

Министерство образования Республики Беларусь

Учреждение образования  
«Полоцкий государственный университет»

О. И. Гордиенко

# СТАТИСТИКА

Учебно-методический комплекс  
для студентов специальности 1-26 02 02  
«Менеджмент»

В двух частях

Часть 1

**Общая теория статистики**

Новополоцк  
ПГУ  
2012

УДК 311(075.8)  
ББК 60.6я73  
Г68

Рекомендовано к изданию  
учебно-методической комиссией финансово-экономического факультета  
в качестве учебно-методического комплекса  
(протокол № 6 от 30.06.2010)

**РЕЦЕНЗЕНТЫ:**

магистр экономических наук,  
первый заместитель начальника Главного статистического  
управления Витебской области И. В. ХОДИКОВА;  
кандидат экономических наук, доцент, заведующая кафедрой  
бухгалтерского учета, анализа и аудита УО «ПГУ» С. Г. ВЕГЕРА

**Гордиенко, О. И.**

Г68      Статистика : учеб.-метод. комплекс для студентов специальности  
1-26 02 02 «Менеджмент». В 2 ч. Ч. 1. Общая теория статистики /  
О. И. Гордиенко. – Новополоцк : ПГУ, 2012. – 316 с.  
ISBN 978-985-531-235-3.

Представлены методология статистической науки, принципы проведения статистического исследования, а также основные методы статистического анализа: метод сводки и группировки, анализ вариации, корреляционно-регрессионный анализ, индексный анализ, анализ динамики.

Разработаны необходимые методические материалы для организации самостоятельной работы студентов: вопросы для самоконтроля знаний, тесты, задачи для практических занятий.

Предназначен для преподавателей и студентов экономических специальностей вузов.

**УДК 311(075.8)**  
**ББК 60.6я73**

**ISBN 978-985-531-235-3 (Ч. 1)**  
**ISBN 978-985-531-234-6**

© Гордиенко О. И., 2012  
© УО «Полоцкий государственный университет», 2012

## СОДЕРЖАНИЕ

Введение .....	6
Часть 1	
Общая теория статистики	
Тема 1	
<b>ПРЕДМЕТ И МЕТОД СТАТИСТИЧЕСКОЙ НАУКИ .....</b>	<b>14</b>
1.1. Предмет статистики .....	14
1.2. Методы статистики и этапы статистического исследования .....	20
1.3. Основные категории статистики .....	24
1.3.1. Статистическая совокупность и единица статистической совокупности .....	24
1.3.2. Признаки как исходный объект статистического изучения .....	25
1.3.3. Типы шкал измерения признаков .....	29
1.4. Организация государственной статистики в Республике Беларусь .....	30
Задания для самостоятельной работы и практических занятий .....	31
Тема 2	
<b>СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ .....</b>	<b>35</b>
2.1. Сущность, цель и этапы статистического наблюдения .....	35
2.2. Проектирование статистического наблюдения .....	36
2.3. Организационные формы статистического наблюдения .....	40
2.4. Виды и способы статистического наблюдения .....	43
2.5. Ошибки статистического наблюдения и способы контроля точности статистических данных .....	48
Задания для самостоятельной работы и практических занятий .....	50
Тема 3	
<b>СВОДКА И ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ.</b>	
<b>СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ .....</b>	<b>55</b>
3.1. Значение и содержание статистической сводки .....	55
3.2. Сущность и виды статистических группировок .....	57
3.3. Определение числа групп и величины интервалов при группировке по количественным признакам .....	72
3.4. Виды и правила построения статистических таблиц .....	78
Задания для самостоятельной работы и практических занятий .....	87
Тема 4	
<b>СИСТЕМА СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ .....</b>	<b>96</b>
4.1. Сущность и классификация статистических показателей .....	96
4.2. Сущность абсолютных величин и их значение в статистике .....	97
4.3. Сущность, виды и порядок расчета относительных величин .....	101
Задания для самостоятельной работы и практических занятий .....	111
Тема 5	
<b>СРЕДНИЕ ВЕЛИЧИНЫ В СТАТИСТИКЕ .....</b>	<b>117</b>
5.1. Сущность средних величин, их виды и значение в статистическом анализе .....	117
5.2. Порядок расчета и свойства средней арифметической .....	121

5.3. Порядок расчета и сфера применения средней гармонической .....	127
5.4. Порядок расчета и сфера применения средней геометрической и средней квадратической .....	129
5.5. Структурные средние .....	132
Задания для самостоятельной работы и практических занятий .....	138
 Тема 6	
<b>СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВАРИАЦИИ</b> .....	145
6.1. Ряд распределения – основа изучения вариации .....	145
6.2. Графическое представление вариационных рядов распределения .....	150
6.3. Показатели вариации количественных признаков .....	154
6.4. Измерение вариации альтернативного признака .....	162
6.5. Правило сложения дисперсий и его применение в статистическом анализе .....	164
Задания для самостоятельной работы и практических занятий .....	169
 Тема 7	
<b>ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ</b> .....	176
7.1. Сущность и значение выборочного наблюдения .....	176
7.2. Основные способы формирования выборочной совокупности .....	179
7.3. Средние и предельные ошибки выборочного наблюдения .....	182
7.4. Определение необходимой численности выборки .....	188
7.5. Малая выборка .....	190
7.6. Распространение данных выборочного наблюдения на генеральную совокупность .....	191
Задания для самостоятельной работы и практических занятий .....	192
 Тема 8	
<b>СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНАМИКИ</b>	
<b>СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ</b> .....	199
8.1. Сущность, виды и правила построения рядов динамики .....	199
8.2. Показатели интенсивности изменения уровней ряда динамики .....	206
8.3. Средние показатели ряда динамики .....	215
8.4. Характеристика и методы выявления основной тенденции развития явлений .....	222
8.5. Понятие и оценка сезонной неравномерности развития .....	226
8.6. Методика оценки компонент динамики .....	228
Задания для самостоятельной работы и практических занятий .....	229
 Тема 9	
<b>ИНДЕКСНЫЙ МЕТОД В СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ</b> .....	237
9.1. Сущность индексов и значение индексного метода анализа .....	237
9.2. Классификация индексов .....	239
9.3. Индексы количественных показателей .....	243
9.4. Индексы качественных показателей .....	245
9.5. Система индексов переменного состава, постоянного состава и структурных сдвигов .....	248
9.6. Индексы пространственно-территориального сопоставления .....	249
9.7. Применение индексов в факторном анализе .....	250

9.8. Решение типовых задач по теме .....	255
Задания для самостоятельной работы и практических занятий .....	259
Тема 10	
<b>СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ .....</b>	<b>266</b>
10.1. Виды и характеристика взаимосвязей в статистике .....	266
10.2. Статистические методы выявления корреляционной связи	
между двумя признаками .....	270
10.2.1. Элементарные методы анализа .....	270
10.2.2. Дисперсионный анализ .....	272
10.2.3. Аналитические возможности и условия применения	
корреляционно-регрессионного анализа (КРА) .....	272
10.3. Измерение степени тесноты корреляционной связи	
между двумя признаками .....	274
10.3.1. Коэффициент корреляции знаков .....	275
10.3.2. Линейный коэффициент корреляции .....	275
10.3.3. Корреляционное отношение .....	278
10.4. Расчет и интерпретация однофакторных уравнений регрессии .....	280
10.5. Понятие о множественной корреляции и факторном анализе .....	286
10.6. Практические примеры анализа корреляционной связи .....	291
Задания для самостоятельной работы и практических занятий .....	301
Литература .....	306
Приложения .....	309
Приложение 1 .....	309
Приложение 2 .....	310
Приложение 3 .....	312
Приложение 4 .....	313
Приложение 5 .....	314

## ВВЕДЕНИЕ

Статистика – одна из общепрофессиональных дисциплин, которая в значительной степени способствует формированию профессионального уровня специалистов экономического профиля, т.к. дает знания, навыки и умения, необходимые для получения достоверной и актуальной информации, используемой при принятии управленческих решений в самых различных сферах деятельности менеджеров-экономистов.

Применение методов статистики в таких областях, как менеджмент, маркетинг, финансы, логистика и др., является важной предпосылкой разработки качественных управленческих решений.

Статистика – это искусство и наука сбора и анализа данных. Статистические методы следует рассматривать как важную часть процесса принятия решений, позволяющую вырабатывать обоснованные стратегические решения, сочетающие знания, опыт и интуицию специалиста с тщательным анализом имеющейся информации.

Почему специалистам в области управления необходимо изучать статистику?

Во-первых, знание методов статистики позволяет более компетентно и грамотно работать с данными.

Во-вторых, статистика позволяет выявить неочевидную информацию, которая не лежит на поверхности, а может быть извлечена и понята на основе применения специальных методов статистического анализа.

В-третьих, статистика может сказать, какова доверительная вероятность сделанных выводов или полученных результатов анализа.

В-четвертых, применение методов статистики позволит усилить аргументацию разработанных предложений и предлагаемых управленческих решений.

*Цель преподавания* дисциплины – изучение методов статистического исследования и возможностей их применения для идентификации проблем и принятия решений в области бизнеса и менеджмента.

В соответствии с образовательным стандартом по специальности 1-26 02 02 «Менеджмент» ОСРБ 1 26 02 02 2007 дисциплина «Статистика» включает два раздела: «Общая теория статистики» и «Статистика предприятий (организаций)».

*Целью изучения* раздела «Общая теория статистики» является освоение студентами методов статистического исследования и формирование практических навыков их применения для идентификации проблем и принятия решений в управлении социально-экономическими процессами.

Целью изучения раздела «Статистика предприятий (организаций)» является приобретение студентами теоретических знаний и формирование у них практических навыков в области построения и анализа системы статистических показателей, характеризующих ресурсы, результаты и эффективность производственно-хозяйственной деятельности предприятий (организаций).

*Задачи изучения дисциплины:*

1) по разделу «Общая теория статистики»:

- формирование знаний о роли и возможностях статистики в изучении социально-экономических явлений и процессов, получение ясного представления об основных категориях статистики, этапах и методах статистического исследования;
- формирование у студентов целостного представления о современных принципах, организационных формах, видах и способах статистического наблюдения;
- освоение методов проведения статистической сводки и группировки;
- освоение методов анализа сведенного статистического материала, методики расчета и интерпретации обобщающих статистических показателей;

2) по разделу «Статистика предприятий (организаций)»:

- усвоение состава показателей, обеспечивающих объективное измерение наличия и уровня использования ресурсов предприятий (организаций), а также методики расчета этих показателей и их аналитических возможностей;
- усвоение состава показателей, обеспечивающих объективное измерение результатов и эффективности производственно-хозяйственной деятельности предприятий (организаций) и методики расчета этих показателей;
- формирование знаний о возможностях использования статико-экономических моделей в статистическом изучении предприятий (организаций);
- формирование представлений о составе и содержании форм статистической отчетности по основным направлениям деятельности предприятий (организаций).

Решение указанных выше задач будет способствовать реализации компетентностно-ориентированного подхода к профессиональной подготовке специалистов по данной специальности при изучении дисциплины «Статистика».

Изучение дисциплины «Статистика» способствует формированию у студентов академических, социально-личностных и профессиональных компетенций.

Выпускник должен обладать следующими *академическими компетенциями*:

- владеть и применять базовые научно-теоретические знания в области статистики для решения теоретических и практических задач управления социально-экономическими процессами на предприятиях различных форм собственности;

- владеть приемами системного и сравнительного анализа при формировании и использовании статистических данных в своей профессиональной деятельности;

- уметь самостоятельно проводить статистические наблюдения (в т.ч. выборочные), обобщать и анализировать сведенный статистический материал;

- быть способным порождать новые идеи в области источников, способов получения и направлений использования статистических данных;

- владеть междисциплинарным подходом при организации проведения статистических исследований;

- иметь навыки, связанные с использованием компьютеров, управлением информацией в области обработки и анализа статистических данных, а также навыки представления их в графической и табличной формах;

- иметь лингвистические навыки (устная и письменная коммуникация) грамотного и компетентного представления аналитических записок по результатам использования статистических методов исследования;

- уметь учиться, повышать свою квалификацию в области владения статистическими методами исследования в течение всей трудовой жизни.

