основе исходных данных наблюдения, **вторичная** использует данные первичной группировки.

Количество групп определяется по формуле Стерджесса:

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

где n - количество групп; N – генеральная совокупность.

Если используются равные интервалы, то **величина интервала** $h = \frac{X_{\max} - X_{\min}}{n}$.

Интервалы могут быть равные и неравные. Последние, в свою очередь, делятся на изменяющиеся по закону арифметической или геометрической прогрессии. Первый и последний интервалы могут быть открытыми или закрытыми. Закрытые интервалы включают или не включают границы интервала.

Если интервалы закрытые и ничего не сказано о включении верхних границ, то считаем, что верхние границы включены.

Если интервалы открытые, то ориентируемся по последнему интервалу.

Признак в этих интервалах может измеряться дискретно и непрерывно (т.е. дробиться). При непрерывном признаке границы смыкаются 1 – 10, 10 – 20, 20 – 30; если признак изменяется дискретно, то можно использовать следующую запись: 1 – 10, 11 – 20, 21 – 30.

Если интервалы открытые, то величина последнего интервала приравнивается к предыдущему, а первого - ко второму.

Классификация – группировка по качественному признаку. Она относительно устойчива, стандартизирована и утверждается органами государственной статистики.

3.2. Ряды распределения: виды и основные характеристики

Под **рядом распределения** понимается ряд данных, характеризующих какое-либо социально-экономическое явление по одному признаку.

Ряды распределения делятся на качественные и количественные, ранжированные и неранжированные, сгруппированные и несгруппированные, с дискретным и непрерывным распределением признака.

Примером несгруппированного, неранжированного ряда по заработной плате является ведомость заработной платы. В то же время список работников может быть ранжированным по алфавиту или по табельным номерам. Примером ранжированного ряда является список команд, рейтинг теннисистов.

Ранжированный ряд распределения - ряд данных, расположенных в порядке убывания или возрастания признака.

Для сгруппированных ранжированных рядов выделяют следующие характеристики: варианту, частоту или частость, кумуляту и плотность распределения.

Варианта (x_i) – среднее интервальное значение признака. Так как при создании группировки должен выполняться принцип равномерного распределения признака в каждом интервале, то варианту можно рассчитывать как полусумму границ интервалов.

Частота (f_i) показывает, сколько раз встречается данное значение признака. Относительное выражение частоты представляет собой **частость** ($f_{i\%}$), то есть долю, удельный вес от суммы частот.

Кумулята (S_i) – накопленная частота или частость, расчет нарастающим итогом. Кумулятивно подсчитываются объем, затраты, доходы, то есть результаты деятельности.

Пример

Рассчитать основные характеристики ряда распределения кре-

дитных организаций РФ по величине уставного капитала (табл. 1).
Результаты расчетов представить в табличной форме.

Таблица 1

*Группировка действующих кредитных организаций
по величине зарегистрированного уставного капитала
в 2008 г. в РФ*

Группировка кредитных организаций по уставному капиталу, млн руб.	Число кредитных организаций f_i	Доля кредитных организаций $f_i\%$, %	Среднее значение уставного капитала по группе x_i	Кумулята по частоте S_i	Кумулята по частоте $S_{\%i}$	Плотность распределения P_i
До 3	37	3,3	1,5	37	3,3	$37/(3-0)=12,33$
3 – 10	61	5,4	6,5	98	8,6	8,71
10 – 30	120	10,6	20	218	19,2	6,00
30 – 60	161	14,2	45	379	33,4	5,37
60 – 150	207	18,2	105	586	51,6	2,30
150 – 300	248	21,8	225	834	73,4	1,65
300 – 450	302	26,6	375	1136	100,0	2,01
<i>Итого</i>	$\Sigma f_i=1136$	100,0	–	–	–	–

Плотность распределения (p_i) используется в группировках с неравными интервалами для обеспечения сопоставимости данных. Определяется как отношение частоты или частости к величине интервала p_i .

3.3. Статистические таблицы: элементы и принципы их построения

Впервые в истории таблицы были использованы в 1727 – 1728 гг. русским статистиком И. Кирилловым при описании царствования Петра Первого.

Для табличного способа характерны компактность, наглядность, доступность.

В любой таблице можно выделить **подлежащее** (о чем говорится в таблице) и **сказуемое** (что говорится о подлежащем) (табл. 2). Если подлежащее включает одну единицу совокупности или перечень единиц, то такая **таблица** называется **простой**, если группировку по одному признаку, – то **сложной**. Если подлежащее или сказуемое содержат группировку по нескольким признакам, то говорят о **комбинационной** таблице.

Таблица 2

Макет статистической таблицы

Сказуемое Подлежащее	Наименование граф			
Наименование строк	2	3	4	Итоговый столбец
				Σ
Итоговая строка	Σ	Σ	Σ	$\Sigma \Sigma$

Принципы построения таблицы:

1. Таблицы должны быть компактны, т.е. содержать те сведения, которые нужны.

2. Материал должен располагаться таким образом, чтобы при анализе таблицы сущность явления раскрывалась чтением информации слева направо и сверху вниз.

3. Заголовок таблицы должен быть лаконичен, содержать название места, времени, объекта и единицу измерения, если она общая для всей таблицы.

4. Графы и строки номеруются цифрами и буквами соответственно.

5. Информация в пределах строки (столбца) должна приводиться с одинаковой степенью точности. Если значение имеет разряд меньший, чем принятая степень точности, то в графе записываются нули заданной степени точности.

6. Если отсутствуют сведения, то в ячейке проставляется многоточие или пишется «Нет сведения», а если отсутствует само явление, то проставляется прочерк.

7. Если ячейка не заполняется, то ставится крестик – X.

8. При переносе таблицы с одной страницы на другую необходимо дублировать «шапку» таблицы и писать «Продолжение таблицы».

3.4. Графическое изображение рядов распределения

Для графического изображения рядов распределения используют графики – гистограмму, полигон, кумуляту, огиву, кривую Лоренца.

Гистограмма строится по интервальным значениям признака и частотам (частостям) (рис. 3). Для построения графика полигона необходимы варианты и частоты (частости) (см. рис. 3). Верхние границы интервалов и накопленные частоты или частости нужны для построения кумуляты (рис. 4). Если поменять местами оси координат, то вместо кумуляты получим огиву. Для отражения динамики показателей используется большое разнообразие диаграмм³.

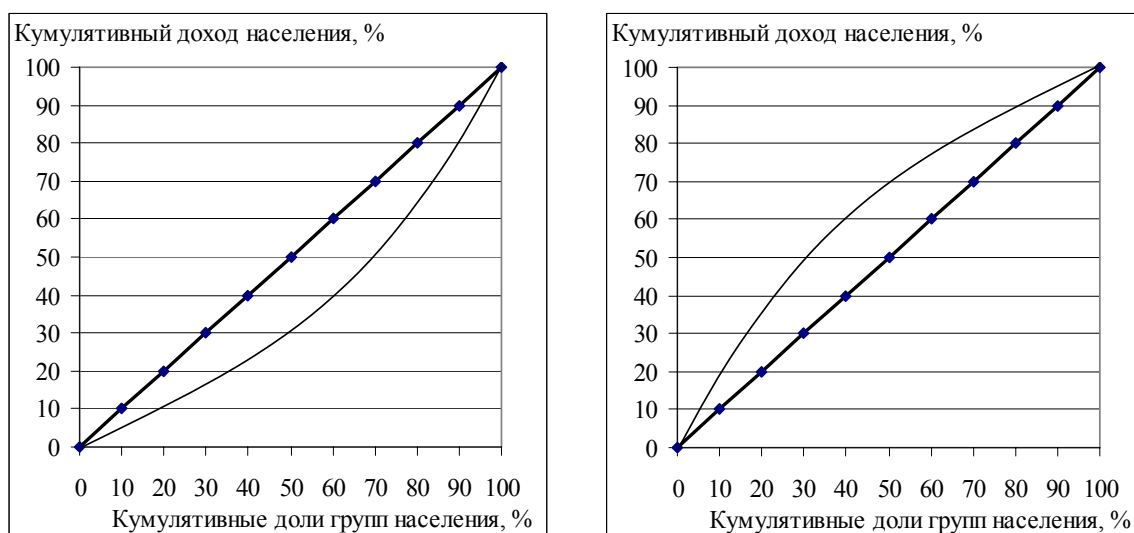


Рис. 2. Кривая Лоренца

³ С видами графиков можно более подробно ознакомиться в [1, 2, 3, 8].

Кривая Лоренца используется для оценки концентрации признака. Кривая Лоренца вогнута, если признак концентрируется в первых группах (см. график слева на рис. 2), и имеет выпуклый характер, если признак сконцентрирован в последних группах (см. график справа на рис. 2).

По оси абсцисс откладывается нарастающим итогом вес групп населения, по оси ординат – кумулятивные доходы групп населения. Чем более неравномерно распределяются доходы, тем больше выпуклость.

Контрольные вопросы к теме 3

- 1. Приведите определение сводки.*
- 2. Чем группировка отличается от классификации?*
- 3. Приведите примеры атрибутивной, структурной и аналитической группировок.*
- 4. Перечислите основные характеристики ряда распределения.*
- 5. Продолжите фразу: «Кумулятивный расчет – это расчет...».*
- 6. Приведите основные принципы построения таблиц.*

4. СТАТИСТИЧЕСКИЕ ВЕЛИЧИНЫ

- *Виды статистических величин* • *Относительные величины* • *Средние величины. Средняя арифметическая: свойства и методы расчета*
- *Другие виды степенных средних* • *Структурные средние*

4.1. Виды статистических величин

Все величины в статистике можно разделить на три группы:

- абсолютные;
- относительные;
- средние.

Абсолютные величины – меры веса, пространства, времени, это исходные величины. Их можно разбить на три группы:

- натуральные (метр, квадратный метр, кубический метр, литр, ящики, штуки, часы, тонны, количество человек);
- условно-натуральные (тысячи условных банок, человеко-день, пассажиро-километр, тонна-километр);
- денежные (все виды валют).

4.2. Относительные величины

Относительные величины – расчетные, производные от натуральных. Различают относительные величины:

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

Относительная величина структуры (ОВС) определяется отношением части и целого. Процентный пункт – разница в процентах разного содержания.

Относительная величина динамики (ОВД) - отношение показателя более позднего периода к одноименному показателю более раннего периода, рассчитанного в сопоставимых условиях.

Относительная величина планового задания (ОВПЗ) - отношение планового показателя более позднего периода к одноименному фактическому показателю более раннего периода.

Относительная величина выполнения плана (ОВВП) - отношение фактического показателя к одноименному плановому показателю одного и того же периода. Выполняется равенство

$$ОВД = ОВВП \cdot ОВПЗ.$$

Пример 4.1

Рассчитать структуру работников в 2007 - 2008 гг. (табл. 3).

Таблица 3

Категории работников

Категория работников	2007 г.		2008 г.	
	Численность, чел.	ОВС, %	Численность, чел.	ОВС, %
Рабочие	70	$70/100 \cdot 100\% = 70$	80	$80/120 \cdot 100\% = 33$
Служащие	30	$30/100 \cdot 100\% = 30$	40	$40/120 \cdot 100\% = 67$
<i>Итого</i>	100	100	120	100

Пример 4.2

По данным табл. 4 рассчитать относительные величины динамики, выполнения плана и планового задания.

Таблица 4

Объем продукции

Показатель, тыс. руб.	2007г.	2008 г.	ОВПЗ	ОВД	ОВВП
План	500	600	$600/400 = 1,5$	$600/500 = 1,2$	$400/500 = 0,8$
Факт	400	800		$800/400 = 2,0$	$800/600 = 1,3$

Относительная величина интенсивности (ОВИ) характеризует плотность распределения признака, например число людей на единицу площади, производство важнейших видов продукции на душу населения.

Относительная величина сравнения (ОВСр) рассчитывается в виде соотношения двух одноименных показателей, определенных за один и тот же период или на момент времени, но по разным территориям, предприятиям, государствам, например соотношение Ивановской и Владимирской областей по площади.

Относительная величина координации показывает соотношение двух частей одного целого (соотношение женщин и мужчин, рабочих и служащих).

При построении относительных величин необходимо соблюдать условия их сопоставимости.