В процессе обучения выпускник должен сформировать следующие *социально-личностные компетенции*:

- гражданственность;

- способность к социальному взаимодействию;

- способность к межличностным коммуникациям, в частности при проведении статистического наблюдения на основе опроса;

- умение работать в командах;

- способность критически мыслить, готовность принимать новые установки, интерес к зарубежному опыту;

- знания идеологических, моральных, нравственных ценностей государства и умение следовать им.



Изучение дисциплины «Статистика» направлено на формирование у студентов важных *профессиональных компетенций*. По видам профессиональной деятельности студенты, обучающиеся по специальности 1-26 02 02, должны быть способны:

1) в *организационно-управленческой деятельности*:

- применять методы статистики в проведении организационно-управленческих обследований предприятия и его подразделений для разработки и обоснования управленческих решений;
- использовать статистические методы в организации контроля качества продукции;
- проводить статистическое изучение ритмичности деятельности предприятия (организации);
- проводить статистическое изучение качества продукции;

2) в *экономической деятельности*:

- проводить статистическое изучение хозяйственно-экономической деятельности предприятия (организации) и осуществлять ее прогнозирование методом экстраполяции;
- применять методы статистики при проведении анализа товарной, ценовой и инновационной политики;
- проводить статистическое изучение финансового состояния и финансовых результатов деятельности организации (предприятия);

3) в *научно-исследовательской деятельности*:

- разрабатывать и применять статистико-экономические модели для подготовки и обоснования управленческих решений;
- применять методы статистики в исследовательской деятельности по формированию и совершенствованию систем управления предприятием (организацией).

Формирование *профессиональных компетенций* специалистов по данной специальности предусматривает, что в результате изучения дисциплины «Статистика» студенты должны:

1) *знать*:

- значение и возможности применения статистических методов исследования в своей сфере профессиональной деятельности;
- основы организации и проведения статистического исследования;
- основные категории и методы статистической науки, ее структуру;
- состав, сущность, порядок расчета, анализа и интерпретации обобщающих статистических показателей;

- роль и возможности статистической методологии в изучении социально-экономических процессов на микроуровне;
- основы построения статистико-экономических моделей и их аналитический потенциал для разработки управленческих решений по совершенствованию деятельности предприятий (организаций);
- статистическую отчетность по основным направлениям деятельности предприятий (организаций);

2) *уметь:*

- проектировать статистические наблюдения по направлениям своей профессиональной деятельности;
- представлять данные в табличной и графической формах;
- проводить расчеты обобщающих статистических показателей и правильно их интерпретировать;
- анализировать статистические данные, выявлять статистические закономерности и представлять их количественное выражение;
- рассчитывать и анализировать показатели, обеспечивающие объективное измерение наличия и уровня использования ресурсов предприятий, а также результатов и эффективности их производственно-хозяйственной деятельности;

3) *иметь навыки:*

- организации и проведения статистических наблюдений;
- выбора и применения методов статистического анализа, обеспечивающих обоснование и аргументацию принимаемых решений в различных областях профессиональной деятельности менеджера-экономиста;
- проведения анализа социально-экономических явлений и процессов на реальных статистических данных предприятий (организаций);
- поиска необходимой статистической информации для решения конкретной исследовательской или профессиональной задачи;
- обобщения результатов статистического анализа и формулирования аналитических выводов;
- использования статистических программных продуктов для проведения статистического анализа.

Достижение цели учебной программы обеспечивается системным и последовательным изучением дисциплины во всех формах организации

учебного процесса: лекции, практические занятия, выполнение контрольных работ, экзамен, организация самостоятельной работы студентов.

Самостоятельная работа осуществляется в тесном сотрудничестве и под методическим и научным руководством преподавателя. Цель самостоятельной работы – научить студентов учиться, развить их творческие способности, повысить качество подготовки специалистов. В учебно-методическом комплексе разработаны следующие формы организации самостоятельной работы студентов:

- вопросы для проверки усвоения теоретического материала, а также тестовые задания;
- задачи для выполнения расчетов по основным направлениям и методам проведения статистического анализа;
- задания, предполагающие работу с периодическими изданиями, использование Интернет-ресурсов для поиска дополнительной информации по той или иной теме;
- задания, предполагающие применение методов статистического анализа на конкретном материале, который студенты формируют самостоятельно;
- тематика рефератов.

### **Диагностика компетенций студентов**

Для мониторинга качества изучения дисциплины и диагностики компетенций студентов используется следующий диагностический инструментарий: тесты, аудиторные и внеаудиторные практические задания, устный опрос во время практических занятий, контрольные работы, подготовка рефератов; экзамен.

### **Методы (технологии) обучения**

При изучении дисциплины «Статистика» применяются следующие эффективные методы и технологии:

- технология учебно-исследовательской деятельности;
- коммуникативные технологии (дискуссия и другие методы активного обучения);
- рейтинговые технологии;
- УМК;
- информационные технологии в форме презентаций;
- тестирование без применения компьютерных технологий;
- использование ПК для проведения статистического анализа.

Дисциплина «Статистика» базируется на общенаучных методах познания количественных закономерностей социально-экономических явлений и связана с такими дисциплинами, как «Микроэкономика», «Экономика предприятий», соответствующие разделы дисциплины «Высшая математика».

Учебным планом по специальности 1-26 02 02 предусмотрен следующий регламент изучения дисциплины:

Виды занятий, формы контроля знаний	Д 1 -26 02 02
Курс	2
Семестр	3
Лекции, ч	54
Экзамен (семестр)	3
Зачет (семестр)	–
Практические (семинарские), ч	36
Лабораторные занятия, ч	–
Расчетно-графические работы (семестр)	–
Контрольные работы (семестр)	–
Курсовая работа (семестр/ч)	–
Курсовой проект (семестр/ч)	–
Управляемая самостоятельная работа (ч)	–

Состав тем, изучаемых дисциплиной «Статистика», количество аудиторных часов в разрезе лекционных и практических занятий для студентов дневного отделения представляет нижеследующая таблица.

### Распределение аудиторных часов по темам и видам учебных занятий

Наименование раздела, темы	Объем, ч	
	Лекции	Практические занятия
<b>Раздел I. Общая теория статистики</b>		
1. Предмет и метод статистической науки	2	1
2. Статистическое наблюдение	2	1
3. Сводка и группировка статистических данных. Статистические таблицы	3	1
4. Система статистических показателей	2	2
5. Графический способ представления статистических данных	–	–
6. Средние величины в статистике	2	2
7. Статистическое изучение вариации	3	2
8. Выборочное наблюдение	2	2
9. Статистическое изучение динамики социально-экономических явлений	4	3
10. Индексный метод в статистических исследованиях	3	3
11. Статистическое изучение связи социально-экономических явлений	3	3
<i>Итого по разделу «Общая теория статистики»</i>	26	20
<b>Раздел II. Статистика предприятий (организаций)</b>		
12. Предприятие как объект статистического изучения	1	–
13. Статистика персонала предприятия (организации)	2	1
14. Статистика производительности труда	3	2
15. Статистика оплаты труда	2	2
16. Статистика основных средств	2	2
17. Статистика оборотных средств	2	1
18. Статистика производства и реализации продукции	3	1
19. Статистика качества продукции	2	1
20. Статистика себестоимости промышленной продукции	2	2
21. Статистика финансовых результатов и финансового состояния предприятия (организации)	4	2
22. Статистика инновационной деятельности предприятий (организаций)	3	1
23. Статистическое изучение экономической эффективности промышленного производства	2	1
<i>Итого по разделу «Статистика предприятий (организаций)»</i>	28	16
<i>Всего по дисциплине</i>	54	36

## Часть I

# ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

### Тема 1

## ПРЕДМЕТ И МЕТОД СТАТИСТИЧЕСКОЙ НАУКИ

- 1.1. Предмет статистики.
- 1.2. Методы статистики и этапы статистического исследования.
- 1.3. Основные категории статистики.
  - 1.3.1. Статистическая совокупность и единица статистической совокупности.
  - 1.3.2. Признаки как исходный объект статистического изучения.
  - 1.3.3. Типы шкал измерения признаков.
- 1.4. Организация государственной статистики в Республике Беларусь.

Изучив данную тему, студенты должны:

- получить ясное представление о предмете статистики и особенностях ее методологии;
- знать этапы статистического исследования;
- понимать роль закона больших чисел (ЗБЧ) в статистике и сущность категории «статистическая закономерность»;
- уметь охарактеризовать массовые социально-экономические явления в терминах основных категорий статистики;
- понимать значение применения методов статистики в экономике, управлении, бизнесе;
- знать основы организации государственной статистики в Республике Беларусь.

### 1.1. Предмет статистики

Термин «статистика» происходит от латинского слова «status», в переводе – состояние, положение вещей. Употреблялось оно в значении «политическое состояние». Отсюда произошло итальянское «stato» – государство и «statista» – знаток государства.

В современном языке термин «статистика» употребляют в нескольких значениях:

1) статистика – это совокупность статистических данных, сведений, характеризующих явления и процессы общественной жизни. В этом аспек-

те со статистикой знаком практически любой человек, т.к. газеты и журналы постоянно публикуют статистические обзоры и бюллетени о конъюнктуре рынков, индексах цен, уровне и динамике безработицы и др.;

2) статистика – практическая деятельность по сбору, обработке, анализу и публикации статистических данных о явлениях и процессах общественной жизни;

3) статистика – это отрасль знаний, т.е. наука, занимающаяся разработкой теоретических положений и методов сбора, обработки, анализа и интерпретации данных о массовых явлениях общественной жизни.

В данном курсе мы будем знакомиться со статистикой как наукой, хотя и другие аспекты будут предметом нашего внимания.

### **История статистики**

Еще в глубокой древности появилась потребность в проведении таких статистических операций, как учет мужского населения, численности войск, числа умерших и родившихся, имущества и др. Первое упоминание о практическом опыте статистического учета относится к XXII в. до н.э., когда в Древнем Китае регулярно проводился учет населения по полу и возрасту. Известно, что в V в. до н.э. учет численности войск персидского царя Дария осуществлялся следующим образом: каждый воин, проходя около определенного места, должен был положить камень. По количеству камней устанавливалось число воинов. В античных Афинах число родившихся и умерших определялось посредством подсчета приношений жрице богини Минервы: за каждого родившегося приносили в дар одну меру пшеницы, за умершего – меру ячменя. В древнем Риме каждый римлянин в присутствии сограждан под клятвой сообщал данные о своем имуществе.

Первоначально статистическая практика была несложной и охватывала немногие общественные явления. По мере развития производительных сил, торговли, товарно-денежных отношений шло и развитие статистических работ. Позднее потребность в статистических операциях возникала уже в связи с необходимостью стимулирования роста народонаселения, производительных сил страны, регулирования потребления, т.е. статистическая практика становилась глубже по содержанию, шире по кругу обследуемых объектов и совершеннее по применяемому инструментарию.

Проведение крупных и сложных статистических работ требовало установления общих правил статистической работы, нуждалось в обобщении

ях, создании теории, что и привело к возникновению особой общественной науки<sup>1</sup> – статистики.

Как наука статистика стала развиваться с середины XVII в. У истоков статистической науки стояли две школы:

- немецкая описательная школа государственоведения;
- английская школа политических арифметиков.

Основателем описательной школы был немецкий ученый Г. Конринг (1606 – 1681). Дальнейшее развитие это направление получило в работах Г. Ахенваля (1719 – 1772) и А. Шлецера (1735 – 1809).

Г. Ахенваль, профессор философии и права, ввел в научный обиход термин «статистика». С 1749 г. в Марбургском, а затем Геттингенском университетах он начал читать новую учебную дисциплину, которой было дано название «статистика».

Представители *школы государственоведения* считали, что задачей статистики является описание «государственных достопримечательностей», а именно: территории, государственного устройства, населения, религии, внешней политики, армии и т.п. Они систематизировали методы описания государства, разработали теорию таких описаний, но вели описание преимущественно в словесной форме без цифр и вне динамики. Содержание, задачи, предмет изучения статистики в понимании этих ученых были далеки от современного взгляда на статистику как науку.

Гораздо ближе к современному пониманию статистики была английская *школа политических арифметиков*. Представителями этой школы были Вильям Петти (1623 – 1687) и Джон Граунт (1620 – 1674).

Политические арифметики путем обобщения и анализа фактов стремились цифрами охарактеризовать состояние и развитие общества, показать закономерности развития общественных явлений, проявляющиеся в массовом материале. Цели и задачи, которые ставили перед собой ученые этой школы, близки и современному пониманию сущности статистики.

Дж. Граунт на основе обработки бюллетеней о естественном движении населения Лондона впервые открыл некоторые закономерности массовых общественных явлений, попытался построить таблицу смертности для стационарного населения.

В. Петти считается создателем экономической статистики. Его основные труды – «Политическая арифметика», «Разное о деньгах».

В первой половине XIX в. возникло третье направление статистической науки – *статистико-математическое*. Представителями этого направ-

---

<sup>1</sup> Поскольку статистика – это наука о методах, а не о социально-экономических явлениях, то она применяется и в науках о природе, т.е. является универсальной.



ления были А. Кетле (1796 – 1874), Ф. Гальтон (1822 – 1911), К. Пирсон (1857 – 1936), Р. Фишер (1890 – 1962), В. Госсет. Большое влияние на развитие математического направления в статистике оказали работы русских математиков П.Л. Чебышева (1821 – 1894), А.А. Маркова (1856 – 1922), А.М. Ляпунова (1857 – 1919).

### **Что изучает статистика и особенности ее методологии**

Статистика как наука имеет:

- свой объект исследования;
- свой предмет исследования;
- свои методы исследования;
- определенную структуру;
- своих пользователей.

*Объектом изучения* статистики являются массовые социально-экономические явления и процессы (далее мы будем использовать для их обозначения термин «статистическая совокупность»). Необходимость исследования массовых социально-экономических явлений, а не единичных, обусловлена тем, что закономерности развития процессов общественной жизни проявляются только через множество фактов, при обобщении данных по достаточно большому числу единиц (действие закона больших чисел). Но массовые социально-экономические явления изучает не только статистика, но и другие науки, например, демография, социология. Но каждая наука имеет свой предмет изучения. Ответ на вопрос «Что изучает статистика в массовых социально-экономических явлениях?» дает определение предмета статистики.

*Предмет статистики* – количественная сторона массовых социально-экономических явлений в неразрывной связи с их качественной стороной в конкретных условиях места и времени.

Свой предмет статистика изучает на основе статистической методологии.

*Статистическая методология* – совокупность приемов, правил и методов, пользуясь которыми статистика изучает свой предмет.

*Важнейшими особенностями* статистической методологии (статистического метода) являются:

- конкретность исследования (выбор определенного объекта исследования, конкретность места и времени его существования);
- примат качественного анализа в статистическом исследовании, т.е. необходимость проведения, прежде всего, содержательного анализа изучаемых явлений, выяснение их сущности с учетом всех факторов и условий развития;

- выделение однородных совокупностей, типов социально-экономических явлений;
- получение обобщающих характеристик изучаемых объектов в виде статистических закономерностей;
- интерпретация любых исключений из выявленных общих характеристик. Если данные содержат особые случаи, которые не «вписываются» в выявленную закономерность, их необходимо исследовать и объяснить;
- получение выводов, которые в ряде случаев носят вероятностный характер, и установление доверительной вероятности, с которой можно ожидать этих результатов;
- применение системы статистических показателей, позволяющих дать всестороннюю характеристику изучаемых явлений и процессов.

Характеристика специфических методов статистики дана в подразделе 1.2 данной темы.

### **Структура статистической науки**

Статистика представляет собой систему научных дисциплин, выделение которых обусловлено сложностью и своеобразием задач, которые общество ставит перед статистической наукой. Структуру статистической науки представляет рис. 1.1.

В составе статистической науки выделяют два *основных раздела: общую теорию статистики и социально-экономическую статистику*. Социально-экономическая статистика, в свою очередь, подразделяется на *экономическую статистику* и ее отрасли, а также *социальную статистику* и ее отрасли.

*Общая теория статистики* разрабатывает общие понятия, категории, принципы и методы статистической науки без конкретизации сферы ее применения. Общая теория статистики является методологической основой всей системы отраслевых статистик. Она формирует общестатистическое мировоззрение экономистов, менеджеров независимо от сферы их профессиональной деятельности.

*Социально-экономическая статистика* разрабатывает методологию статистической оценки и анализа процессов и явлений общественной жизни, учитываемых при формировании экономической политики страны и принятии решений по регулированию и прогнозированию ее социально-экономического развития. В рамках социально-экономической статистики особое место занимает *система национального счетоводства (СНС)*. Система национальных счетов используется для описания и анализа развития рыночной экономики на макроуровне, характеризует производство, обра-

зование, распределение, перераспределение и использование доходов, накопление национального богатства и связи страны с остальным миром.

Задачей *экономической статистики* является разработка и анализ показателей, отражающих состояние национальной экономики, взаимосвязи отраслей, особенности размещения производительных сил, наличие материальных, трудовых, финансовых ресурсов, достигнутый уровень их использования.

В структуре экономической статистики выделяется макроэкономическая статистика, разрабатывающая методы комплексного изучения экономики страны, межотраслевые связи и др. Помимо этого в совокупность отраслей экономической статистики входят статистика промышленности, сельского хозяйства, строительства, транспорта и др.

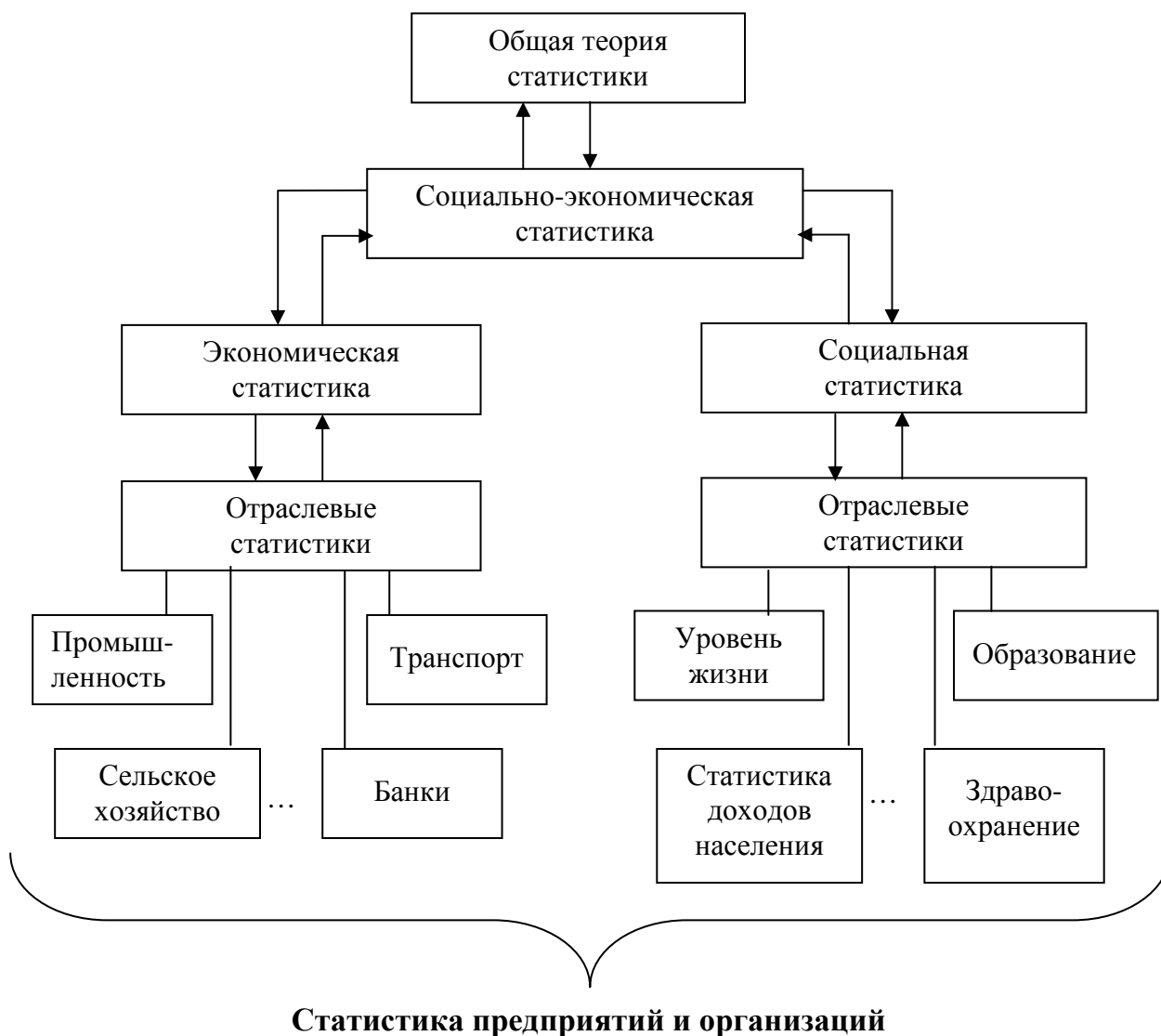


Рис. 1.1. Структура статистической науки

Источник: [1, с. 39]

*Социальная статистика* дает количественную характеристику структуры общества, жизни и деятельности людей, их взаимоотношений с государством и правом, позволяет выявить и измерить основные закономерности в поведении людей, в распределении благ между ними.

В состав социальной статистики входят статистика народонаселения, уровня жизни, культуры, образования, общественного мнения и других отраслей.

Задачей данного курса является изучение общей теории статистики, а также статистики предприятий (организаций).

### **Пользователи статистики**

В настоящее время статистические методы и статистические данные применяются для анализа различных социально-экономических явлений: при исследовании товарных рынков и рынков факторов производства, проведении аудиторских проверок, оценке хозяйственных рисков и финансового состояния хозяйствующих субъектов, в планировании и прогнозировании; страховании, маркетинге, управлении качеством продукции и др. Благодаря статистике управляющие органы получают объективную и всестороннюю характеристику управляемых объектов, будь то вся национальная экономика или отдельные ее отрасли, предприятия и их подразделения. В состав пользователей статистики входят:

- органы государственного управления всех уровней;
- коммерческие организации;
- научно-исследовательские и образовательные организации и учреждения;
- общественные организации;
- международные организации;
- население.

Во всех случаях задачей статистики является представление объективной информации для разработки и принятия рациональных управленческих решений.

## **1.2. Методы и этапы статистического исследования**

Объективные статистические данные получают в результате проведения *статистического исследования*. Важнейшим условием качественного статистического исследования является содержательный анализ изучаемых явлений и процессов, понимание их сущности и законов развития. Поэтому при проведении статистического исследования, наряду со *специфи-*

*ческими методами статистики*, используют общенаучные методы исследования и знания специальных дисциплин. Таким образом, выделяют три группы методов, используемых в статистическом исследовании:

1) *общенаучные методы познания*: диалектический метод, методы индукции и дедукции, логики, сравнения, аналогий и др. В соответствии с диалектическим методом познания статистика изучает все явления во взаимосвязи, движении и развитии, причинной обусловленности. Диалектический подход требует принимать во внимание процесс перехода количественных изменений в коренные качественные;

2) *методы и знания специальных наук*: экономической теории, экономики предприятий, финансов, маркетинга, демографии и др. Использование знаний специальных наук необходимо в связи с тем, что статистика изучает количественную сторону массовых явлений и процессов в неразрывной связи с их качественной стороной, определение которой и входит в задачи специальных наук;

3) *специфические методы статистики*.