4.3. Средние величины. Средняя арифметическая: свойства и методы расчета

Средняя величина является основной в статистике, поскольку она характеризует центр распределения признака. Средние можно разделить на две группы:

- степенные (средняя арифметическая, геометрическая, гармоническая, квадратическая, кубическая);
- структурные (мода, медиана, квартиль, дециль, перцентиль).

Все виды **степенных** средних получаются на основе формулы

$$\bar{x} = \sqrt[m]{\frac{\sum x^m f}{\sum f}},$$

где x – значение признака; f – частота, вес; m – показатель степени.

Так, **средняя арифметическая** формируется при $m=1$:

$$\bar{x}_{ap.} = \frac{\sum xf}{\sum f}.$$

Если ряд сгруппированный, то используется средняя арифметическая взвешенная.

Для несгруппированного ряда используется средняя арифметическая простая: $\bar{x}_{ap.} = \frac{\sum x}{n}$.

Средняя арифметическая обладает рядом свойств:

1. Средняя от постоянной величины равна ей самой: $\bar{A} = A$.

2. Произведение средней на сумму частот равно сумме произведений вариант на частоты: $\bar{x} \sum f = \sum xf$.

3. Изменение каждой варианты на одну и ту же величину изменяет среднюю на ту же величину: $\frac{\sum (x + A)f}{\sum f} = \bar{x} + A$.

4. Изменение всех вариантов в одно и то же число раз во столько же раз изменяет среднюю: $\frac{\sum (x \cdot A)f}{\sum f} = \bar{x}A$.

5. Изменение всех весов (частот) в одно и то же число раз не изменяет значение средней: $\frac{\sum x \cdot (Af)}{\sum Af} = \bar{x}$.

6. Алгебраическая сумма отклонений всех вариантов от средней равна нулю: $\frac{\sum (x - \bar{x})f}{\sum f} = 0$.

7. Средняя суммы равна сумме средних: $\overline{(x + y)} = \bar{x} + \bar{y}$.

8. Сумма квадратов отклонений вариант от средней меньше, чем от любой другой величины: $\sum (x - \bar{x})^2 < \sum (x - A)^2$.

Методы расчета средней величины

При расчете средней величины в числителе собираются все значения признака, а в знаменателе – общее количество единиц, обладающих данным признаком. Необходимо контролировать, чтобы и в числителе и в знаменателе были показатели, имеющие экономический смысл, и чтобы полученная средняя не выходила за границы минимального и максимального значений признака.

$$\overline{ЗП} = \frac{\PhiЗП}{\overline{Ч}}, \text{ где } \overline{ЗП} - \text{средняя заработная плата; } \PhiЗП -$$

фонд заработной платы; $\overline{Ч}$ – среднесписочная численность.

$$B_{\Gamma} = \frac{Q}{\overline{Ч}}, \text{ где } Q - \text{количество товаров, услуг; } B_{\Gamma} - \text{годовая}$$

выручка.

$$P = \frac{\Pi}{\Phi}, \text{ где } P - \text{рентабельность; } \Pi - \text{прибыль.}$$

$$\bar{x} = \frac{\sum xf}{\sum f}, \text{ где } \sum xf - \text{все значение признака; } \sum f - \text{количе-}$$

ство единиц признака.

Пример 4.3

30 предприятий выполнили план на 80 %, 20 предприятий – на 102 %, 50 предприятий – на 98 %. Определить средний процент выполнения плана по всем предприятиям.

Решение

$$\bar{x} = \frac{30}{100\%} \cdot 80\% + \frac{20}{100\%} \cdot 102\% + \frac{50}{100\%} \cdot 98\% = 93,4\% .$$

Пример 4.4

Рассчитать среднедушевой доход по приведенным данным (табл. 5).

Таблица 5

Среднесписочные доходы работников, руб.

Группировка работников по доходам	Количество человек, f	Средний доход по группам x_i , руб.	$x_i f_i$, руб.
До 5000	10	2500	25000
5000 - 10000	20	7500	150000
10000 - 15000	25	12500	312500
15000 - 20000	30	17500	525000
20000 и выше	15	22500	337500
<i>Итого</i>	$\sum f_i = 100$	–	$\sum x_i f_i = 1350000$

Решение

$$\bar{x} = \frac{\sum xf}{\sum f} = 1350000 / 100 = 13500.$$

Метод моментов (метод условного нуля)

$$\bar{x} = m_1 h + A, \quad m_1 = \frac{\sum \frac{(x - A)}{h} f}{\sum f}, \quad \text{где } m_1 \text{ – момент первого по-}$$

рядка; h – величина интервала; A – варианта, соответствующая максимальной частоте.

Пример 4.5

По данным табл. 6 рассчитать среднедушевой доход методом моментов.

Решение

Для нашего примера $h = 5000$; $A = 17500$; $m_1 = \frac{-80}{100} = -0,8$.

$$\bar{x}_{ар.} = -0,8 \cdot 5000 + 17500 = 13500.$$

Таблица 6

Исходные данные

Среднедушевой доход, руб.	Количество человек, f	x_i , руб.	$\frac{(x - A)}{h}$	$\frac{(x - A)}{h} f$
До 5000	10	2500	-3	-30
5000 - 10000	20	7500	-2	-40
10000 - 15000	25	12500	$\frac{(12500 - 17500)}{5000} = -1$	-25
15000 - 20000	30	17500	$\frac{(17500 - 17500)}{5000} = 0$	0
20000 и выше	15	22500	1	15
<i>Итого</i>	$\sum f = 100$	-	-	$\sum \frac{(x - A)}{h} f = -80$

4.4. Другие виды степенных средних

Из общей формулы $\bar{x} = \sqrt[m]{\frac{\sum x^m f}{\sum f}}$ можно получить другие

виды степенных средних.

При $m = -1$ рассчитывается **средняя гармоническая**:

взвешенная – $\bar{x}_{гарм.} = \frac{\sum xf}{\sum xf \frac{1}{x}}$ для сгруппированного ряда;

простая – $\bar{x}_{гарм.} = \frac{n}{\sum \frac{1}{x}}$ для несгруппированного ряда.

Средняя гармоническая используется, когда неизвестны весовые коэффициенты (частоты или частоты).

Пример 4.6

Определить среднюю цену реализации товара по данным трех торговых точек (табл. 7).

Таблица 7

Цены реализации

Торговая точка	Цена, руб.	Выручка, тыс. руб.
1	20	20
2	25	50
3	18	180

Решение

$V = ЦQ_{ед}$, где V – выручка; $Ц$ – цена единицы продукции.

$$\bar{x}_{гарм.} = \frac{20\ 000 + 50\ 000 + 1\ 80\ 000}{\frac{20\ 000}{20} + \frac{50\ 000}{25} + \frac{180\ 000}{18}} = \frac{250\ 000}{13} = 19,2 \text{ руб.}$$

При $m = 0$ рассчитывается **средняя геометрическая**

$\bar{x}_{геом.} = \sqrt[n-1]{\prod k_{2/1} \dots k_{n/(n-1)}}$, используется в рядах динамики при расчете среднего темпа роста.

При $m = 2$ рассчитывается **средняя квадратическая**

$$\bar{x}_{кв.} = \sqrt{\frac{\sum x^2 f}{\sum f}}.$$

При $m = 3$ рассчитывается **средняя кубическая**

$$\bar{x}_{куб.} = \sqrt[3]{\frac{\sum x^3 f}{\sum f}}.$$

Средняя квадратическая и средняя кубическая используются для расчета моментов 2-го и 3-го порядков.

Свойство мажорантности (ранжирование степенных средних): чем больше степень, тем больше средняя.

$$\bar{x}_{гарм.} \leq \bar{x}_{геом.} \leq \bar{x}_{арифм.} \leq \bar{x}_{кв.} \leq \bar{x}_{куб.}$$

4.5. Структурные средние

К структурным средним относят моду, медиану, квартиль, дециль, перцентиль.

Мода – наиболее часто встречающееся значение признака, которое используют для оценки спроса на продукцию, то есть мода – это варианта, соответствующая максимальной частоте.

Если для определения моды для дискретного признака не требуется никаких расчетов, то для определения моды в интервальном ряду необходимо:

1. Определить модальный интервал по максимальной частоте.
2. Рассчитать значение моды M_o по формуле

$$M_o = x_{m_o} + h_{m_o} \frac{f_{m_o} - f_{m_o-1}}{(f_{m_o} - f_{m_o-1}) + (f_{m_o} - f_{m_o+1})},$$

где m_o – модальный интервал; x_{m_o} – нижняя граница модального интервала; h_{m_o} – величина модального интервала; f_{m_o} , f_{m_o-1} ,

f_{m_0+1} – частоты модального интервала, предшествующего модальному и следующего за модальным соответственно.

Пример 4.7

Рассчитать значение моды по приведенной группировке (табл. 8).

Таблица 8

Данные для расчета

Среднедушевой доход, руб.	Количество человек	Кумулята
До 5000	10	10
5000 – 10000	20	30
10000 – 15000	25	55
15000 – 20000	30	85
20000 и более	15	100

Для данного примера значение моды $M_0 = 15000 + 5000 \times \frac{(30 - 25)}{(30 - 25) + (30 - 15)} = 16250$.

Графически определить моду можно в соответствии с рис. 3.

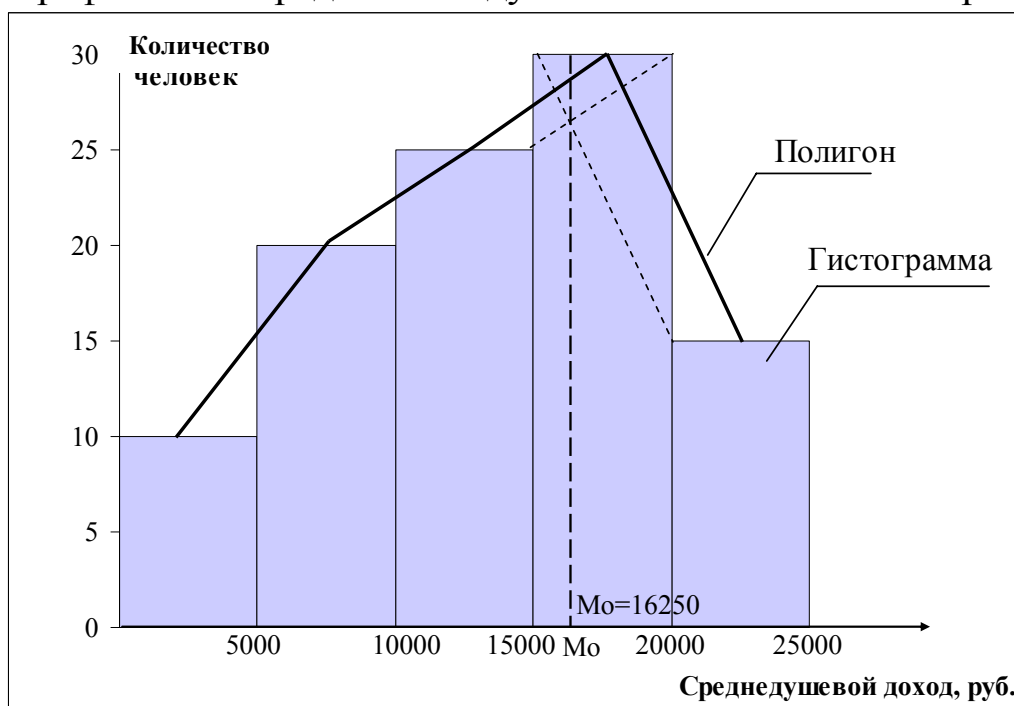


Рис. 3. Гистограмма и полигон.

Графическое определение моды на основании гистограммы

Медиана – значение признака, соответствующее середине ранжированного ряда. Используется для характеристики ряда распределения при сильной дифференциации признака.

Номер медианного представителя дискретного ряда определяется по формуле $h_{me} = \frac{n+1}{2}$, если число членов ряда распределения нечетное. Если ряд включает четное число членов, то в середине находятся две единицы наблюдения, и значение медианы рассчитывается как средняя арифметическая их значений.

Пример 4.8

Определить медианный размер обуви по данным табл. 9.