К специфическим методам статистики относят:

- методы массового наблюдения;
- методы статистической сводки;
- методы расчета и анализа обобщающих статистических показателей.

*Методы массового статистического наблюдения* используют при сборе первичных статистических данных. Это **первый этап** любого статистического исследования.

*Методы статистической сводки и группировки* дают возможность собранные в процессе массового статистического наблюдения первичные данные подвергнуть систематизации и классификации. Это *второй этап* статистического исследования.

Методы расчета и анализа обобщающих статистических показателей позволяют выявить количественные закономерности, присущие изучаемым совокупностям и характеризующие уровни развития, структуру, динамику, взаимосвязи социально-экономических явлений. Эти методы используют на *третьем этапе* статистического исследования.

Таким образом, своеобразие методов статистики и определенная логическая последовательность их применения определяют состав и содержание этапов статистического исследования (табл.1.1).

Все стадии статистического исследования тесно связаны между собой. Недостатки, возникающие на одной из них, сказываются на всем исследовании в целом. Поэтому строгое соблюдение правил статистической науки обязательно на всех стадиях статистического исследования.

**Этапы и методы статистического исследования**

Этапы статистического исследования	Содержание этапа	Применяемые методы
1. Статистическое наблюдение	1. Проектирование статистического наблюдения. 2. Подготовка статистического наблюдения. 3. Проведение статистического наблюдения. 4. Контроль качества первичных статистических данных	1. Метод массового статистического наблюдения. 2. Выборочное наблюдение
2. Статистическая сводка	1. Расчет численности совокупности и значений признаков по совокупности (простая сводка). 2. Группировка статистических данных (сложная сводка). 3. Расчет абсолютных и других обобщающих показателей по группам и всей совокупности. 4. Оформление результатов сводки в форме статистических таблиц, графиков	Статистическая сводка и группировка
3. Расчет и анализ обобщающих статистических показателей	1. Расчет обобщающих статистических показателей. 2. Выявление и количественное выражение закономерностей и связей в изучаемых совокупностях. 3. Формулировка выводов	1. Расчет обобщающих показателей в форме <ul style="list-style-type: none"> <li>– абсолютных величин;</li> <li>– относительных величин;</li> <li>– средних величин.</li> </ul> 2. Методы анализа динамики. 3. Методы анализа вариации. 4. Методы анализа связей между явлениями. 5. Индексный анализ

**Роль закона больших чисел в статистике**

Основной задачей статистики является установление и измерение закономерностей, присущих изучаемым массовым социально-экономическим явлениям и процессам. Чтобы такие закономерности выявить, статистика должна изучать не отдельные явления, а их совокупности. В связи с тем, что статистика изучает массовые социально-эко-

номические явления, для применения ее методов важно учитывать действие закона больших чисел.

*Закон больших чисел (ЗБЧ) устанавливает, что количественные закономерности, присущие массовым общественным явлениям, отчетливо проявляются лишь в достаточно большом числе наблюдений.*

ЗБЧ выражает диалектику случайного и необходимого. В каждом отдельном явлении необходимое – т.е. то, что присуще всем явлениям данного вида, проявляется в единстве со случайным, индивидуальным, присущим лишь этому конкретному явлению. Единичные явления в большей степени подвержены воздействию случайных и несущественных факторов, чем совокупность в целом. В большом числе наблюдений разнонаправленные случайные отклонения взаимопогашаются. В результате взаимопогашения случайных отклонений типичные черты и закономерности, сформировавшиеся под действием существенных факторов в данных условиях места и времени, проявляются отчетливо.

Знание действия закона больших чисел позволяет грамотно применять метод средних величин, выборочное наблюдение, методы анализа связей между явлениями.

ЗБЧ не связан с сущностью явлений, не влияет на их содержание, их внутренние причинные связи, не определяет уровень развития исследуемых явлений. Он представляет собой лишь форму проявления *статистических закономерностей*.

### **Статистическая закономерность**

*Закономерности, которые проявляются лишь в большой массе явлений через преодоление случайностей, присущих единичным явлениям, называются статистическими.*

Статистические закономерности проявляются тем полнее и отчетливее, чем больше единичных явлений охвачено наблюдением. Важной особенностью статистической закономерности является то, что она не соблюдается в каждом конкретном единичном явлении, а проявляется только во всей их массе. Так, например, реклама какого-либо товара может не оказать влияния на рост объема продажи этого товара, однако обобщение данных о затратах на рекламу товаров и объеме их реализации показывает наличие прямой связи между этими показателями.

Приведем другой пример статистической закономерности. Изучая ожидаемую продолжительность жизни при рождении, можно видеть, что при всем различии в уровне развития стран, их культуры общим для всех является большая ожидаемая и фактическая продолжительность жизни женщин. Хотя в конкретной семье эта закономерность может не соблю-

даться. Другой пример. Установлено, что на 100 девочек рождается 104 – 106 мальчиков. Однако в отдельной семье и даже небольшом населенном пункте это соотношение может быть совершенно иным.

Статистическая закономерность по своей природе близка к закону. Она так же, как и закон, объективна, причинно обусловлена, но ее причинно-следственные связи менее устойчивы и относятся к определенному пространству и времени. Любое заметное изменение условий существования данной статистической совокупности окажет воздействие на статистическую закономерность.

Резюмируя рассмотрение статистической закономерности, укажем основные ее черты:

- статистическая закономерность объективна;
- причинно обусловлена;
- проявляется только в массе явлений;
- относительно неустойчива.

Для того чтобы лучше понять сущность статистической закономерности, рассмотрим понятие *динамической закономерности*.

*Динамическими называются закономерности, которые, являясь объективными и причинно обусловленными, проявляются в каждом отдельно взятом элементе совокупности.* Например, в каждой молекуле воды содержится два атома водорода и один атом кислорода; площадь круга зависит от радиуса и эта зависимость выражается формулой  $S = 2\pi r^2$ , которая справедлива для любого круга.

### **1.3. Основные категории статистики**

Теоретическую основу любой науки, в т.ч. и статистики, составляют понятия и категории, в совокупности которых выражаются исходные положения данной науки. В статистике к важнейшим категориям относятся: статистическая совокупность, признак, вариация, статистический показатель. Рассмотрим эти категории.

#### **1.3.1. Статистическая совокупность и единица статистической совокупности**

##### **Статистическая совокупность**

*Статистическая совокупность – это множество объективно существующих объектов, явлений, связанных общими существенными признаками и общностью процессов развития.*

Чтобы сформировать статистическую совокупность, необходимо выделить существенные признаки, присущие ее единицам и важные для



достижения целей конкретного статистического исследования. Например, общими существенными признаками статистической совокупности «промышленные предприятия (организации) частной формы собственности, находящиеся на территории Витебской области» являются: вид экономической деятельности предприятий (промышленность); форма собственности (частная) и территориальное расположение (Витебская область).

Статистическая совокупность может быть однородной и неоднородной.

*Однородная статистическая совокупность - это совокупность, в которой ее элементы (единицы) сходны между собой по существенным для данного исследования признакам и относятся к одному и тому же типу явлений.*

Понятие однородности относительно и вовсе не означает полного соответствия всех единиц совокупности, а лишь подразумевает близость основного свойства или существенных признаков. Одна и та же совокупность может быть однородна по одним признакам и неоднородна по другим. Например, предприятия могут быть однородными по виду перерабатываемого сырья, но разнородными по размеру, уровню рентабельности.

*Неоднородная статистическая совокупность – это совокупность, в которой элементы (единицы), ее составляющие, относятся к различным типам изучаемого явления.*

Для неоднородной совокупности исчисление обобщающих характеристик, в особенности в виде средней величины, неправомерны. Для образования однородных совокупностей используют метод группировки.

Статистическая совокупность состоит из *единиц совокупности*.

#### **Единица статистической совокупности**

*Единица совокупности – отдельные объекты или явления, образующие статистическую совокупность.* Например, в совокупностях предприятий, работников, вузов единицами совокупности соответственно являются: предприятие, работник, вуз.

Отдельная единица может рассматриваться как элемент различных статистических совокупностей. Человек может изучаться как единица населения страны, как работник предприятия, как потребитель товара. Включение единиц в совокупность определяется целями и объектом исследования: демография, рабочая сила, потребление товара.

#### **1.3.2. Признаки как исходный объект статистического изучения**

Единицы совокупности обладают определенными свойствами, качествами. Эти свойства принято называть *признаками*.

*Признак – конкретное свойство (характеристика, качество), присущее единице статистической совокупности.*

Например, признаки человека: возраст, образование, занятие, семейное положение, рост, вес и др.; признаки предприятия: форма собственности, организационно-правовая форма, численность работников, уровень рентабельности и т.д.

Признаки – непосредственный объект статистического изучения. Только через признаки можно охарактеризовать природу статистической совокупности и выявить присущие ей закономерности.

Признаки различаются способами их измерения и другими особенностями, влияющими на приемы статистического изучения. Поэтому проводят *классификацию признаков* по ряду оснований.

1. *С содержательной точки зрения* различают:

- существенные признаки;
- несущественные (второстепенные) признаки.

*Существенные признаки* выражают сущность совокупности, определяют ее однокачественность, обособляя от окружающих явлений, формируя закономерности, присущие исследуемой совокупности.

Все существенные признаки тесно связаны между собой, образуют единую систему, поэтому изменение одного признака ведет к изменению других.

Например, существенными признаками транспортного предприятия являются объем перевозок, численность работников, количество транспортных средств и др.; существенные признаки студента – успеваемость, количество пропусков, участие в научно-исследовательской работе, общественная активность и др.

*Несущественные признаки* играют второстепенную роль в формировании определяющих совокупность закономерностей.

Несущественные признаки дают дополнительную информацию о свойствах явлений и интересны для статистического исследования, если они оказывают влияние на существенные признаки.

Возвращаясь к вышеприведенным примерам, в качестве несущественных признаков транспортного предприятия можно указать его название, у студента – возраст, национальность, группу крови и др.

Отнесение признаков к существенным или несущественным зависит от задач проводимого исследования и от того, в какую совокупность входят исследуемые явления. Если национальность, группа крови, возраст несущественны для характеристики студента, то при переписи населения возраст будет существенным признаком, но при этом студент рассматривается как единица населения страны.

2. *По степени изменчивости (вариации)* различают:

- статические (неварьирующие) признаки;
- варьирующие признаки:

- с альтернативной вариацией;
- с множественной вариацией.

*Статические признаки – это признаки, которые имеют неизменные (одинаковые) значения у всех единиц совокупности.* Пример абсолютно статического признака (т.е. неизменного во времени и при переходе от одной единицы совокупности к другой): у каждого человека одно сердце. Существуют признаки, статические во времени для данной единицы, например, группа крови у конкретного человека, однако в пространстве, т.е. при переходе от человека к человеку, группа крови меняется (варьирует).

Статистика изучает варьирующие признаки.

*Вариация – это изменение (колеблемость) значений признака при переходе от одной единицы совокупности к другой.*

*Варьирующие признаки – признаки, которые изменяют свои значения при переходе от одной единицы совокупности к другой.*

Например, возраст людей может принимать разные значения от 0 до 100 и более; такой признак, как национальность, также может принимать разные значения: белорус, еврей, русский и т.д.

Если изменения изучаемого явления происходят во времени, то говорят уже не о вариации значений признака, а о его динамике (см. тему 8 «Статистическое изучение динамики социально-экономических явлений»).