Таблица 9

	<i>Объем продаж</i>				
Размер обуви	36	37	38	39	40
Количество пар	5	7	10	6	5
Кумулята	5	12	22	28	33

Решение

$h_{me} = \frac{33+1}{2} = 17$, пара соответствует 38-му размеру обуви, так как $22 > 17$.

Медиана интервального ряда определяется в два этапа:

1. Находится медианный интервал по полусумме частот и кумуляте. Кумулята медианного интервала впервые превышает полусумму частот.

2. Рассчитывается значение медианы по формуле

$$Me = x_{me} + h_{me} \frac{\left(\frac{\sum f}{2} - S_{me-1} \right)}{f_{me}},$$

где x_{me} – нижняя граница медианного интервала; h_{me} – величина медианного интервала; S_{me-1} – кумулята интервала, предшествующего медианному; f_{me} – частота медианного интервала.

Медиана показывает, что 50 % единиц совокупности имеют значения признака, не превышающие медианного значения.

Квартили, децили, перцентили рассчитываются по аналогии с медианой и характеризуют соответственно, до какого предела значения признака имеет 25, 10, 1 % представителей.

Для примера 4.7: $Me = 10000 + 5000 \frac{(50 - 30)}{25} = 14000$.

Графически определить медиану можно в соответствии с рис. 4.

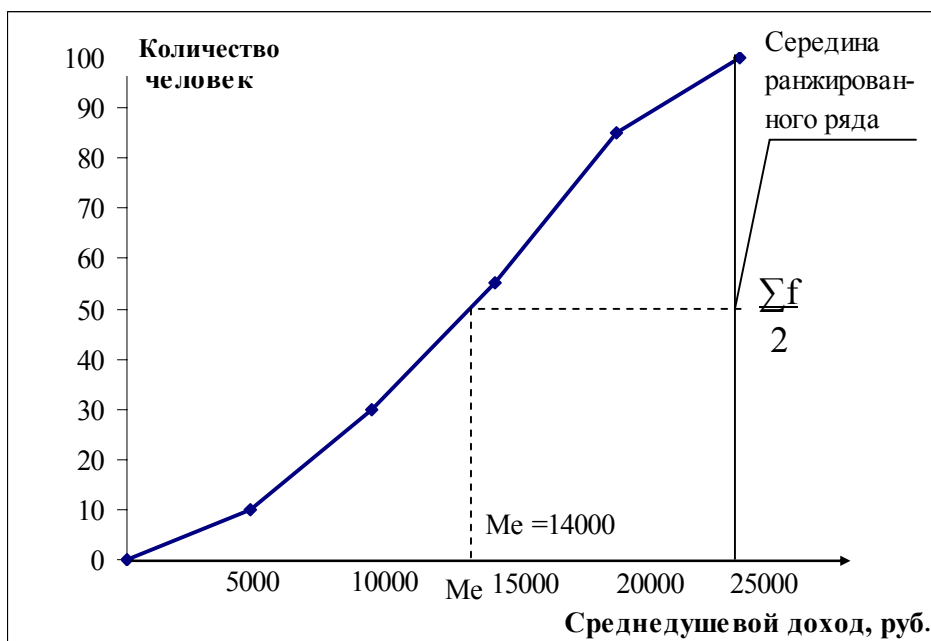


Рис. 4. Кумулята. Графическое определение медианы на основании кумуляты

В примере 4.7 значения $\bar{x}_{ар.}$, Me и Mo говорят о том, что среднедушевой доход работников составляет 13500 руб., у 50 % работников доход не превышает 14000 руб., чаще всего встречаются работники с уровнем дохода 16250 руб.

Контрольные вопросы к теме 4

1. Приведите пример условно-натуральной величины.
2. Перечислите все относительные величины.
3. Перечислите степенные средние в порядке возрастания их значения.
4. Какие Вы знаете структурные средние?
5. Приведите определение медианы.
6. Какое значение признака называется модальным?

5. ПОКАЗАТЕЛИ ВАРИАЦИИ. ИЗМЕНЧИВОСТЬ

- *Виды показателей вариации* • *Свойства и методы расчета дисперсии* • *Дисперсия альтернативного признака*
- *Правило сложения дисперсий* • *Нормальное распределение и его характеристики* • *Моменты распределения.*
Показатели формы распределения

5.1. Виды показателей вариации

Показатели вариации характеризуют величину отклонений всех значений признака от среднего уровня, то есть однородность совокупностей. Если совокупность признака однородна, то средняя, рассчитанная по данной совокупности, будет надежна, типична. Если отклонения значительные, то совокупность неоднородна, а рассчитанная средняя - случайна.

Пример 5.1

Один и тот же уровень потребления мяса и мясопродуктов $\bar{x}_{ар.} = 53$ кг можно получить с помощью распределения уровня потребления указанных продуктов населением $f_1\%$ или $f_2\%$ (табл. 10). Но если в первом распределении 70 % населения потребляют от 40 до 60 кг мясопродуктов, то во втором распределении – лишь 11 %. Поэтому доверие к средней в первом распределении будет выше, чем во втором

Таблица 10

Группировка населения

Объем потребления мясопродуктов, кг	$f_1\%$	$f_2\%$
20 – 40	10	50
40 – 60	70	11
60 – 80	15	13
80 – 100	5	26
<i>Итого</i>	100 %	100 %

Оценить степень доверия к средней, ее надежность позволяют показатели вариации, к которым относят: размах вариации R , среднее линейное отклонение d , дисперсию σ^2 , среднее квадратическое отклонение σ , коэффициент вариации v .

R – **размах вариации**, показывает амплитуду колебаний признака от минимума до максимума, но не дает распределения признака между ними.

$$R = x_{\max} - x_{\min}.$$

d – **среднее линейное отклонение**, характеризует среднее отклонение по модулю индивидуальных значений признака от среднего уровня.

$$d = \frac{\sum |x - \bar{x}| f}{\sum f}.$$

σ^2 – **дисперсия**, не имеет экономического смысла и соответственно единиц измерения. Используется для определения среднего квадратического отклонения.

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

σ – **среднее квадратическое отклонение**, имеет ту же размерность, что и средняя величина.

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

Поскольку в соответствии со свойством средней сумма отклонений всех вариантов от средней всегда равна нулю, для расчета средней используют либо среднее линейное отклонение d , либо среднее квадратическое отклонение σ . В силу различий способов расчета $\sigma > d$.

Относительной мерой вариации служит **коэффициент вариации** $v = \frac{\sigma}{\bar{x}} 100 \%$.

Для экономических расчетов **средняя считается надежной**, типичной, а исследуемая совокупность однородной, если **коэф-**

фициент вариации не превышает 33,3 %. В случае превышения данной величины можно говорить о случайности средней характеристики и неоднородности исходных данных. Для повышения надежности рекомендуется исключить экстремальные значения, если они носят случайный характер, и увеличить объем выборки.

5.2. Свойства и методы расчета дисперсии

Свойства дисперсии:

1. Величина дисперсии не изменится, если из всех значений вариант вычесть какое-либо число A .

$$\sigma^2(x - A) = \sigma^2(x).$$

2. Если все значения вариант увеличить или уменьшить в A раз, то σ^2 увеличится (уменьшится) в A^2 раз, а σ - в A раз.

$$\sigma^2(xA) = A^2\sigma^2(x).$$

3. Если все частоты увеличить (уменьшить) в A раз, то σ^2 и σ не изменятся.

4. Дисперсия равна разности среднего квадрата и квадрата средней величины.

$$\begin{aligned} \sigma^2 &= \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{\sum (x^2 - 2x\bar{x} + \bar{x}^2) f}{\sum f} = \frac{\sum (x^2 f - 2x\bar{x}f + \bar{x}^2 f)}{\sum f} = \\ &= \frac{\sum (x^2 f)}{\sum f} - 2\bar{x} \frac{\sum (xf)}{\sum f} + \frac{\sum (\bar{x}^2 f)}{\sum f} = \frac{\sum (x^2 f)}{\sum f} - 2\bar{x}^2 + \bar{x}^2 = \overline{x^2} - \bar{x}^2, \end{aligned}$$

где $\overline{x^2}$ – средний квадрат индивидуальных значений признака;
 \bar{x}^2 – квадрат средней величины.

Расчет дисперсии методом моментов

$$\sigma^2 = h^2(m_2 - m_1^2),$$

$$\text{где } m_1 = \frac{\sum \left(\frac{x-A}{h}\right) f}{\sum f}; \quad m_2 = \frac{\sum \left(\frac{x-A}{h}\right)^2 f}{\sum f}.$$

Пример 5.2

Оценить надежность средней $\bar{x} = 53$ кг для группировки населения по мясопотреблению (табл. 11).

Таблица 11

Данные для расчета

Объем потребления мясопродуктов, кг	$f\%$	\bar{x}_i	$\frac{(x-A)}{h}$	$\frac{(x-A)}{h} f$	$\left(\frac{x-A}{h}\right)^2$	$\left(\frac{x-A}{h}\right)^2 f$
20–40	10	30	-1	-10	1	10
40–60	70	50	0	0	0	0
60–80	15	70	1	15	1	15
80–100	5	90	2	10	4	20
Итого	$\Sigma=100\%$		-	$\Sigma=15$	-	$\Sigma=45$

Решение

$$h = 20; A = 50;$$

$$m_1 = 15/100 = 0,15;$$

$$m_2 = 45/100 = 0,45;$$

$$\sigma^2 = 400(0,45 - 0,0225) = 400 \cdot 0,1285 = 171;$$

$$\sigma = \sqrt{171} = 13,08 \text{ кг.}$$

$$\bar{x} = 53, \quad v = \frac{\sigma}{\bar{x}} 100 \% = \frac{13,08}{53} 100 \% \approx 25 \%, \quad v < 33,3 \%$$

Вывод: средняя надежна и типична для данной группировки.

5.3. Дисперсия альтернативного признака

Единицы совокупности могут либо обладать альтернативным признаком, либо нет.

Приняты обозначения:

1 – наличие признака; 0 – отсутствие признака; p – доля единиц, обладающих данным признаком; q – доля единиц, не обладающих данным признаком; $q + p = 1$ (или 100 %).

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{(1 \cdot p + 0 \cdot q)}{q + p} = \frac{p}{1} = p.$$

Для альтернативного признака средняя величина равна доле единиц, обладающих данным признаком.

Дисперсию альтернативного признака рассчитывают следующим образом:

$$\begin{aligned} \sigma^2 &= \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1 - p)^2 p + (0 - p)^2 q}{p + q} = \frac{q^2 p + p^2 q}{1} = \\ &= pq(q + p) = pq ; \\ \sigma &= \sqrt{pq} . \end{aligned}$$

5.4. Правило сложения дисперсий

Если совокупность разбить на группы, то средние величины и дисперсию можно рассчитать как для всей совокупности, так и для каждой группы.

Различают внутригрупповую среднюю из групповых, межгрупповую и общую дисперсии. Общая дисперсия отражает влияние всех возможных факторов. Внутригрупповая дисперсия отражает влияние всех факторов, кроме группировочного признака. Средняя из групповых аналогична внутригрупповым дисперсиям. Межгрупповая дисперсия характеризует влияние только группировочного признака.

В соответствии с правилом сложения дисперсии **общая дисперсия равна сумме средней из групповых и межгрупповой дисперсии.**

$$\sigma^2 = \overline{\sigma_i^2} + \delta^2 ,$$

где $\overline{\sigma_i^2}$ – средняя из групповых дисперсий; δ^2 – межгрупповая дисперсия.

Средняя из групповых дисперсий определяется по формуле

$$\overline{\sigma_i^2} = \frac{\sum \sigma_i^2 f_i}{\sum f_i} .$$

Межгрупповая дисперсия определяется по формуле

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

Пример 5.3

Проверим правило сложения дисперсий на примере расчета выработки рабочих разных уровней квалификации (табл. 12).