Вариация статистических признаков может быть *альтернативной и множественной*.

*Альтернативным признаком (признаком с альтернативной вариацией) является признак, который имеет только два значения.* Следовательно, конкретная единица совокупности может принимать только одно из двух возможных значений альтернативного признака. Например, признак «пол» имеет только два значения: мужской или женский. Признак «качество продукции» принимает также два значения: либо продукция годная (соответствует установленным требованиям), либо нет (брак).

*Признаком с множественной вариацией является признак, который имеет более двух возможных значений.* Например, национальность человека, его рост, вес и т.д.

Значение признака у конкретной единицы совокупности называется *вариантой* и обозначается как  $x$ . Варианты могут принимать любые значения в пределах размаха вариации признака. Они могут повторяться при переходе от одних единиц к другим.

3. По *форме выражения значений признака у единиц совокупности различают:*

- атрибутивные (качественные) признаки;
- количественные признаки.

*Атрибутивные признаки – это признаки, варианты которых не имеют числового выражения и выражаются словесно.*

Для этих признаков можно подсчитать только число единиц совокупности, обладающих определенными значениями признака, и затем на этой основе рассчитать такие статистические показатели, как мода, показатели структуры.

*Количественные признаки – это признаки, значения которых выражаются числами, имеющими содержательную интерпретацию.* К количественным признакам не относят числа, отражающие порядковый номер, код.

Количественные признаки могут характеризоваться *дискретной* и *непрерывной* вариацией.

*Количественные признаки с непрерывной вариацией* могут принимать любое числовое значение в пределах размаха вариации признака. Например, вес, рост человека, объем производства продукции на предприятиях и др. На практике значения непрерывных признаков округляют с определенной степенью точности, так что они становятся квазидискретными (мнимодискретными).

*Количественные признаки с дискретной вариацией* (дискретные признаки) – это признаки, значения которых отличаются друг от друга на фиксированную величину без промежуточных значений между ними. Дискретные признаки, как правило, имеют целочисленные значения. Например, тарифный разряд, число членов семьи, количество этажей в здании и т.п.

Существует понятие *квазинепрерывной вариации*, под которой понимают вариацию дискретного признака, который имеет множественную вариацию в весьма широких пределах. Например, численность работников предприятия, численность населения страны и др. Строго говоря, эти признаки характеризуются дискретной вариацией, но с ними проводят статистические операции как с непрерывными признаками, например, строят интервальные ряды распределения, а не дискретные.

4. По роли в причинно-следственных связях выделяют:

- факторные признаки;
- результативные признаки.

Деление признаков на эти две группы основано на содержательном анализе причинно-следственных связей между ними.

*Признаки, которые оказывают влияние на другой, связанный с ними, признак и обуславливают его вариацию, называют факторными.*

Существует, однако, более развернутая характеристика факторных признаков, в соответствии с которой под названием факторного признака объединяют *признак-причину* и *признак-условие*.

*Признаком-причиной* называется такой признак, который в определенной связи является причиной изменения другого признака, а *признак-условие* создает возможность такого изменения. Поскольку признаки-причины и

признаки-условия часто трудно различимы, на практике их объединяют под названием факторного признака.

*Признак, который изменяет свое значение под влиянием других, связанных с ним факторных признаков, называется результативным, или признаком-следствием.*

Следует обратить внимание, что деление признаков на факторные и результативные не является раз и навсегда данным и зависит от целей исследования, т.е. в одном исследовании признак может выступать как факторный, в другом – как результативный. Например, в системе признаков «себестоимость единицы продукции» и «прибыль от реализации продукции» первый является факторным признаком, второй – результативным. Однако в системе признаков «прибыль от реализации продукции» и «рентабельность реализованной продукции» прибыль выступает уже как факторный признак, а рентабельность как результативный.

### **1.3.3. Типы шкал измерения признаков**

Для статистического изучения признаков их необходимо измерить. Измерение признаков означает присвоение определенного значения каждой единице совокупности в соответствии с заранее определенными правилами (типами шкал измерения). В теории измерений используют четыре типа шкал измерения:

- 1) номинальная шкала (шкала наименований);
- 2) порядковая шкала;
- 3) интервальная шкала;
- 4) относительная шкала.

*Шкала наименований* (номинальная шкала) используется для измерения качественных признаков. Шкала наименований позволяет различать, идентифицировать объекты. Ее применяют в тех случаях, когда нельзя определить количественные или порядковые соотношения между значениями признака.

*Для шкалы наименований применимы такие виды статистических показателей, которые основаны на подсчете частоты. Например, удельный вес, мода.*

*Порядковая шкала* позволяет не только различать объекты, но и ранжирует их по степени проявления некоторого свойства.

Обратите внимание: в порядковых шкалах численные характеристики отличаются только по порядку (рангу), но не по величине. Поэтому, *в дополнение к статистическим показателям, доступным для данных шкалы наименований, порядковые шкалы позволяют рассчитать квартили, медиану.*

*Интервальная шкала* – это шкала, в которой числа используют для ранжирования объектов таким образом, чтобы численно равные различия на шкале представляли равные различия измеряемого признака.

Интервальная шкала содержит всю ту информацию, которая есть в порядковой шкале, но в дополнение к этому она позволяет сравнивать (измерять) различия между объектами. В интервальной шкале начало отсчета и единицы измерения выбираются произвольно.

Шкалы этого типа широко используются в маркетинговых исследованиях для определения отношения потребителей к товару. Например, респонденты оценивают отношение к товару по семибалльной шкале, где «1 = абсолютно не нравится», а «7 = очень нравится».

*Для интервальной шкалы применимы все статистические показатели, используемые для номинальной и порядковой шкал, а также в дополнение к ним возможен расчет средней арифметической и показателей вариации.*

*Шкала отношений* позволяет идентифицировать объекты, ранжировать их, а также сравнивать интервалы или различия и вычислять отношения между значениями признака.

Таким образом, как следует из вышеприведенного определения, шкала отношений обладает всеми свойствами шкалы наименований, порядковой и интервальной шкал. В шкале отношений начало отсчета строго фиксировано, поэтому отношение между значениями шкалы имеет смысл.

*К данным шкалы отношений применимы все статистические методы обработки данных.*

Измерение признаков позволяет получать статистические показатели (см. тему 4).

#### **1.4. Организация государственной статистики в Республике Беларусь**

В Республике Беларусь осуществляется государственная статистическая деятельность.

Государственная статистика – система сбора, обработки, обобщения, накопления, хранения, защиты, представления распространения и использования органами государственной статистики и уполномоченными государственными организациями первичных статистических данных об экономическом, демографическом, социальном и экологическом положении в Республике Беларусь.

Государственная статистика в Республике Беларусь осуществляется в соответствии с законом Республики Беларусь «О государственной статистике» [2].

Основными задачами государственной статистики являются:

– разработка научно-обоснованной статистической методологии и ее совершенствование в соответствии с национальными и международными стандартами в области статистики;

– сбор, обработка, обобщение, накопление, хранение и защита статистических данных (информации) на основе статистической методологии;

– распространение сводных статистических данных.

Основными принципами государственной статистики являются:

– научная обоснованность и объективность статистических данных;

– актуальность и своевременность статистических данных;

– конфиденциальность первичных статистических данных;

– сопоставимость статистических данных;

– доступность и открытость сводных статистических данных;

– профессионализм и независимость при осуществлении государственной статистической деятельности.

Система органов государственной статистики имеет иерархическую структуру и включает:

– республиканский орган государственного управления в области государственной статистики;

– территориальные органы государственной статистики.

Республиканский орган государственной статистики представлен Национальным статистическим комитетом. Он проводит государственную политику, осуществляет руководство государственной статистикой в Республике Беларусь и координирует деятельность в этой области.

К территориальным органам государственной статистики относятся областные и Минское городское управление статистики.

Территориальные органы государственной статистики подчиняются республиканскому органу государственного управления в области государственной статистики и действуют на основе положений, утвержденных этим органом.

В структуру областных и Минского городского управлений статистики входят отделы статистики в районах и городах (районах г. Минска).

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Что означает термин «статистика»?
2. Чем обусловлено возникновение и развитие статистической науки и практики?
3. Что является объектом изучения статистики?

4. Что является предметом изучения статистической науки?
5. В чем заключается сущность статистической методологии?
6. Охарактеризуйте структуру статистической науки.
7. Назовите и охарактеризуйте этапы статистического исследования.
8. Назовите методы, используемые в процессе статистического исследования.
9. Каковы специфические методы статистики?
10. Охарактеризуйте сущность закона больших чисел и поясните, какова его роль в статистике.
11. Что такое статистическая закономерность? Назовите ее основные черты.
12. В чем состоит отличие динамической закономерности от статистической?
13. Что такое статистическая совокупность?
14. Чем определяется однородность статистической совокупности?
15. Охарактеризуйте принятую в статистике классификацию признаков наблюдаемых единиц совокупности.
16. Что такое вариация признака?
17. Какие типы шкал используются для измерения признаков?
18. Принципы организации государственной статистики в Республике Беларусь.
19. Охарактеризуйте систему органов государственной статистики в Республике Беларусь.

### Тест 1

1. Предметом статистики как науки является ...
  - а) метод статистики;
  - б) группировка и классификация;
  - в) количественная сторона массовых социально-экономических явлений;
  - г) статистические признаки.
2. Статистическая совокупность – это ...
  - а) математическое множество;
  - б) любое множество явлений природы и общества;
  - в) множество разнородных элементов;
  - г) множество объективно существующих элементов, связанных общими существенными признаками и общностью процессов развития.
3. Укажите черты, присущие статистической закономерности:
  - а) проявляется лишь в массе явлений;
  - б) проявляется в каждом отдельном взятом явлении;



- в) носит объективный характер;
- г) причинно обусловлена.

4. Укажите черты, присущие динамической закономерности:

- а) проявляется лишь в массе явлений;
- б) проявляется в каждом отдельном взятом явлении;
- в) носит объективный характер;
- г) причинно обусловлена.

5. Укажите верное утверждение:

- а) ЗБЧ определяет сущность статистической закономерности;
- б) ЗБЧ представляет собой форму проявления статистической закономерности.

6. Является ли верным утверждение, что значительное изменение условий существования и развития данной статистической совокупности окажет воздействие на статистическую закономерность?

- а) да;
- б) нет.

7. Единица статистической совокупности – это ...

- а) признак совокупности;
- б) элемент математического множества;
- в) элемент таблицы Менделеева;
- г) носитель значений признаков.

8. Какие из перечисленных признаков являются статистическими:

- а) цена одного килограмма товара;
- б) курс доллара;
- в) количество атомов в молекуле воды;
- г) температура кипения воды?

9. Какие из перечисленных признаков являются дискретными:

- а) количество этажей в здании;
- б) вес человека;
- в) рост человека;
- г) количество членов семьи?

10. Признаки единиц статистической совокупности бывают ...

- а) только количественные;
- б) количественные и качественные;
- в) только качественные;
- г) только неименованные.

11. Вариация признака – это ...

- а) изменение массовых явлений во времени;
- б) изменение состава совокупности;
- в) изменение структуры совокупности в пространстве;
- г) изменение значений признака у единиц совокупности.

12. Укажите, какой из показателей является характеристикой совокупности работников предприятия:

- а) заработная плата экономиста Иванова И.И. в январе 2010 г. составила 1 214 тыс. руб.;
- б) средняя месячная заработная плата одного работника предприятия составила в январе 2010 г. 1 153 тыс. руб.

13. Все ли атрибуты статистического показателя содержит показатель «Среднегодовая выработка одного среднесписочного работника составила 114, 6 млн руб.»?

- а) да;
- б) нет.

### **Задания по теме**

1. Какие совокупности можно выделить в высшем учебном заведении для статистического изучения?

2. Какие совокупности могут быть выделены среди торговых предприятий города?