Решение

Рассчитаем групповые средние:

$$\bar{x}_{i1} = \frac{36}{4} = 9 \text{ дет.}, \quad \bar{x}_{i2} = \frac{63}{6} = 10,5 \text{ дет.}$$

$$\bar{x}_i = \frac{36 + 63}{10} = 9,9 \text{ дет.}$$

Таблица 12

Группировка рабочих по уровню квалификации

Рабочие 5-го разряда			Рабочие 6-го разряда		
№ п/п	Количество деталей \bar{x}_i	\bar{x}_i^2	№ п/п	Количество деталей \bar{x}_i	\bar{x}_i^2
1	8	64	1	9	81
2	8	64	2	9	81
3	9	81	3	10	100
4	11	121	4	10	100
-	-	-	5	12	144
-	-	-	6	13	169
$\Sigma=4$	$\Sigma=36$	$\Sigma=330$	$\Sigma=6$	$\Sigma=63$	$\Sigma=675$

Определим групповые дисперсии методом разности:

$$\sigma_1^2 = \overline{x_1^2} - \bar{x}_1^2 = \frac{330}{4} - 9^2 = 82,5 - 81 = 1,5;$$

$$\sigma_2^2 = \overline{x_2^2} - \bar{x}_2^2 = \frac{675}{6} - 10,5^2 = 112,5 - 110,25 = 2,25;$$

$$\overline{\sigma_i^2} = \frac{(1,5 \cdot 4 + 2,25 \cdot 6)}{10} = 1,95.$$

Рассчитаем межгрупповую дисперсию:

$$\begin{aligned} \delta^2 &= \frac{\sum (\bar{x}_i - \bar{x})^2 f_i}{\sum f_i} = \frac{(9 - 9,9)^2 \cdot 4 + (10,5 - 9,9)^2 \cdot 6}{10} = \\ &= \frac{0,81 \cdot 4 + 0,0256 \cdot 6}{10} = 0,54. \end{aligned}$$

Общая дисперсия

$$\sigma^2 = \overline{x^2} - \bar{x}^2 = \frac{330 + 675}{10} - 9,9^2 = 2,49;$$

$$\sigma^2 = \overline{\sigma_i^2} + \delta^2 = 1,95 + 0,54 = 2,49, \text{ что подтверждает правило сложения дисперсий.}$$

Для оценки влияния группировочного признака (уровня квалификации) используют показатели, построенные на соотношении межгрупповой и общей дисперсии: эмпирический коэффициент детерминации η^2 и эмпирическое корреляционное отношение η .

Эмпирический коэффициент детерминации η^2 рассчитывается по формуле $\eta^2 = \frac{\delta^2}{\sigma^2} 100\%$ и показывает, какой процент общей вариации изучаемого признака определяется вариацией группировочного признака.

$$\text{Для рассматриваемого примера } \eta^2 = \frac{0,54}{2,49} 100\% = 21,7\%, \text{ т.е.}$$

вариация выработки рабочих на 21,7% определяется вариацией уровня их квалификации.

Эмпирическое корреляционное отношение η характеризует тесноту связи между признаками и рассчитывается по формуле $\eta = \sqrt{\eta^2}$. $0 \leq \eta \leq 1$. Связь отсутствует, если $\eta = 0$, связь функциональная, если $\eta = 1$.

Сила связи определяется в соответствии со шкалой Чеддока, которая представлена в табл. 12.

Таблица 12
Шкала Чеддока для определения
силы связи

η	Сила связи
0,1 – 0,3	Слабая
0,3 – 0,5	Умеренная
0,5 – 0,7	Заметная
0,7 – 0,9	Высокая
0,9 – 0,99	Очень высокая

В рассматриваемом примере связь выработки рабочих с уровнем квалификации умеренная, так как $\eta=0,47$.

5.5. Нормальное распределение и его характеристики

Анализ эмпирических частот вариационного ряда позволяет выявить тип распределения, подобрать теоретическую кривую, определить ее параметры и проверить правильность гипотезы о типе распределения данного ряда. Характер кривой лучше выявляется при большом объеме наблюдений и малых интервалах.

Большое значение имеет сопоставление фактических кривых с теоретическими. Наиболее часто в качестве теоретической кривой используют **кривую нормального распределения** (рис. 5), плотность распределения которой выражается уравнением

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}t^2},$$

где $f(x)$ – ордината кривой нормального распределения (частость);

t – нормированное отклонение $t = \frac{x-\bar{x}}{\sigma}$; $\pi = 3,14$; $e = 2,78$.

Характеристики кривой нормального распределения:

- кривая симметрична и имеет максимум $f(x) = \frac{1}{\sigma\sqrt{2\pi}}$ в

точке $\bar{x} = Mo = Me$;

- кривая асимптотически стремится к оси абсцисс;

- кривая имеет две точки перегиба на расстоянии $\pm\sigma$ от \bar{x} ;

- при $x = \text{const}$ с ростом σ кривая становится более полой (рис. 5, а);

- при $\sigma = \text{const}$ с изменением средней кривая не меняет своей формы, а лишь сдвигается вправо или влево по оси x (рис. 5, б);

- в интервале $\bar{x} \pm \sigma$ находится 68,3 % всех единиц совокупности, в интервале $\bar{x} \pm 2\sigma$ – 95,4 %, в интервале $\bar{x} \pm 3\sigma$ – 99,7 % (практически все значения);

- коэффициенты эксцесса и асимметрии равны нулю.

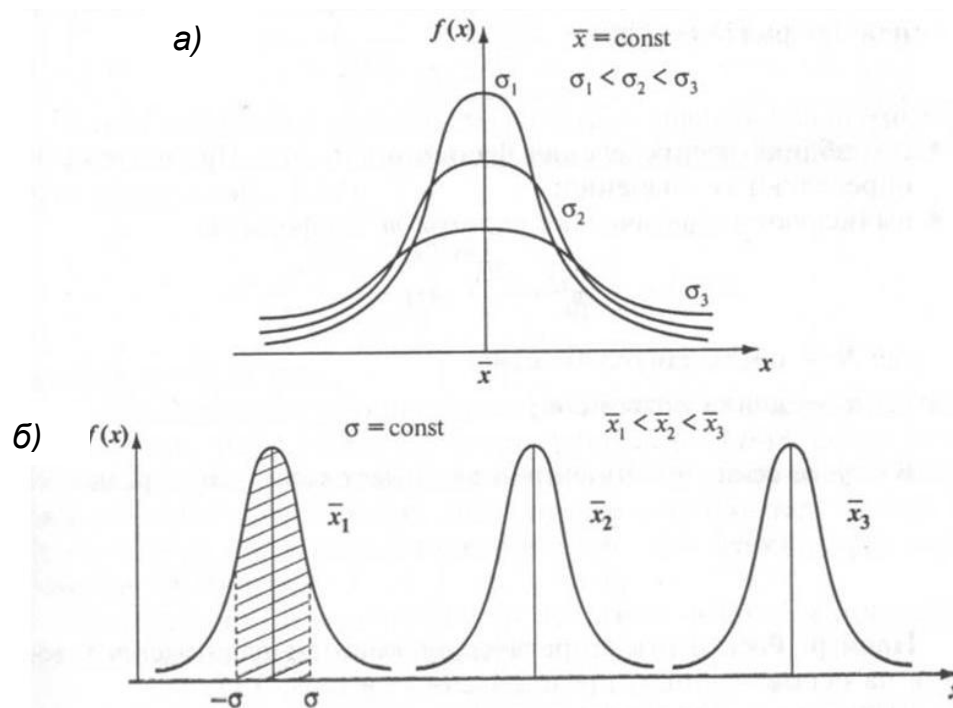


Рис. 5. Кривые нормального распределения

5.6. Моменты распределения. Показатели формы распределения

Для описания особенностей распределения используют такую характеристику, как момент. **Момент k -го порядка** – это

средний размер k -й степени отклонений всех значений признака

$$x \text{ от какой-либо величины: } M_k = \frac{\sum (x_i - A)^k f_i}{\sum f_i}.$$

Моменты характеризуются видом и порядком. В зависимости от величины A рассматривают следующие **виды** моментов:

- начальные, при $A = 0$;
- центральные, при $A = \bar{x}$;
- условные, при $A \neq \bar{x} \neq 0$.

Порядок момента характеризуется показателем степени отклонения. Различают моменты первого, второго, третьего, четвертого порядка.

Формулы расчета моментов представлены в табл. 13.

Таблица 13

Формулы расчета моментов

Порядок моментов	Виды моментов		
	Начальные	Центральные	Условные
1	$M_1 = \frac{\sum x_i f_i}{\sum f_i} = \bar{x}$	$\mu_1 = \frac{\sum (x_i - \bar{x}) f_i}{\sum f_i} = 0$	$m_1 = \frac{\sum (x_i - A) f_i}{\sum f_i} = \bar{x} - A$
2	$M_2 = \frac{\sum x_i^2 f_i}{\sum f_i}$	$\mu_2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$ $\mu_2 = \sigma^2$	$m_2 = \frac{\sum (x_i - A)^2 f_i}{\sum f_i}$
3	$M_3 = \frac{\sum x_i^3 f_i}{\sum f_i}$	$\mu_3 = \frac{\sum (x_i - \bar{x})^3 f_i}{\sum f_i}$	$m_3 = \frac{\sum (x_i - A)^3 f_i}{\sum f_i}$
4	$M_4 = \frac{\sum x_i^4 f_i}{\sum f_i}$	$\mu_4 = \frac{\sum (x_i - \bar{x})^4 f_i}{\sum f_i}$	$m_4 = \frac{\sum (x_i - A)^4 f_i}{\sum f_i}$

При сравнении эмпирической и теоретической кривых распределения обращают внимание на показатели асимметрии распределения и на эксцесс – крутизну распределения по сравнению с нормальной.

Асимметрию можно измерить с помощью коэффициента асимметрии Пирсона, который рассчитывается по формуле $As = \frac{\bar{x} - Mo}{\sigma}$. Если $\bar{x} > Mo$, то говорят о правосторонней асимметрии, если $\bar{x} < Mo$, то асимметрия левосторонняя.

Эксцесс – крутизна распределения по сравнению с нормальной – вычисляется с помощью нормированного момента 4-го порядка по формуле $Ex = \frac{\mu_4}{\sigma^4} - 3$. При $Ex < 0$ говорят о низкововершинности распределения, а при $Ex > 0$ - о высоковершинности распределения.

Контрольные вопросы к теме 5

- 1. Для чего нужны показатели вариации?*
- 2. Какие показатели относятся к показателям вариации?*
- 3. Какой признак называется альтернативным?*
- 4. Чему равны средняя и дисперсия альтернативного признака?*
- 5. Приведите правило сложения дисперсий.*
- 6. Приведите характеристики нормального распределения.*

6. СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

• Виды и формы взаимосвязей. Понятие корреляционного и регрессионного анализа • Парная регрессия • Множественная (многофакторная) регрессия • Параметрические методы изучения связи • Принятие решений на основе уравнения регрессии

6.1. Виды и формы взаимосвязей.

Понятие корреляционного и регрессионного анализа

Различают функциональные и стохастические (вероятностные) взаимосвязи между признаками.

Функциональная связь – жесткая обратимая зависимость, когда каждому изменению факторного признака (независимого) соответствует строго определенное изменение результативного, т.е. зависимого, признака. Например, заработная плата рабочего-сдельщика прямо пропорциональна выполненному объему. Из этой зависимости можно определить расценку или объем произведенной продукции.

Если изменение факторного признака или комбинации признаков приводит к вариации значений результативного, то говорят о наличии стохастической, или корреляционной (частного случая), связи. Этот вид взаимосвязи нежесткий, необратимый.

Корреляционная связь – частный случай стохастической связи, при этом среднее изменение факторного признака приводит к изменению среднего значения результативного признака. Например, с ростом стажа, как правило, растет средняя заработная плата. С увеличением вносимых удобрений, как правило, растет урожайность. Если с ростом факторного признака растет среднее значение результативного, – связь прямая, если уменьшается, – связь обратная.

Если рассматривать влияние только одного факторного признака, имеют место парные корреляции.

Если исследуется влияние нескольких факторных признаков на результативные, говорят о множественной, многофакторной корреляции.

Для исследования функциональной зависимости применяют балансовый и индексный методы.

Для исследования корреляционной связи используют группу методов для изучения качественных и количественных признаков. Связь количественных признаков изучается параметрическими методами (параллельных рядов; графическим; аналитических группировок; корреляционно-регрессионным анализом), а качественных – с помощью непараметрических методов на основе таблиц сопряженности.⁴ Наиболее точным и детальным является корреляционно-регрессионный анализ (КРА).

Задача корреляционного анализа – определение тесноты связей. Цель регрессионного анализа - определение вида и параметров уравнения теоретической кривой. Недостатками корреляционно-регрессионного анализа является необходимость обработки большого массива сопоставимых данных. Выводы, полученные на основе КРА, можно использовать только для данных, сопоставимых с исходными, на основе которых проводился КРА.

6.2. Парная регрессия

Парная регрессия отражает связь между двумя признаками – результативным и факторным. Аналитически эту связь можно представить уравнениями

$$y = a_0 + a_1 x \text{ - прямая,}$$

⁴ Непараметрические методы подробно рассматриваются в курсе «Экономико-математические методы и модели».

$$y = a_0 + a_1 \frac{1}{x} - \text{гипербола,}$$

$$y = a_0 + a_1 x + a_2 x^2 - \text{парабола.}$$

Для того чтобы определить, какая это связь, необходимо изучить само явление, то есть выявить: возможно ли монотонное возрастание или убывание функции, нет ли точек насыщения, перегиба, асимптот. После выбора возможного типа теоретической кривой определяются ее параметры с помощью метода наименьших квадратов «МНК».

Параметры выбранного типа теоретической кривой подбираются исходя из условия, что сумма квадратов отклонений фактических значений результативного признака от теоретических будет минимальной. $S = \sum (y - \bar{y})^2 \rightarrow \min$.

С этой целью на основе решения системы нормальных уравнений, определенных для каждого вида теоретической кривой, рассчитываются коэффициенты регрессии. Например, для линейной парной регрессии $y = a_0 + a_1 x$ система нормальных уравнений имеет вид

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

где x и y – данные наблюдения.

Решение этой системы относительно a_0 и a_1 позволит определить коэффициенты регрессии в уравнении регрессии и составить уравнение теоретической кривой. Коэффициент при x показывает, на сколько своих единиц измерения изменяется результативный признак при изменении факторного признака на свою единицу измерения. Свободный член уравнения характеризует влияние всех прочих факторов кроме x .

Пример

Составить уравнение регрессии по данным табл. 14.

Зависимость прибыли от затрат на 1 руб. продукции

№ п/п	Прибыль y , тыс. руб.	Затраты на 1 руб. x , руб.	x^2	xy	yx
1	1200,00	0,96	0,92	1152,00	1869,19
2	2400,00	0,77	0,59	1848,00	1873,00
3	2480,00	0,77	0,59	1909,60	1873,00
4	1600,00	0,89	0,79	1424,00	1870,60
5	1750,00	0,82	0,67	1435,00	1872,00
6	1800,00	0,81	0,66	1458,00	1872,20
<i>Итого</i>	11230,00	5,02	4,23	9226,60	11230,00

Решение

$$\begin{cases} a_0 = \frac{11230 - 5,02a_1}{6}, \\ 5,02a_0 + 4,23a_1 = 9226,6. \end{cases}$$

$$a_0 = 1888,46, \quad a_1 = -20,1.$$

$$y_x = 1888,46 - 20,1x.$$

На основе уравнения регрессии можно сделать вывод, что с ростом затрат на единицу продукции на 1 руб. прибыль предприятия снижается на 20,1 тыс. руб.

6.3. Множественная (многофакторная) регрессия

Если определяется зависимость между двумя и более признаками, то говорят о наличии множественной регрессии.

$$y_x = a_0 + a_1x_1 + a_2x_2 + a_3x_3.$$

Этапы исследования множественной регрессии:

- выбор формы связи;
- отбор факторных признаков;
- построение уравнения.

Теоретически уравнение множественной регрессии может быть представлено любой формой связи. Но одновременное влияние нескольких факторов, как правило, затушевывает конкретный вид зависимости.

Как правило, все виды нелинейных уравнений сводятся к линейным.

Количество факторов, включенных в модель, с одной стороны, должно быть ограничено. С другой стороны, оно должно быть достаточно, чтобы отразить влияние всех факторов.

В модель необходимо включать независимые или взаимосвязанные факторные признаки при условии, что их влияние на результат признака будет больше их взаимного влияния друг на друга.

Из двух взаимосвязанных факторных признаков из модели исключается тот, чье влияние на результирующий признак будет меньше.

Система нормальных уравнений для трехфакторной модели:

$$\begin{cases} na_0 + a_1 \sum x_1 + a_2 \sum x_2 = \sum y \quad , \\ a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum x_1 y \quad . \\ a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum x_2 y \quad . \end{cases}$$

Коэффициенты регрессии a_0, a_1, a_2 характеризуют влияние конкретных факторов: a_1 – показывает, на сколько единиц измерений изменяется y при изменении x на свою единицу измерения.

6.4. Параметрические методы изучения связи

Для определения тесноты связей используются следующие показатели: линейный коэффициент корреляции, эмпирическое и теоретическое корреляционное отношения, множественный коэффициент корреляции, частные коэффициенты корреляции.

Линейный коэффициент корреляции характеризует тесноту и направление связи между **двумя признаками** при наличии между ними **линейной** зависимости.

Линейный коэффициент корреляции обозначается r и рассчитывается:

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

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\left(n \sum x^2 - (\sum x)^2 \right) \left(n \sum y^2 - (\sum y)^2 \right)}}$$

x, y находят из наблюдений, n – количество наблюдений.
 r изменяется от -1 до +1, то есть $-1 \leq r \leq 1$.

При $r=0$ связь отсутствует, при $r=1$ связь функциональная (нет различных случайных факторов), при $r \in [-1; 0)$ связь обратная, при $r \in (0; 1]$ – прямая.

Между линейным коэффициентом корреляции и коэффициентом регрессии существует определенная зависимость

$$r_{yx_i} = \frac{a_i \sigma_{x_i}}{\sigma_y}$$

Эмпирическое отношение: $\eta_{эмп.} = \sqrt{\frac{\delta^2}{\sigma^2}}$,

где σ^2 – общая величина дисперсии эмпирических значений признака; δ^2 – межгрупповая дисперсия, которая характеризует формирование группового признака.

Теоретическое отношение: $\eta_{теор.} = \sqrt{\frac{\delta_{теор.}^2}{\sigma^2}}$.

Корреляционное отношение изменяется от 0 до 1.

Множественный коэффициент корреляции R рассчитывается при наличии связи между результативным признаком и несколькими факторными. Этот коэффициент позволяет оценить тесноту связи факторов, включенных в модель, и результативного признака.

R изменяется от 0 до 1, при этом чем ближе значение к 1, тем больше взаимосвязь.

$$R_{y/x_1 x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} r_{yx_2} r_{x_1 x_2}}{(1 - r_{x_1 x_2}^2)}}$$

Частные коэффициенты корреляции характеризуют тесноту связи при фиксированном значении других факторных признаков, то есть связь в чистом виде.

Частные коэффициенты корреляции для трехфакторной модели имеют следующий вид:

$$r_{y/x_1x_2} = \frac{r_{yx_1} + r_{yx_2} r_{x_1x_2}}{\sqrt{\left(1 - r_{x_2y}^2\right) \left(1 - r_{x_1x_2}^2\right)}},$$

$$r_{yx_2/x_1} = \frac{r_{yx_2} + r_{yx_1} r_{x_1x_2}}{\sqrt{\left(1 - r_{x_1y}^2\right) \left(1 - r_{x_1x_2}^2\right)}}.$$

Частные коэффициенты изменяются от -1 до $+1$.

6.5. Принятие решений на основе уравнения регрессии

Полученные регрессионные модели позволяют использовать их для принятия обоснованного управленческого решения, но при этом следует иметь в виду, что полученная закономерность верна только для тех условий, для которых она была рассчитана, или для подобных условий.

С целью расширения возможности анализа **используют коэффициенты эластичности**, которые показывают, на сколько процентов в среднем изменяется значение результативного признака при изменении факторного на 1 %.

В уравнении множественной регрессии используют частные коэффициенты детерминации.

Коэффициент эластичности связан коэффициентом регрессии: $d_{x_i} = r_{yx_i} \beta_{x_i}$, где r_{yx_i} - парный коэффициент корреляции между результативным и факторным признаком; β_{x_i} - стандартизованный коэффициент детерминации.

Стандартизованные коэффициенты детерминации связаны с коэффициентом регрессии: $\beta_{x_i} = a_i \frac{\sigma_{x_i}}{\sigma_y}$.

Стандартизованные коэффициенты детерминации показывают, на сколько средних квадратических отклонений изменяется результативный признак при изменении факторного признака на одно среднеквадратическое отклонение.

Коэффициенты эластичности и стандартизованные коэффициенты позволяют сопоставить варианты достижения одного и того же результата.

Контрольные вопросы к теме 6

- 1. Сравните характеристики функциональной и корреляционной связей.*
- 2. Перечислите показатели тесноты связи между признаками в парной корреляции.*
- 3. Приведите показатели тесноты связи между признаками во множественной корреляции.*
- 4. О чем говорят коэффициенты регрессии в уравнении регрессии?*
- 5. Что характеризуют коэффициенты детерминации?*

7. ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

- *Характеристика выборочного наблюдения* • *Ошибки выборки*
- *Определение необходимой численности выборки* • *Малая выборка*

7.1. Характеристика выборочного наблюдения

Выборочное наблюдение представляет собой один из методов сплошного наблюдения, характеризуется тем, что отобранная в случайном порядке часть единиц дает представление о всей изучаемой совокупности по какому-либо признаку.

При проведении выборочного наблюдения необходимо обеспечить случайность выборки и её репрезентативность (представительность).

Принцип случайности означает, что все единицы совокупности должны иметь равные шансы попасть в выборку.

Для обеспечения принципа репрезентативности необходимо, чтобы выборка была достаточной. В этом случае ошибка выборки не превысит допустимые размеры.

Преимущества выборочного наблюдения:

1. Экономия времени, материальных и денежных ресурсов.
2. Возможность проведения там, где невозможно проведение сплошного наблюдения (например при проверке качества продукции, связанного с её уничтожением, – вскрытие консервов, проверка лампочек).

Недостаток выборочного наблюдения – возникновение ошибок выборки.

Совокупность, из которой осуществляется выборка, называется генеральной. Отобранная часть представляет выборочную совокупность, или выборку.

Ошибками выборки называются расхождения между характеристиками генеральной и выборочной совокупностей. Причина

появления ошибок выборки – отличие структуры выборки от структуры генеральной совокупности.

Для определения и обозначения основных характеристик генеральной и выборочной совокупностей используют следующие условные обозначения (табл. 15).

Таблица 15

Условные обозначения основных характеристик генеральной и выборочной совокупностей

Показатель	Генеральная совокупность	Выборочная совокупность
Численность выборки	N	n
Среднее значение признака	\bar{x}	\tilde{x}
Общая дисперсия	σ^2	σ_v^2
Межгрупповая дисперсия	δ^2	δ_v^2
Доля единиц, обладающих признаком	p	ω
Доля единиц, не обладающих признаком	$1-p$	$1-\omega$
Частота	f	f_v

7.2. Ошибки выборки

Различают два вида ошибок выборки:

- стандартная, или средняя;
- предельная.

Под **средней ошибкой выборки** понимают такое расхождение между средней генеральной совокупности и средней выборочной совокупности, которое не превышает $\pm\sigma$ (среднего квадратического отклонения).

$$-\sigma \leq \tilde{x} - \bar{x} \leq +\sigma .$$

Предельной ошибкой выборки считают максимально возможное расхождение между средней генеральной и средней выборочной совокупности при заданной вероятности её появления.

В основе определения ошибок выборки лежит закон нормального распределения. Формула средней ошибки выборки зависит от метода (способа) проведения выборочного наблюдения (собственно случайный отбор, серийный, механический отбор), объема выборки и вариации признака.

Для **собственно случайного повторного отбора** стандартная, или средняя, ошибка выборки определяется по формуле

$$\mu = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}},$$

где n – численность выборки.

Для **собственно случайного бесповторного отбора** ошибка выборки определяется по формуле $\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$.

Организовать собственно случайный повторный отбор сложнее, чем собственно случайный бесповторный отбор. Так как рассчитывать ошибку легче по формуле случайного повторного отбора, а организовывать выборку удобнее как случайную бесповторную, то **на практике используют случайный бесповторный отбор, а ошибку выборки рассчитывают как при повторном отборе, несколько завышая её величину.**

Предельная ошибка выборки рассчитывается по формуле

$$\Delta = \mu t,$$

где t – коэффициент доверия, который определяется по таблице нормального распределения.