3. Какие совокупности могут быть выделены в жилом фонде города?

4. Какими наиболее существенными признаками можно охарактеризовать следующие единицы:

- а) промышленные предприятия;
- б) торговое предприятие;
- в) город;
- г) вуз;
- д) семья.

5. Выберите по своему усмотрению единицу статистического наблюдения и перечислите основные ее признаки, выделив среди них количественные и атрибутивные.

6. Назовите основные факторные признаки, определяющие вариацию успеваемости студентов.

7. Назовите основные факторные признаки, влияющие на уровень производительности труда рабочих.

8. Приведите примеры статистических совокупностей, используя периодические издания за текущий год и материалы сайта Национального статистического комитета Республики Беларусь. Укажите признаки, характеризующие совокупность.

## Тема 2

### СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

- 2.1. Сущность, цель и этапы статистического наблюдения.
- 2.2. Проектирование статистического наблюдения.
- 2.3. Организационные формы статистического наблюдения.
- 2.4. Виды и способы статистического наблюдения.
- 2.5. Ошибки статистического наблюдения и способы контроля точности статистических данных.

Изучив данную тему, студенты должны:

- понимать роль статистического наблюдения в проведении качественного статистического исследования;
- знать основные этапы статистического наблюдения;
- уметь разработать программно-методологическое обеспечение статистического наблюдения;
- иметь представление об организационных вопросах статистического наблюдения;
- знать организационные формы, виды и способы статистического наблюдения;
- понимать значение точности статистического наблюдения и знать способы выявления ошибок наблюдения.

#### 2.1. Сущность, цель и этапы статистического наблюдения

Статистическое наблюдение – первый этап статистического исследования.

*Статистическое наблюдение – научно-организационный сбор данных о социально-экономических процессах и явлениях путем регистрации их существенных признаков по специально разработанной программе.*

Цель статистического наблюдения – получение объективной и полной информации об объекте наблюдения.

Собранные и зафиксированные в процессе статистического наблюдения данные называются *первичным статистическим материалом*.

Качественно подготовленное и правильно проведенное статистическое наблюдение должно отвечать следующим требованиям:

- 1) иметь научно-практическую ценность, т.е. способствовать решению определенных научных или практических задач;
- 2) обеспечивать получение достоверных данных, которые соответствуют тому, что есть на самом деле;

- 3) обеспечивать полноту охвата наблюдаемого объекта за счет:
  - полного охвата единиц наблюдения;
  - охвата наиболее существенных признаков единиц наблюдения;
  - получения данных за максимально длительные периоды, если явление изучается во времени;
- 4) способствовать своевременному и оперативному получению статистических данных;
- 5) учитывать правила сопоставимости данных по методике получения, времени, пространственным объектам;
- 6) иметь программно-методологическое обеспечение и четко продуманную организацию.

Проведение статистического наблюдения включает следующие этапы:

- проектирование статистического наблюдения;
- подготовка к наблюдению (размножение статистических формуляров и инструкций; подбор и подготовка статистиков; проведение пробного статистического наблюдения);
- непосредственное статистическое наблюдение;
- контроль материалов статистического наблюдения.

Для того чтобы лучше понять характер задач, возникающих в процессе статистического наблюдения, охарактеризуем основные его элементы (рис. 2.1).

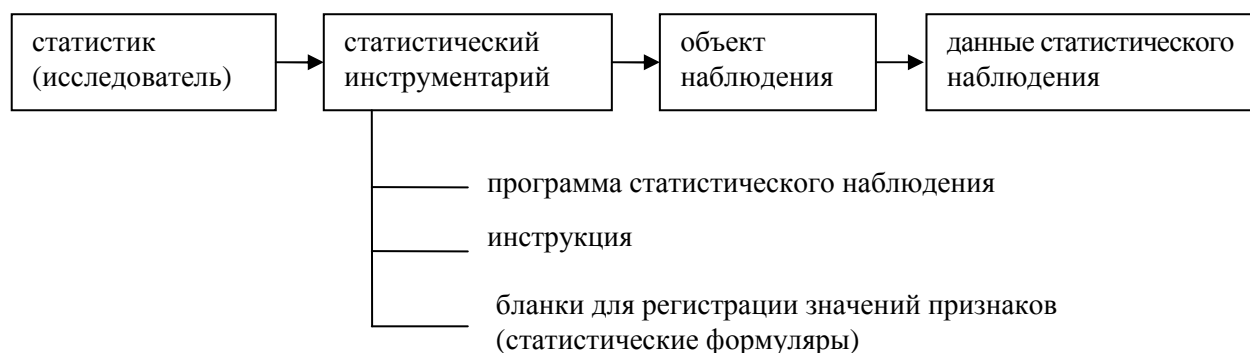


Рис. 2.1. Элементы статистического наблюдения

## 2.2. Проектирование статистического наблюдения

Проектирование статистического наблюдения включает:

- разработку программно-методологического обеспечения;
- решение организационных вопросов наблюдения.

## **Программно-методическое обеспечение статистического наблюдения**

Разработка *программно-методического обеспечения* статистического наблюдения включает определение:

- цели статистического наблюдения;
- объекта наблюдения;
- единицы наблюдения;
- программы наблюдения.

При проектировании статистического наблюдения в первую очередь формулируется его цель. *Цель наблюдения – это познавательная задача наблюдения, которая в конечном итоге сводится к информационному обеспечению процесса принятия решения по исследуемой проблеме.* Так, например, при разработке экономической и социальной политики государства важной предпосылкой является наличие объективной информации о рынке труда. Для сбора данных о состоянии рынка труда организуются наблюдения по вопросам занятости, безработицы, заработной платы, затрат на рабочую силу.

Нечетко поставленная цель может привести к тому, что в процессе наблюдения будут собраны ненужные данные или, наоборот, не будут получены сведения, необходимые для анализа.

*Цель определяет объект статистического наблюдения.*

*Объект статистического наблюдения – статистическая совокупность, о которой в соответствии с поставленной целью должны быть собраны необходимые статистические данные.*

Чтобы определить объект статистического наблюдения необходимо установить границы изучаемой совокупности, которые будут четко отделять данный объект от других, смежных с ним. В связи с этим в статистике используют такое понятие, как *ценз*. *Ценз – это существенные признаки, при наличии которых явления и процессы относятся к изучаемой совокупности.*

Важность правильного определения объекта наблюдения можно проиллюстрировать на следующем примере [3, с. 25]. Существует порог экономической безопасности по уровню безработицы в размере 6 – 8 %, превышение которого ведет к социальной напряженности и конфликтам в обществе. Поэтому государство должно отслеживать и регулировать уровень занятости и безработицы. Что же в этом случае следует считать объектом наблюдения? Оказывается, недостаточно указать, что объектом наблюдения является население, надо конкретизировать, что это *экономически активное население*, т.е. часть населения, которая предлагает свой труд для производства товаров и услуг.

Всякий объект статистического наблюдения состоит из отдельных элементов – *единиц наблюдения*. Характеристики объекта наблюдения могут быть получены только на основе изучения единиц наблюдения.

*Единица наблюдения – составной элемент объекта наблюдения, который является носителем признаков, подлежащих регистрации.*

В зависимости от предмета исследования единицами наблюдения могут выступать:

- физические лица (люди, составляющие население страны, региона, города и т.п.; работники предприятий и учреждений; потребители продукции; студенты и др.);
- реально существующие физические объекты (станки, машины, здания и т.п.);
- организационно обособленные единицы (это могут быть как юридические лица, например предприятия, банки, вузы и др., так и образования, не имеющие такого статуса, например, цеха, бригады, студенческие группы и т.п.);
- отдельные события и состояния исследуемых процессов (например, факты рождения, вступления в брак, развода; простои оборудования; случаи прогулов, неявки на работу по болезни и др.).

В отдельных наблюдениях может быть не одна, а несколько единиц наблюдения. Например, при переписи населения единицей наблюдения может быть человек и семья.

Следует обратить внимание на возможное расхождение между *единицей наблюдения* и *единицей совокупности*. *Единица совокупности* – то, что подвергается обследованию. *Единица наблюдения* – источник сведений о единицах совокупности. Единицы наблюдения могут совпадать и не совпадать с единицами совокупности. При переписи скота, например, единицей совокупности является животное, а единицей наблюдения – сельскохозяйственный двор. При переписи населения и единицей совокупности, и единицей наблюдения является человек.

Для определения состава регистрируемых признаков у единиц наблюдения разрабатывают *программу наблюдения*.

*Программа наблюдения – перечень признаков единицы наблюдения, подлежащих регистрации в процессе статистического наблюдения.* Содержание программы зависит:

- от целей исследования;
- особенностей наблюдаемого объекта;
- установленных сроков получения материалов наблюдения;

- установленных требований к степени детализации исследования;
- располагаемых ресурсов (материально-технических, трудовых, финансовых).

От того, насколько хорошо разработана программа статистического наблюдения, во многом зависит качество собранной информации.

При составлении программы следует ориентироваться на соблюдение следующих требований:

- программа должна содержать только такие вопросы, которые, безусловно, необходимы для данного статистического исследования;
- в программу следует включать лишь те вопросы, на которые можно получить достоверные ответы;
- программу наблюдения целесообразно строить так, чтобы ответами на одни вопросы можно было контролировать ответы на другие;
- не следует включать в программу вопросы, способные вызвать подозрение, что ответы на них могут быть использованы против отвечающих.

Для обеспечения единообразия получаемых сведений от каждой единицы наблюдения программа оформляется в виде документа, называемого *статистическим формуляром*. Формуляр может быть представлен в форме отчета, переписного листа, карточки учета, анкеты и т.д.

Различают две системы статистического формуляра: индивидуальную и списочную. *Индивидуальный формуляр* предусматривает регистрацию ответов на вопросы программы только об одной единице наблюдения, а списочный о нескольких единицах. Например, все формы статистической отчетности заполняются каждым предприятием (организацией) в отдельности, т.е. в этих случаях используют индивидуальный формуляр.

Кроме формуляра разрабатывается *инструкция*, определяющая порядок проведения наблюдения и заполнения формуляра. В зависимости от сложности программы наблюдения инструкция публикуется в виде отдельной брошюры или помещается на оборотной стороне формуляра. Формуляр и инструкция по его заполнению составляют *инструментарий статистического наблюдения*.

### **Организационные вопросы статистического наблюдения**

При проведении статистического наблюдения наряду с программно-методологической подготовкой большое значение имеет решение следующих *организационных вопросов*:

- определение времени и места статистического наблюдения;
- определение органов, проводящих наблюдение;

- разработку календарного плана статистического наблюдения;
- составление сметы затрат.

Выбор места проведения наблюдения зависит, главным образом, от цели наблюдения и охватывает все места нахождения единиц наблюдения. Если необходимо получить данные для изучения состава населения по стране, то в этом случае наблюдение охватывает территорию всей страны. При сборе информации о развитии малого бизнеса в Витебской области, местом проведения обследования будет территория этой области.

Выбор времени наблюдения заключается в решении двух вопросов:

- установление *критического момента* (даты) или *интервала времени*;
- установление срока (периода) наблюдения.

Под *критическим моментом* (датой) понимается конкретный день года, час дня, по состоянию на который должна быть проведена регистрация признаков по каждой единице наблюдения. Критический момент устанавливается с целью получения сопоставимых данных.

Если же надо проанализировать изменение значений признаков в данном периоде по сравнению с принятым за базу, то устанавливается не критический момент, а *критический интервал*, за который следует получить данные.

*Срок (период) наблюдения* – это время, в течение которого заполняются статистические формуляры, т.е. время, необходимое для массового наблюдения. Этот срок определяется исходя из объема работы (числа регистрируемых признаков и единиц в объекте наблюдения), численности персонала, занятого сбором информации. Следует учитывать, что отдаление периода наблюдения от критического момента или интервала времени может привести к снижению достоверности получаемых сведений.