Предельная ошибка выборки используется при определении доверительного интервала, который выглядит так:

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

Чем выше вероятность, с которой гарантируется попадание в доверительный интервал, тем больше величина доверительного интервала.

Наряду с абсолютной величиной рассчитывается **относительная величина ошибки выборки**, которая в общем случае опре-

деляется по формуле $\Delta\% = \frac{\Delta}{\tilde{x}} 100\%$. Для альтернативного признака ошибки определяются по следующим формулам:

$$\mu = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}},$$

$$\Delta = \mu t,$$

$$\Delta\% = \frac{\Delta}{\omega} 100\%.$$

В статистике доказано, что общая величина дисперсии генеральной совокупности связана с дисперсией выборки следующим соотношением: $\sigma^2 = \sigma_e^2 \frac{n}{n-1}$.

При большом объеме выборки ($N > 30$) $\frac{n}{n-1}$ стремится к единице и $\sigma^2 \approx \sigma_e^2$, поэтому среднюю ошибку выборки можно рассчитывать на основании выборочной дисперсии по формуле

$$\mu = \sqrt{\frac{\sigma_e^2}{n}} = \frac{\sigma_e}{\sqrt{n}}.$$

Пример

При проверке веса импортируемого груза на таможенную методом случайной повторной выборки было отобрано 200 изделий. В результате был установлен средний вес изделия 30 г при среднем квадратическом отклонении 4 г. С вероятностью 99,7 % определить доверительный интервал, в котором находится средний вес изделия всего импортируемого груза.

Решение

$$\sigma = 4, V = 200, \mu = \frac{4}{\sqrt{200}}, \Delta = \mu t = \frac{4}{\sqrt{200}} \cdot 3 = 0,84.$$

Определим интервал, в котором находится средний вес по всему грузу: $30 - 0,84 \leq \bar{x} \leq 30 + 0,84$.

$$29,16 \leq \bar{x} \leq 30,84.$$

Ответ. С вероятностью 99,7 % можно гарантировать, что вес изделия импортируемого груза будет находиться в интервале от 29,16 до 30,84 г.

7.3. Определение необходимой численности выборки

Необходимый объем выборки можно вывести исходя из формул предельной и средней ошибок выборки:

$$\mu = \sqrt{\frac{\sigma^2}{n}} \Rightarrow n = \frac{\sigma^2}{\mu^2}, \quad \Delta = \mu t \Rightarrow \mu = \frac{\Delta}{t}, \quad \text{отсюда } n = \frac{\sigma^2 t^2}{\Delta^2},$$

где n – необходимая численность выборки.

Необходимая численность выборки n прямо пропорциональна вариации признака, вероятности, с которой гарантируется результат, и обратно пропорциональна предельной ошибке выборки.

Таким образом, для определения необходимого объема выборки следует заранее знать как вероятность, так и дисперсию. Но последнюю характеристику можно рассчитать только по результатам выборочного наблюдения. Получается замкнутый круг, когда объем выборки нельзя рассчитать без дисперсии, и наоборот. Чтобы решить эту проблему, предварительно проводят пробные наблюдения. Полученные результаты по выборочной дисперсии используют для расчета необходимой численности выборки. При этом дисперсию стараются несколько завысить. Чем выше σ^2 , тем больше объем выборки, следовательно, уменьшается ошибка выборки.

Численность выборки альтернативного признака определяется на основе формул

$$\sigma^2 \approx \sigma_g^2 \Rightarrow \mu = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\omega(1-\omega)}{n}} \Rightarrow n = \frac{\omega(1-\omega)}{\mu^2},$$

$$\Delta = \mu t \Rightarrow \mu = \frac{\Delta}{t}, \quad \text{отсюда } n = \frac{\omega(1-\omega)t^2}{\Delta^2},$$

где ω – доля единиц выборки, обладающих данным признаком;

$1 - \omega$ – доля единиц выборки, не обладающих данным признаком.

Для альтернативного признака используют максимальную величину дисперсии. Известно, что произведение $\omega(1 - \omega)$ будет максимально при $\omega = 1 - \omega = 0,5$, тогда формула для определения численности выборки будет следующей: $n = \frac{0,25t^2}{\Delta^2}$.

Для каждого метода выборочного наблюдения (собственно случайный отбор, серийный отбор, механическая или типологическая выборка) необходимый объем выборки рассчитывается на основе своей формулы ошибки выборки.⁵

7.4. Малая выборка

В ряде случаев, связанных с проверкой качества продукции, проведением таможенного контроля и другими, невозможно проводить отбор единиц в большом объеме, поэтому используют малую выборку, то есть выборку с объемом, не превышающим 30 единиц. В этом случае необходимо учитывать влияние коэффициента $\frac{n}{n-1}$. Поэтому $\sigma^2 = \sigma_e^2 \frac{n}{n-1}$, соответственно средняя ошибка малой выборки рассчитывается по формуле

$$\mu_{me} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{\sigma_e^2 \cdot \frac{n}{n-1}}{n}} = \sqrt{\frac{\sigma_e^2}{n-1}}.$$

Предельную ошибку малой выборки определяют по формуле

$$\Delta = \mu t_{cm.} = \sqrt{\frac{\sigma_e^2}{n-1}} t_{cm.},$$

где $t_{cm.}$ – коэффициент Стьюдента.

⁵ С формулами расчета ошибок выборки и необходимой численности выборки для других способов отбора можно ознакомиться в [1, 2, 3, 8].

Если объем выборки больше 30 единиц или равен ему, то вероятность ошибок выборки подчинена закону нормального распределения. Для малой выборки вероятность ошибок выборки определяется по закону распределения Стьюдента (табл. 16).

Таблица 16

Вероятность распределения ошибок по закону нормального распределения и распределению Стьюдента

t	По закону Стьюдента	По закону нормального распределения
$t=1$	$p(t)=0,67=67\%$	$p(t)=0,683=68,3\%$
$t=2$	$p(t)=0,94=94\%$	$p(t)=0,954=95,4\%$
$t=3$	$p(t)=0,992=99,2\%$	$p(t)=0,997=99,7\%$

Контрольные вопросы к теме 7

1. *Приведите определение выборочного наблюдения.*
2. *Какие методы организации выборочного наблюдения Вы знаете?*
3. *Назовите виды ошибок выборки.*
4. *Какая выборка называется малой выборкой?*
5. *От каких факторов зависит объем необходимой численности выборки?*
6. *Как решить проблему взаимосвязи расчета дисперсии и необходимого объема выборки?*

8. РЯДЫ ДИНАМИКИ

• Виды рядов динамики и их характеристика • Условия сопоставимости рядов динамики • Расчет среднего уровня ряда динамики • Показатели рядов динамики • Смыкание рядов динамики

8.1. Виды рядов динамики и их характеристика

Ряд динамики представляют собой ряд данных, характеризующих социально-экономическое явление во времени.

Основными характеристиками рядов динамики являются уровень и время. **Уровень** ряда динамики выражается абсолютными, относительными или средними величинами. Время может быть представлено моментными или интервальными показателями.

В зависимости от представления временной характеристики ряды динамики делят:

- на **моментные**, характеризуют состояние на момент, определенную дату. Моментные ряды динамики характеризуются результатами проверок на всех уровнях:

 предприятия – остатков запасов на складе, средств на счетах предприятия, балансовых показателей;

 города – наличия жилого фонда, посадочных мест в библиотеке;

 страны – населения по результатам переписи;

- **интервальные** ряды динамики представлены показателями, характеризующими результаты деятельности (объема, доходов, затрат) и средними показателями (средней заработной платы, выработки, затрат на единицу продукции).

Необходимо помнить, что уровни интервальных рядов суммировать можно, а моментных рядов динамики нельзя.

8.2. Условия сопоставимости рядов динамики

Основные условия работы с рядами динамики – сопоставимость и правильный выбор начального уровня ряда динамики. По умолчанию начальным уровнем принимается первый уровень ряда динамики.

Ряды динамики должны быть сопоставлены по единицам измерения; масштабу, территории; структуре предприятия и продукции; методике исчисления показателей; уровню цен.

8.3. Расчет среднего уровня ряда динамики

Средний уровень интервального ряда рассчитывается по средней арифметической простой для равных интервалов и по средней арифметической взвешенной для неравных.

Средний уровень моментного ряда определяется по средней хронологической для равных интервалов или путем перехода от моментного ряда к интервальному с последующим расчетом по средней арифметической взвешенной для неравных интервалов.

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

Пример 8.1

Определить среднегодовые остатки денежных средств (табл. 17).

Таблица 17

Средние остатки на депозитах

Средние остатки по депозитам, млрд руб.	01.01.07	01.04.07	01.07.07	01.10.07	01.01.08
	1,5	4,0	3,2	6,0	8,6

Решение

$$\bar{x}_{xp.} = \frac{\frac{1}{2} \cdot 1,5 + 4 + 3,2 + 6 + \frac{1}{2} \cdot 8,6}{4} = 4,6 \text{ млрд руб.}$$

Пример 8.2

По данным табл. 18 рассчитать среднегодовые остатки денежных средств.

Таблица 18

Оборотные средства предприятия

Дата	Оборотные средства предприятия на отчетную дату, тыс. руб.	Расчетные графы	
		Длительность периода, мес	Средняя величина оборотных средств за период, тыс. руб.
На 01.01.00	17000	1	17500
На 01.02.00	18000		
На 01.05.00	12000	3	15000
На 01.09.00	6000	4	9000
На 01.11.00	20000	2	13000
На 01.01.01	14000	2	17000

Решение

$$\bar{x}_{\text{взвеш.}} = \frac{17500 \cdot 1 + 15000 \cdot 3 + 9000 \cdot 4 + 13000 \cdot 2 + 17000 \cdot 2}{12} =$$
$$= 13200 \text{ тыс. руб.}$$

8.4. Показатели рядов динамики

К показателям рядов динамики относят абсолютный прирост, темп роста, темп прироста, абсолютное значение одного процента прироста. Все показатели, за исключением абсолютного значения одного процента прироста, могут быть цепными, базисными, средними. Формулы расчета всех показателей представлены в табл. 19.

Цепные показатели характеризуют данный уровень по отношению к предыдущему.

Базисные показатели сравнивают данный уровень с уровнем, принятым за базу сравнения. Как правило, в качестве базисного принимают первый уровень ряда динамики.

Таблица 19

Формулы расчета основных показателей рядов динамики

Показатель	Вид показателей		
	Цепные	Базисные	Средние
Абсолютный прирост, тыс. руб. ⁶	$\Delta x_i^y = x_i - x_{i-1}$	$\Delta x_i^{\bar{b}} = x_i - x_{\bar{b}}$	$\overline{\Delta x} = \frac{\sum x_i^y}{n-1}$ или $\overline{\Delta x} = \frac{x_n^{\bar{b}}}{n-1}$
Коэффициент роста	$T_i^y = \frac{x_i}{x_{i-1}}$	$T_i^{\bar{b}} = \frac{x_i}{x_{\bar{b}}}$	$\bar{T} = \sqrt[n-1]{T_{2/1}^y \cdot \dots \cdot T_{n/n-1}^y}$ или $\bar{T} = \sqrt[n-1]{T_n^{\bar{b}}}$
Темп роста, %	$T_{i\%}^y = \frac{x_i}{x_{i-1}} 100\%$	$T_{i\%}^{\bar{b}} = \frac{x_i}{x_{\bar{b}}} 100\%$	$\bar{T}_{\%} = \sqrt[n-1]{T_n^{\bar{b}}} 100\%$
Темп прироста: – в коэффициентах – в процентах	$\Delta T_i^y = T_i^y - 1$ $\Delta T_{i\%}^y = T_{i\%}^y - 100\%$	$\Delta T_i^{\bar{b}} = T_i^{\bar{b}} - 1$ $\Delta T_{i\%}^{\bar{b}} = T_{i\%}^{\bar{b}} - 100\%$	$\overline{\Delta T} = \bar{T} - 1$ $\overline{\Delta T}_{i\%} = \bar{T}_{i\%} - 100\%$
Абсолютное значение одного процента прироста, тыс. руб.	$A3\Pi_i = \frac{x_{i-1}}{100} = \frac{\Delta x_i^y}{\Delta T_i^y}$	–	–

⁶ Сумма цепных абсолютных приростов равна базисному абсолютному приросту последнего периода $\sum \Delta x_i^y = x_n^{\bar{b}}$.