### **2.3. Организационные формы статистического наблюдения**

В статистической практике используются три организационные формы статистического наблюдения (рис. 2.2.):

- *статистическая отчетность* (предприятий, организаций, учреждений и т.п.);
- *специально организованное статистическое наблюдение* (переписи, единовременные учеты, специальные обследования сплошного и несплошного характера, мониторинг);
- *регистры*.





Рис. 2.2. Организационные формы статистического наблюдения

### **Статистическая отчетность**

*Статистическая отчетность – основная форма статистического наблюдения, при которой организации и предприятия представляют в органы статистики сведения о своей деятельности в виде отчетов, регламентированных по содержанию и срокам представления за подписью лиц, ответственных за своевременность и достоверность представленных данных.*

Более подробно организация статистической отчетности как формы статистического наблюдения рассматривается в теме 11 (часть 2).

### **Специально организованное наблюдение**

Специально организованное наблюдение проводится с целью получения сведений, отсутствующих в отчетности, или проверки ее данных.

Наиболее простым примером такого наблюдения является *перепись*.

*Перепись – это специально организованное наблюдение, повторяющееся, как правило, через равные промежутки времени с целью получения информации об объекте наблюдения на определенный момент времени.*

Характерными особенностями переписи являются:

- одновременность ее на всей территории, которая должна быть охвачена наблюдением;
- единство программы наблюдения;
- регистрация всех единиц наблюдения по состоянию на один и тот же критический момент времени.

Практическая статистика проводит переписи населения, многолетних насаждений, объектов незавершенного строительства, оборудования и др. Из всех переписей наиболее известны *переписи населения*. Их цель состоит в установлении численности и размещения населения на территории страны, получении данных о структуре населения по полу, возрасту, занятиям и другим показателям. Первая всеобщая перепись населения России была проведена в 1897 г. В Республике Беларусь последняя перепись населения проводилась в 2009 г.

Переписи проводятся как работниками статистических органов, так и специально подготовленными лицами, которые путем организованной регистрации фактов заполняют статистические формуляры переписи.

Разновидностью специально организованных наблюдений являются *единовременные учеты*. Единовременный учет отличается от переписи тем, что при его проведении статистические формуляры, разработанные статистическими органами, заполняются по инструкции этих органов работниками предприятий и организаций на основе данных первичного учета.

Статистическое наблюдение может проводиться также в форме специально организованного обследования, направленного, например, на изучение покупательского спроса, общественного мнения, использования рабочего времени (фотография рабочего времени, хронометраж) и др. Чаще такие обследования организуют как выборочные.

Необходимость постоянного отслеживания динамически меняющихся процессов в социальной и экономической сферах привела к распространению такой формы специальных наблюдений как *мониторинг*.

*Мониторинг – система мероприятий, позволяющих непрерывно следить за состоянием определенного объекта, регистрировать его важнейшие характеристики, оценивать их, оперативно выявлять результаты воздействия на объект различных факторов.*

Мониторинг позволяет разрабатывать предложения по развитию объекта в нужном направлении и делать заключения об эффективности мер по управлению объектом. Например, социальный мониторинг обеспечивает комплексный оперативный контроль и управление процессами общественной жизни всего населения или групп людей (занятых, безработных, пенсионеров, молодежи) в регионах.

## **Регистры**

*Регистровое наблюдение – это форма непрерывного статистического наблюдения за долговременными процессами, имеющими фиксированное начало, стадию развития и фиксированный конец. Оно основано на ведении статистического регистра.*

Регистр представляет собой систему, постоянно следящую за состоянием единицы наблюдения и оценивающую силу воздействия различных факторов на изучаемые показатели. Одни из них остаются неизменными в течение всего времени наблюдения и регистрируются один раз; другие показатели, периодичность изменения которых неизвестна, обновляются по мере изменения; третьи – представляют динамические ряды показателей с заранее известным периодом обновления. Все показатели хранятся до полного завершения наблюдения за единицей обследуемой совокупности.

Например, для организации статистического наблюдения за предприятиями органы государственной статистики формируют и ведут *статистический регистр*. *Статистический регистр – перечень респондентов с указанием сведений о них, необходимых для организации государственных статистических наблюдений*. Одним из основных источников сведений для формирования статистического регистра является Единый государственный регистр юридических лиц и индивидуальных предпринимателей.

#### **2.4. Виды и способы статистического наблюдения**

Статистическая информация может быть получена различными способами (рис. 2.3), основными из которых являются:

- непосредственное наблюдение;
- документальный учет фактов;
- опрос.

*Непосредственное наблюдение* осуществляется путем фиксации в формуляре фактов, лично установленных регистраторами в результате осмотра, измерения, подсчета признаков изучаемого явления. Таким способом осуществляется наблюдение за вводом в действие жилых домов. Другим примером этого способа наблюдения является проведение инвентаризации, изучение использования рабочего времени методом моментных наблюдений, фотографии рабочего времени, хронометража.

*Документальное наблюдение*, основано на использовании в качестве источника информации данных различных документов. Например, при составлении статистической отчетности по финансам используют документы бухгалтерского учета, в частности бухгалтерский баланс. Статистический отчет об использовании рабочего времени основан на данных табельного учета рабочего времени.

*Опрос* как способ наблюдения основан на получении данных в форме ответов опрашиваемых лиц (респондентов). Это способ используют в

тех случаях, когда исследуемые признаки не поддаются прямому измерению и непосредственному наблюдению.

Опрос используют при проведении переписи населения, социологических обследований, изучении общественного мнения.

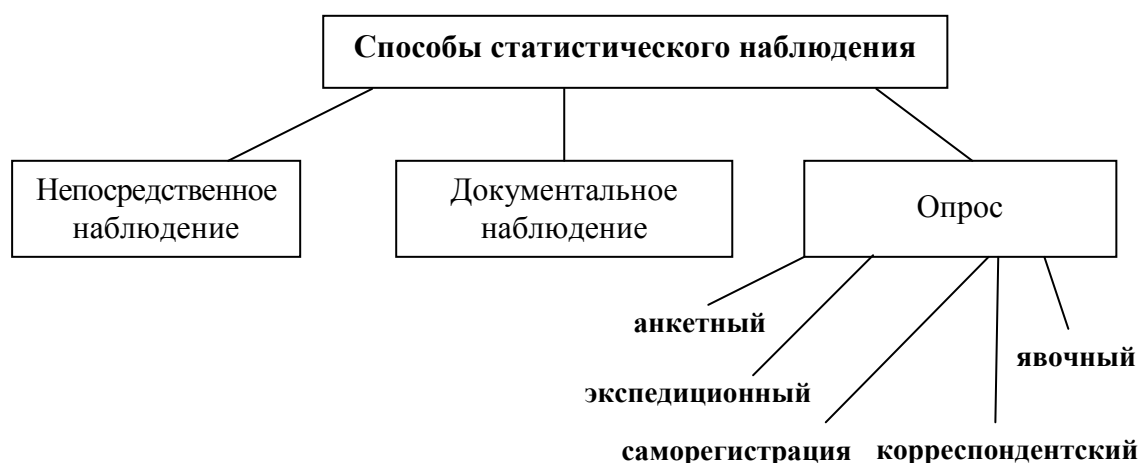


Рис. 2.3. Способы статистического наблюдения  
Источник: [1, с. 69]

Рассмотрим применяемые в статистике виды опросов.

*Явочный способ* – предусматривает представление сведений в органы, ведущие наблюдение в явочном порядке. Например, при регистрации браков, рождений, разводов.

*Экспедиционный способ* (иначе такой опрос называют *устным*) заключается в том, что к каждой единице наблюдения посылаются специальные лица (счетчики, регистраторы), которые в статистических формулярах фиксируют сведения о наблюдаемом явлении. Этот способ реализуется не только в варианте непосредственного личного контакта с опрашиваемым – возможен опрос по телефону.

*Саморегистрация* предполагает, что специальные работники снабжают опрашиваемых бланками и дают инструкции о порядке их заполнения. Заполняют бланки сами же опрашиваемые. Этот способ лежит в основе проведения *бизнес-обследований*. В международной практике этот вид опросов применяется более 40 лет [4, с. 208 – 214]. Первые бизнес-обследования были проведены в Европе, вначале в Германии IFO Институтом (1949 г.), а затем INSEE во Франции и ISCO в Италии. Сегодня бизнес-обследования проводятся более чем в 50 странах мира. В соответствии с методикой бизнес-обследования респондентам предлагается оценить фактическое состояние и ожидаемое изменение основных показателей хозяйственной деятельности в рамках альтернатив «увеличение-уменьшение», «улучшение-ухудшение»,

«осталось на том же уровне». Кроме того, респонденты отвечают на вопросы, связанные с выявлением факторов, ограничивающих деятельность их структур, и на ряд других качественных вопросов.

*Корреспондентский способ* основан на том, что сведения в органы, ведущие наблюдение, сообщает штат добровольных корреспондентов. Этот вид опроса требует наименьших затрат, но не дает уверенности в том, что полученные данные являются высококачественными, т.к. трудно проверить их точность.

*Анкетный способ* состоит в том, что разработанная анкета вручается определенному кругу лиц (или публикуется в периодической печати) и после заполнения возвращается статистическому органу или другим организациям, проводившим исследование.

При проведении анкетирования следует учитывать, что, как правило, возвращается органу, проводившему опрос не более 80 % разосланных анкет, т.к. не все респонденты заинтересованы в опросе. Кроме того, у респондентов могут возникнуть трудности в заполнении анкет, если они сформулированы недостаточно ясно или некорректно.

Укажем самые простые правила разработки анкет:

- в них должна быть определена и четко сформулирована цель исследования;
- вопросы не должны быть двусмысленными и должны соответствовать уровню образованности потенциальных респондентов;
- вопросы следует формировать корректно, чтобы они не вели к возрастанию коэффициента Non-Response Rate («неполучение ответа»);
- формулировать вопросы следует по возможности нейтрально, т.е. избегать формулировок «за» и «против».

Анкетный опрос применяется при изучении общественного мнения, например, о качестве работы городского транспорта, торговли и т.д.

### **Виды статистического наблюдения**

Охарактеризуем *виды статистического наблюдения* на основе их классификации по следующим признакам:

- времени регистрации признаков;
- охвату единиц совокупности.

По времени проведения наблюдение бывает:

- непрерывным (текущим);
- прерывным, которое может быть:
  - периодическим;
  - единовременным.

Рассмотрим указанные виды наблюдения (рис. 2.4).