Пример расчета показателей ряда динамики представлен в табл. 20.

Пример 8.3

Таблица 20

Пример расчета основных показателей ряда динамики

Показатель	Год			Средние показатели
	2006	2007	2008	
Объем работ, тыс. руб.	200	210	240	$\bar{x} = 216,7$
Абсолютный прирост, тыс. руб.:				
– <i>цепной</i>	–	+10	+30	$\overline{\Delta x} = \frac{10+30}{3-1} = 20$
– <i>базисный</i>	0	+10	+40	$\overline{\Delta x} = \frac{40}{3-1} = 20$
Темп роста:				
– <i>цепной</i> :	–	$\frac{210}{200} = 1,05$	$\frac{240}{210} = 1,14$	$\bar{T} = \sqrt{1,05 \cdot 1,14} = 1,095$
– коэффициент				$\bar{T}_{\%} =$
– в процентах	-	$\frac{210}{200} 100 \% = 105 \%$	$\frac{240}{210} 100 \% = 114 \%$	$= \sqrt{1,05 \cdot 1,14} \cdot 100 \% = 109,5 \%$
– <i>базисный</i> :				
– коэффициент	1	$\frac{210}{200} = 1,05$	$\frac{240}{200} = 1,2$	$\bar{T} = \sqrt{T_2^{\bar{\delta}}} = \sqrt{1,2} = 1,095$
– в процентах				$\bar{T}_{\%} = 1,095 \cdot 100 = 109,5 \%$
Темп прироста:				
– <i>цепной</i>	–	$1,05 - 1 = 0,05$	$1,14 - 1 = 0,14$	$\overline{\Delta T} = 1,095 - 1 = 0,095$
– <i>базисный</i>	0	$1,05 - 1 = 0,05$	$1,2 - 1 = 0,2$	$\overline{\Delta T}_{\%} = 0,095 \cdot 100 = 9,5 \%$
Абсолютное значение одного процента прироста, тыс. руб.	–	$\frac{200}{100} = 2$	$\frac{210}{100} = 2,1$	–

Все показатели рассчитываются одинаково для моментных и интервальных рядов динамики.

8.5. Смыкание рядов динамики

Смыкание рядов динамики – это процесс приведения рядов динамики к сопоставимому виду по методике исчисления, ценам, структуре, единицам измерения, базисному уровню и т.д.

Пример 8.4

По приведенным в табл. 21 данным составить ряд динамики базисных темпов роста выработки к 2002 г.

Решение

$$T_v = \frac{T_Q}{T_{\bar{C}}}, \text{ где } T_Q \text{ – темп роста объема; } T_{\bar{C}} \text{ – темп роста среднесписочной численности работников.}$$

Расчеты представлены в табл. 21.

Таблица 21

Темпы роста объемов и численности работников

Показатель, %	Год					
	2002	2003	2004	2005	2006	2007
Темп роста объема базисный к 2002 г.	100	103	104	107	110	115
Темп роста среднесписочной численности работников цепной	101	102	100	103	101	102
<i>Решение</i>						
Темп роста среднесписочной численности работников базисный к 2002 г.	100	101	99	102	100	101
Темп роста выработки базисный к 2002 г.	100	102	105	105	110	114

Пример 8.5

По данным табл. 22 рассчитать темпы роста выработки по годам.

Таблица 22

Исходные данные

Показатель	2002 г.	2003 г.	2004 г.	2005 г.	2006 г.	2007 г.
Выработка, тыс. шт.	150	200	220	280	300	310
Численность работников по списку на 01.01. чел.	30	40	30	50	32	48
<i>Решение</i>						
Среднесписочная численность работников $T_{\bar{c}}$	$(30+40)/2=35$	$(40+30)/2=35$	40	41	40	—
Выработка на одного работника $B=Q/T$	$150/35=4,29$	$200/35=5,71$	5,5	6,8	7,5	—

Пример 8.6

По данным табл. 23 составить единый ряд динамики объемов работ предприятия.

Таблица 23

Объемы работ

Показатель	2000 г.	2001 г.	2002 г.	2003 г.	2004 г.	2005 г.
Объем работ, млн руб., по методике 2000 г.	15,0	20,0	22,0	28,0	—	—
Объем работ, млн руб., по методике 2003 г.	—	—	—	30,0	42,0	44,0
<i>Решение</i>						
Коэффициент пересчета объемов 2003 г. по отношению к 2000 г.	$30/28=1,07$					
Объем работ, млн руб., по методике 2003 г.	$15 \cdot 1,07=16,1$	21,4	23,5	30,0	42,0	44,0

Контрольные вопросы к теме 8

- 1. Какие ряды динамики относятся к моментным?*
- 2. Какие ряды динамики называются интервальными?*
- 3. Как рассчитать средний уровень моментного ряда для равных и неравных интервалов?*
- 4. Как определить средний уровень интервального ряда для равных и неравных интервалов?*
- 5. Перечислите показатели рядов динамики.*

9. ЭКОНОМИЧЕСКИЕ ИНДЕКСЫ

- *Понятие и виды экономических индексов* • *Агрегатные индексы*
- *Средневзвешенные, или средние из индивидуальных индексов*
 - *Средние индексы* • *Индексные методы анализа*

9.1. Понятие и виды экономических индексов

Индекс – это относительная величина. Все индексы можно разделить на две группы: динамические и территориальные.

Территориальные индексы представлены относительными величинами сравнения, **динамические** индексы – относительными величинами динамики или темпами роста.

Классификация индексов представлена на рис. 6.

Индексы позволяют выявлять влияние отдельных составляющих на результат. Область применения индексов – функциональные зависимости. Величина влияния факторов зависит от порядка расчета индексов, в чем проявляется недостаток индексного метода анализа.

Правило построения индексов формулируется так: «Количество изменяется при старом качестве, а качество - при новом количестве».

9.2. Агрегатные индексы

Агрегатные индексы позволяют определять изменения валовых показателей в целом по предприятию, структурным подразделениям, например выручки, затрат, фонда заработной платы.

$$\text{Общий агрегатный индекс } I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}.$$

Если I_{pq} – индекс выручки, тогда p – цена единицы товаров, q – объем реализации данного вида товаров.

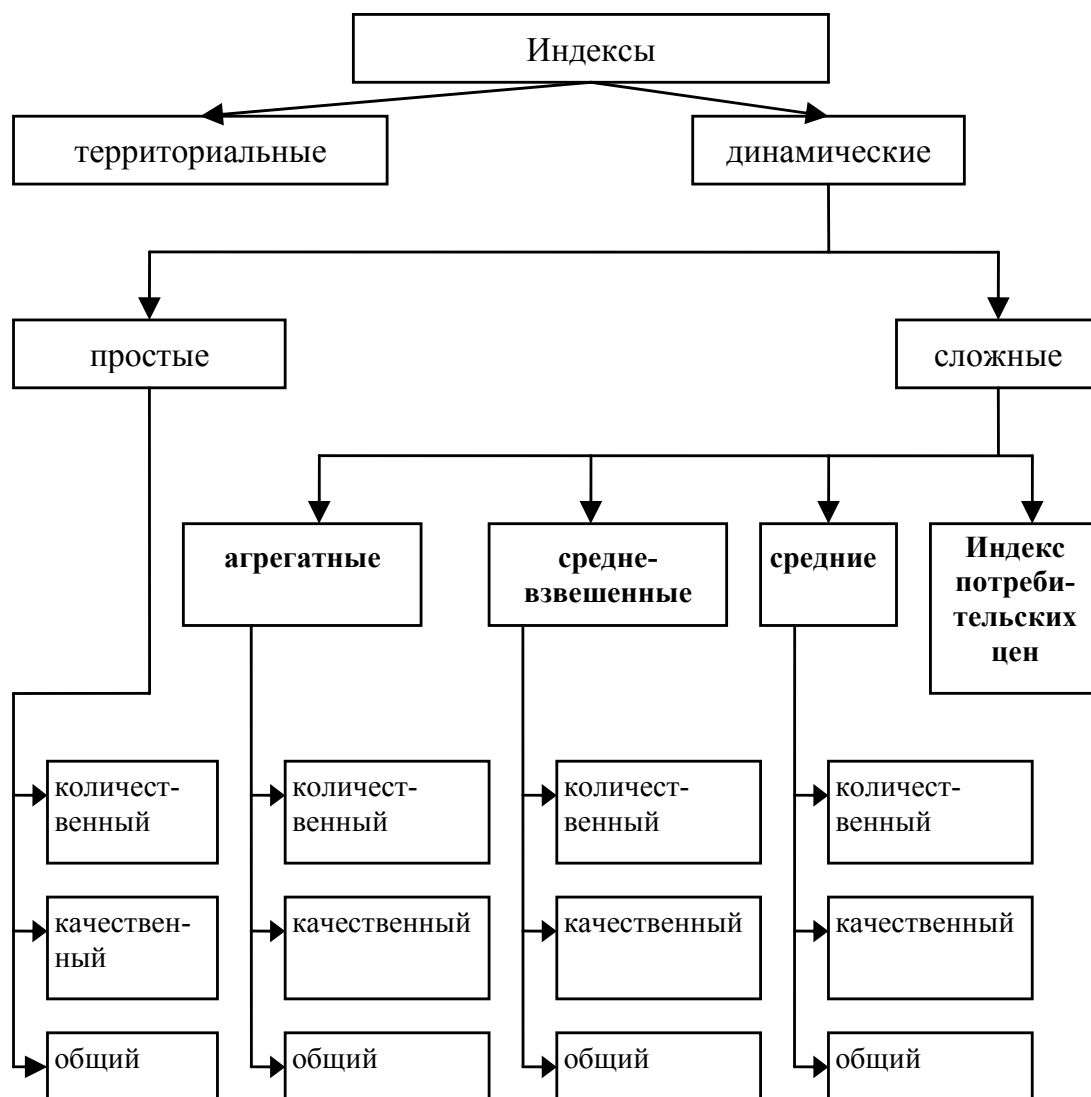


Рис. 6. Классификация индексов

Индекс физического объема $I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$ показывает, как

изменяется выручка при изменении объема в базисных ценах.

Индекс цен $I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$ отражает влияние цен на изменение

выручки при закреплении объемов на отчетном уровне.

Имеет место взаимосвязь индексов: $I_{pq} = I_p I_q$. Для определения абсолютного прироста рассчитывается разность числителя и знаменателя соответствующего индекса. При этом сумма абсо-

лютных изменений выручки за счет влияния отдельных факторов составляет общее изменение выручки: $\Delta B_p + \Delta B_q = \Delta B_{pq}$.

Пример 9.1

По данным табл. 24 определить изменение выручки предприятия в целом, в том числе за счет изменения цен на автомобили и за счет изменения объема реализации.

Таблица 24

Данные о продажах

Вид товара	Отчетный период (2004 г.)		Базисный период (2003 г.)	
	Цена, тыс. руб.	Объем реализации, шт.	Цена, тыс. руб.	Объем реализации, шт.
ВАЗ-2115	200	37	150	20
ВАЗ-2112	218	40	190	24

Решение

$$I_{pq} = \frac{200 \cdot 37 + 218 \cdot 40}{150 \cdot 20 + 190 \cdot 24} = \frac{16120}{7560} = 2,13.$$

Определим, какое влияние на изменение выручки оказали цены и объем реализации:

$$I_q = \frac{150 \cdot 37 + 190 \cdot 40}{150 \cdot 20 + 190 \cdot 24} = \frac{13150}{7560} = 1,74,$$

$$I_p = \frac{200 \cdot 37 + 218 \cdot 40}{150 \cdot 37 + 190 \cdot 40} = \frac{16120}{13150} = 1,23.$$

Абсолютное изменение выручки за счет:

- изменения количества реализованной продукции: $\Delta B_q = 13150 - 7560 = 5590$ тыс. руб.;

- изменения цен: $\Delta B_p = 16120 - 13150 = 2970$ тыс. руб.;

- $\Delta B_p = 16120 - 13150 = 2970$ тыс. руб.

В целом выручка за отчетный период (за 2004 г.) выросла на 8560 тыс. руб. (на 113 % по сравнению с выручкой базисного периода), в том числе за счет изменения объемов реализации – на 5590 тыс. руб. (т.е. влияние объемов реализации обуславливает 74 % общего изменения выручки), за счет роста цен – на 2870 тыс. руб. (т.е. влияние цен обуславливает 23 % общего изменения выручки).

9.3. Средневзвешенные, или средние из индивидуальных индексов

Средневзвешенные индексы – это сложные индексы, так как они учитывают динамику по группам товаров.

Данный вид индексов выводится из агрегатных и используется в том случае, если неизвестны абсолютные значения объемов или цен, но есть информация по динамике изменения этих показателей.

$$\text{Общий индекс } \bar{I}_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0}.$$

Количественный индекс представлен средним арифметическим индексом физического объема и выводится из количественного агрегатного: $\bar{I}_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$. Из равенства $i_q = \frac{q_1}{q_0}$ выразим $q_1 = i_q q_0$ и, подставив это выражение в формулу количественного индекса, получим

$$\bar{I}_q = \frac{\sum p_0 q_0 i_q}{\sum p_0 q_0}.$$

Качественный индекс представлен средним гармоническим индексом цен: $\bar{I}_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$.

Из равенства $i_p = \frac{p_1}{p_0}$ выразим $p_0 = \frac{p_1}{i_p}$. Подставив это выражение в формулу качественного индекса, получим

$$\bar{I}_p = \frac{\sum p_1 q_1}{\sum p_1 q_1 \frac{1}{i_p}} .$$

Для средних агрегатных индексов справедливо соотношение $\bar{I}_{pq} = \bar{I}_p \bar{I}_q$.

Абсолютное значение влияния каждого фактора рассчитывается как разность числителя и знаменателя соответствующего индекса (как и у агрегатных индексов).

Пример 9.2

Выручка по предприятию в 2007 г. по отношению к 2006 г. снизилась на 2 % и составила 300 тыс. руб. Цена продукции А выросла на 5 %, цена продукции В – на 7 %. Определить изменение выручки за счет изменения объема выпуска продукции А и В. Доля продукции А в выручке 2007 г. составила 30 %.

Решение

$$\bar{I}_p = \frac{300}{90 \cdot \frac{1}{1,05} + 210 \cdot \frac{1}{1,07}} = 1,06 . \text{ За счет изменения цен вы-}$$

ручка выросла на 6 %. $\bar{I}_q = 0,98 / 1,06 = 0,92$. За счет изменения объема продукции выручка снизилась на 8 %.

9.4. Средние индексы

Средние индексы используются для определения изменения средних величин (средней заработной платы, цены реализации, издержек на единицу продукции) по однородной продукции.

Средние индексы включают индексы переменного состава, структурных сдвигов, постоянного состава.

Индекс переменного состава

$$\bar{I}_{n.c.} = \frac{\sum p_1 q_1}{\sum q_1} / \frac{\sum p_0 q_0}{\sum q_0} .$$

Индекс структурных сдвигов

$$\bar{I}_{стр.сд.} = \frac{\sum p_0 q_1}{\sum q_1} / \frac{\sum p_0 q_0}{\sum q_0}.$$

Индекс фиксированного состава

$$\bar{I}_{ф.с.} = \frac{\sum p_1 q_1}{\sum q_1} / \frac{\sum p_0 q_1}{\sum q_1}.$$

$$\bar{I}_{н.с.} \cdot \bar{I}_{стр.сд.} = \bar{I}_{ф.с.}$$

Для определения абсолютного значения показателя нужно из числителя вычесть знаменатель.

Пример 9.3

Два завода одной фирмы выпускают однородную продукцию (табл. 25). Определить изменение средних затрат на производство единицы продукции, в том числе за счет изменения себестоимости единиц продукции каждого завода, и изменения структуры производства продукции на фирме.

Таблица 25

Затраты на производство продукции

Завод	2007 г.		2006 г.	
	Затраты на однородную продукцию p_1 , руб.	Объем производства q_1 , тыс. шт.	Затраты на однородную продукцию p_0 , руб.	Объем производства q_0 , тыс. шт.
1-й	20	100	24	70
2-й	28	300	30	150

Решение

$$\bar{I}_{н.с.} = \frac{20 \cdot 100 + 28 \cdot 300}{400} / \frac{24 \cdot 70 + 30 \cdot 150}{220} = 26 / 28,1 = 0,93.$$

Средние затраты на производство продукции в 2007 г. снизились на 7 %.

$$\bar{I}_{ф.с.} = \frac{20 \cdot 100 + 28 \cdot 300}{400} / \frac{24 \cdot 100 + 30 \cdot 300}{400} = 26 / 28,5 = 0,91.$$

За счет изменения себестоимости затраты на единицу продукции снизились на 9 % по предприятию.

$$\bar{I}_{стр.сд.} = \frac{24 \cdot 100 + 30 \cdot 300}{400} / \frac{24 \cdot 70 + 30 \cdot 150}{220} = 28,5 / 28,1 = 1,01.$$

За счет изменения структуры производства затраты выросли на 1 %.

В целом по предприятию затраты на производство единиц продукции снизились на 2,1 руб., или 7 %, в том числе за счет изменения себестоимости снизились на 2,5 руб., или 9 %, и за счет изменения структуры выросли на 0,4 руб., или 1 %.

Все предыдущие индексы (агрегатные, средневзвешенные и средние) взаимосвязаны, то есть зная два индекса, можно найти третий.

Индекс потребительских цен используется для оценки уровня инфляции, он не подчиняется правилу построения индексов.

Если для предыдущих индексов цен используется формула индекса Пааше $I_{ц} = \sum p_1 q_1 / \sum p_0 q_1$, то для индекса потребительских цен применяется формула индекса Ласпейреса $I_{н.ц.} = \sum p_1 q_0 / \sum p_0 q_0$. Достоинством данного индекса является возможность рассчитывать базисные индексы на основе произведения цепных индексов:

$$\sum p_1 q_0 / \sum p_0 q_0 = (\sum p_2 q_0 / \sum p_1 q_0) (\sum p_3 q_0 / \sum p_2 q_0) (\sum p_4 q_0 / \sum p_3 q_0) = \sum p_4 q_0 / \sum p_1 q_0.$$

Это позволяет отслеживать динамику цен, одной и той же потребительской корзины. Недостатком данного индекса является нарушение взаимосвязей индексов: $I_{pq} \neq I_p I_q$.

9.5. Индексные методы анализа

Какая зависимость существует между признаками, связанными между собой знаками «×» или «/», такая же связь сохраняется и между индексами.

Пример 9.4.

Определить, как изменится объем производства продукции, если цены выросли на 12 %, а выручка – на 7 %.

Решение

$Выручка = Объем \cdot Цена$, отсюда $Объем = Выручка / Цена$.

Изменение объема будет равно $I_Q = 1,07 / 1,12 = 0,95$. Объем производства продукции снизился на 5 %.

К индексным методам анализа относят метод разности и метод цепных подстановок.

Метод цепных подстановок используют в том случае, если количество факторных показателей больше трёх. Заключается этот метод в последовательной замене базисных значений на отчетные на основе правила построения индекса.

Если объем выражается формулой $Q = B_{ч} t n \bar{C}$, то индекс объема равен $I_Q = B_{ч1} t_1 n_1 \bar{C}_1 / B_{ч0} t_0 n_0 \bar{C}_0$.

Анализ начинается от количественных факторов, переходя к более качественному.

$$I_{Q_{\bar{C}}} = B_{ч0} t_0 n_0 \bar{C}_1 / B_{ч0} t_0 n_0 \bar{C}_0 \text{ (влияние численности);}$$

$$I_{Q_n} = B_{ч0} t_0 n_1 \bar{C}_1 / B_{ч0} t_0 n_0 \bar{C}_1 \text{ (влияние количества рабочих дней);}$$

$$I_{Q_t} = B_{ч0} t_1 n_1 \bar{C}_1 / B_{ч0} t_0 n_1 \bar{C}_1 \text{ (влияние продолжительности рабочего дня);}$$

$$I_{Q_{B_{ч}}} = B_{ч1} t_1 n_1 \bar{C}_1 / B_{ч0} t_1 n_1 \bar{C}_1 \text{ (влияние часовой выработки).}$$

$$I_Q = I_{Q_{\bar{C}}} I_{Q_n} I_{Q_t} I_{Q_{B_{ч}}}.$$

Для определения абсолютных изменений из числителя соответствующего индекса вычитается знаменатель.

Метод разности дает абсолютное изменение показателя за счет влияния факторов. Используется для мультипликативных моделей.

$$\Delta Q_{z\bar{C}} = (\bar{C}_1 - \bar{C}_0) B_{ч0} t_0 n_0, \quad \Delta Q_{zn} = (n_1 - n_0) B_{ч0} t_0 \bar{C}_1, \quad \Delta Q_{zt} = (t_1 - t_0) B_{ч0} \bar{C}_1 n_1,$$

$$\Delta Q_{zB_{ч}} = (B_{ч1} - B_{ч0}) \bar{C}_1 n_1 t_1.$$

$$\Delta Q_{z\bar{C}} + \Delta Q_{zn} + \Delta Q_{zt} + \Delta Q_{zB_{ч}} = \Delta Q_z.$$

Контрольные вопросы к теме 9

- 1. Какие виды индексов Вы знаете?*
- 2. В чем преимущества и недостатки индексных методов анализа?*
- 3. Приведите правило построения индексов.*
- 4. Как по индексу определить абсолютное изменение показателей?*
- 5. В каком случае используются агрегатные индексы?*
- 6. Для каких целей используются средние индексы?*

Рекомендательный библиографический список

1. *Громыко, Г. Л.* Теория статистики : практикум / Г. Л. Громыко. – 3-е изд., доп. и перераб. – М. : Инфра-М, 2004. – 160 с. (Высшее образование). – ISBN 5-16-000583-8.
2. Статистика : учеб. для вузов / под ред. И. И. Елисеева. – М. : Проспект, 2005. – 444 с. – ISBN 5-482-00031-1.
3. *Ефимова, М. Р.* Практикум по общей теории статистики : учеб. пособие для вузов по специальности "Экономика" / М. Р. Ефимова, О. И. Ганченко, Е. В. Петрова. – М. : Финансы и статистика, 2001. – 278 с. – ISBN 5-279-02011-7.
4. *Ефимова, М. Р.* Общая теория статистики : учеб. для вузов / М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев. – 2-е изд., испр. и доп. – М. : Инфра-М, 2006. – 413 с. (Высшее образование). – ISBN 5-16-002179-5.
5. *Громыко, Г. Л.* Теория статистики : учеб. для вузов / Г. Л. Громыко [и др.] ; под ред. Г. А. Громыко. – 2-е изд., перераб. и доп. – М. : Инфра-М, 2006. – 475 с. (Классический университетский учебник). – ISBN 5-16-002158-2.
6. *Шмойлова, Р. А.* Теория статистики : учеб. для вузов / Р. А. Шмойлова [и др.] ; под ред. Р. А. Шмойловой. – 4-е изд., доп. и перераб. – М. : Финансы и статистика, 2006. – 655 с. – ISBN 5-279-02559-3.
7. *Она же.* Практикум по теории статистики : учеб. пособие для эконом. специальностей вузов / Р. А. Шмойлова, В. Г. Миначкин, Н. А. Садовникова ; под ред. Р. А. Шмойловой. – 2-е изд., перераб. и доп. – М. : Финансы и статистика, 2006. – 415 с. – ISBN 5-279-02558-5.
8. *Четыркин, Е. М.* Статистические методы прогнозирования / Е. М. Четыркин. – М. : Статистика, 1975. – 184 с.

Учебное издание

МАРЧЕНКО Елизавета Маратовна

СТАТИСТИКА

Учебное пособие

Ч.1. Общая теория статистики

Подписано в печать 06.05.09.
Формат 60x84/16. Усл. печ. л. 4,65. Тираж 200 экз.
Заказ
Издательство
Владимирского государственного университета.
600000, Владимир, ул. Горького, 87.