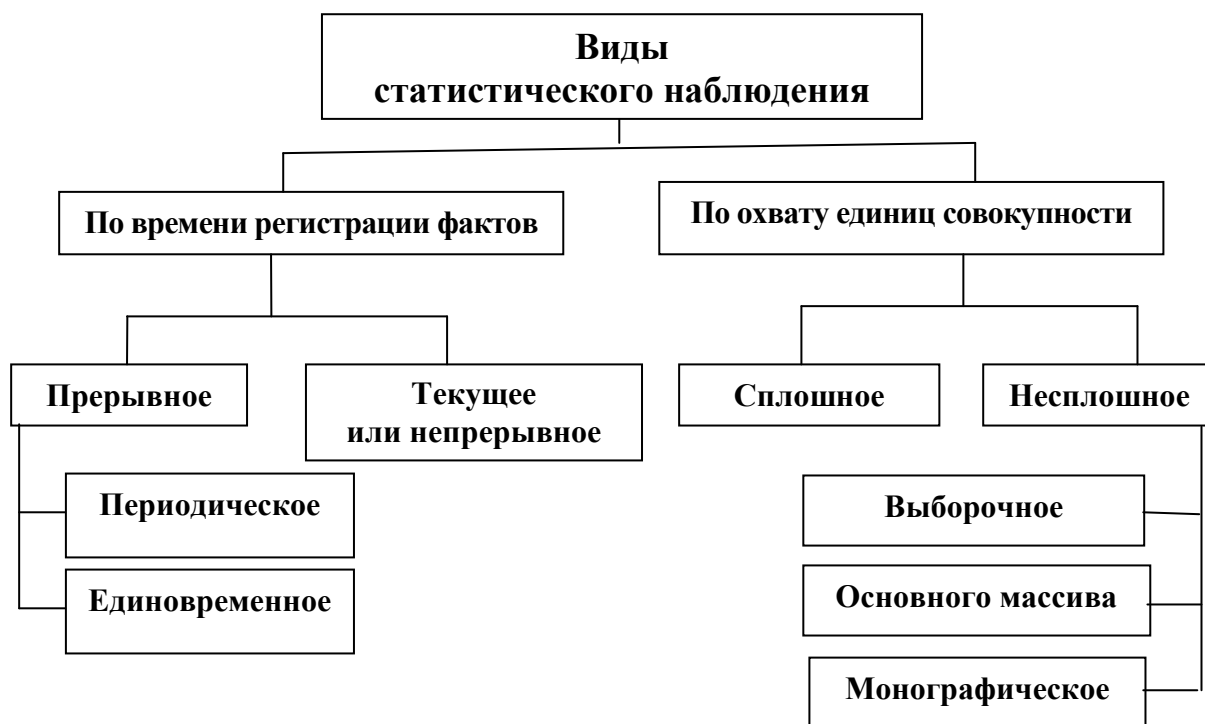


Рис. 2.4. Виды статистического наблюдения  
Источник: [1, с. 71]

*Непрерывное наблюдение* осуществляется путем непрерывной регистрации фактов по мере их возникновения. В этом случае фиксируются все изменения изучаемого процесса или явления. Примерами такого наблюдения являются: запись актов гражданского состояния (рождение, смерть, состояние в браке), учет выработанной продукции, учет отпуска продукции со склада и др.

*Прерывное наблюдение* проводится либо регулярно, через определенные промежутки времени (*периодическое наблюдение*), либо нерегулярно и даже однократно по мере необходимости (*единовременное*).

Примерами периодического статистического наблюдения являются переписи населения, которые проводятся через каждые десять лет; регистрация цен на товары и услуги для ежемесячного определения *индекса потребительских цен*. С определенной периодичностью предприятия и организации представляют в органы государственной статистики формы статистической отчетности. Примером одновременных наблюдений могут служить переписи зеленых насаждений, жилого фонда.

По охвату единиц совокупности наблюдение делится на *сплошное* и *несплошное*.

Сплошное наблюдение обеспечивает полноту информации. Но оно не всегда возможно, например, в случае физического уничтожения и порчи обследованных единиц; когда совокупность велика и на сплошное ее обследо-

ние требуются большие затраты средств и времени. Кроме того, сплошное наблюдение не всегда обеспечивает требуемую точность. Поэтому наряду со сплошным наблюдением широко используется *несплошное наблюдение*.

*Несплошное* наблюдение предполагает, что обследованию подлежит лишь часть единиц изучаемой совокупности. Несплошное наблюдение может проводиться как выборочное, монографическое и наблюдение основного массива (см. рис. 2.4.).

*Выборочным наблюдением* является такое, при котором характеристика всей совокупности (*генеральной совокупности*) дается по некоторой ее части (*выборочной совокупности*), отобранной в случайном порядке. Случайность отбора единиц гарантируется независимостью результатов выборки от воли лиц ее производящих. Таким образом, результат выборки освобождается от тенденциозных ошибок. Возникающие случайные ошибки выборки определяются на основе теорем закона больших чисел (ЗБЧ) и надлежащей организацией наблюдения сводят их к допустимому минимуму. При этом определяется доверительная вероятность возможной ошибки выборки.

*Наблюдение основного массива* предполагает обследование определенных, наиболее существенных по значимости изучаемых признаков единиц совокупности. Например, обследование 15 – 20 % крупных предприятий отрасли выпускающих основной объем (70 – 80 %) определенной продукции.

*Монографическое наблюдение* применяют для глубокого и всестороннего исследования отдельных единиц совокупности, обладающих какими-либо особенностями и являющихся представителями каких-либо, часто новых, типов явлений. Например, изучение деятельности высокорентабельных предприятий, имеющих рентабельность, значительно более высокую, чем в среднем по отрасли. Так как при проведении монографического наблюдения цель изучить всю совокупность не ставится, обеспечивается детальное статистико-монографическое изучение одного предприятия, бюджета семьи и т.д., что позволяет выявить те пропорции и связи, которые ускользают из поля зрения при массовом наблюдении.

В каждом конкретном статистическом исследовании выбор конкретного вида и способа наблюдения будет зависеть:

- от целей и задач наблюдения;
- специфики наблюдаемого объекта;
- установленной срочности получения материалов статистического наблюдения;
- наличия подготовленных регистраторов (статистиков);
- возможности применения технических средств сбора и обработки данных;
- размера выделенных средств.

## 2.6. Ошибки статистического наблюдения и способы контроля точности статистических данных

Точность статистического наблюдения определяется степенью соответствия значения того или иного признака, определенного в результате статистического наблюдения, истинному его значению. Для проверки точности статистического наблюдения его материалы подлежат обязательному контролю по следующим направлениям:

- полнота охвата единиц наблюдения;
- полнота заполнения каждого формуляра;
- правильность заполнения формуляра.

Чтобы качественно провести контроль данных статистического наблюдения, нужно знать характер возможных *ошибок наблюдения*.

*Ошибка наблюдения* – это величина расхождения между результатом наблюдения и истинным значением признака.

В зависимости от причин возникновения выделяют *ошибки регистрации* и *ошибки репрезентативности* (рис. 2.5).

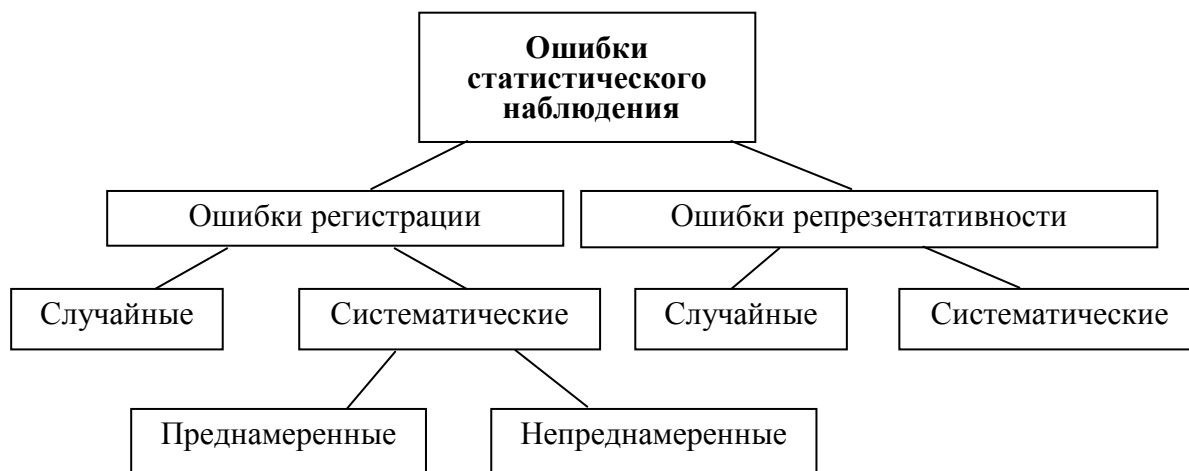


Рис. 2.5. Виды ошибок статистического наблюдения

*Ошибки регистрации* возникают при сплошном и несплошном видах наблюдения. По своей природе ошибки регистрации могут быть *случайными* и *систематическими*.

*Случайные ошибки* возникают в результате действия различных случайных факторов, например, переставлены местами цифры, перепутаны соседние строки или графы при заполнении статистического формуляра. Такие ошибки имеют разную направленность: они могут исказить значе-



ние признака как в сторону увеличения, так и в сторону уменьшения. При достаточно большой обследуемой совокупности в результате действия ЗБЧ эти ошибки взаимопогашаются.

*Систематические ошибки* регистрации, как правило, имеют определенную тенденцию либо к увеличению, либо к уменьшению значений исследуемого признака. Это приводит к формированию накопленной ошибки, которая будет искажать истинное значение признака по совокупности в целом.

Систематические ошибки регистрации могут быть *преднамеренными* и *непреднамеренными*.

Преднамеренная систематическая ошибка регистрации возникает, если лица, проводящие наблюдение, по каким-либо причинам заинтересованы в сокрытии или искажении действительности. Преднамеренные ошибки выражаются чаще всего в виде приписок, сознательного завышения или занижения значений признаков, приукрашивающих результаты производственно-хозяйственной деятельности в целях извлечения выгоды.

Непреднамеренная систематическая ошибка возникает в силу строго определенных причин, однонаправлено искажающих значения признаков. Например, такого рода ошибки возникают в случае неисправности регистрирующих приборов, при неправильном понимании регистратором инструкций по заполнению формуляров.

*Ошибки репрезентативности* возникают при несплошном обследовании в силу того, что состав отобранной для него части единиц совокупности недостаточно полно отображает состав всей исследуемой совокупности, хотя регистрация значений признака по каждой отобранной для обследования единицы была произведена точно.

Поэтому при несплошном обследовании общая ошибка наблюдения является суммой ошибки регистрации и ошибки репрезентативности.

Ошибки репрезентативности также бывают *случайные* и *систематические*.

*Случайные ошибки репрезентативности* объясняются тем, что отобранная совокупность не полностью воспроизводит генеральную совокупность. Эта ошибка оценивается.

*Систематическая ошибка репрезентативности* появляется вследствие нарушения принципов отбора единиц из исходной совокупности.

Более подробно ошибки репрезентативности будут рассмотрены в теме 7 «Выборочное наблюдение».

После получения статистических формуляров проводят *проверку* полноты собранных данных, т.е. определяют, все ли единицы наблюдения заполнили формуляры, значения всех ли признаков отражены в них.

После проверки полноты собранных данных проводят контроль правильности регистрации значений признаков (проверяют правильность заполнения формуляра). С этой целью используют две основные формы контроля: *арифметический* и *логический контроль*.

*Арифметический контроль* основан на использовании количественных связей между значениями различных показателей и состоит в проверке соблюдения этих связей. Например, в ходе арифметического контроля проверяются итоги, балансовые соотношения, функциональные связи между показателями. К примеру, если среди собранных данных имеются сведения о среднесписочной численности промышленно-производственного персонала, выработке продукции в среднем на одного работающего и объеме произведенной продукции, то произведение первых двух показателей должно дать значение третьего показателя. Если арифметический контроль покажет, что данная зависимость не выполняется, это будет свидетельствовать о недостоверности собранных данных.

*Логический контроль* осуществляется путем проверки содержательной связи между значениями признаков по каждой единице наблюдения. Например, человек в возрасте 6 лет не может иметь высшее образование. Поэтому если в бланке переписи имеются одновременно обе записи, это показывает, что одна из них не соответствует действительности. Логический контроль, в отличие от арифметического, непосредственно не устанавливает ошибку наблюдения, а лишь сигнализирует о ее возможности и требует дополнительной связи с единицей наблюдения.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Какое место занимает статистическое наблюдение в процессе статистического исследования?
2. Дайте определение статистического наблюдения.
3. Какова роль статистического наблюдения в получении качественных результатов статистического исследования?
4. Что входит в задачи проектирования статистического наблюдения?
5. Что включает разработка программно-методологического обеспечения статистического наблюдения?
6. Что такое объект и цель статистического наблюдения?
7. Что такое программа статистического наблюдения?

8. Каким требованиям должна отвечать программа статистического наблюдения?
9. Что такое статистический формуляр?
10. Какие организационные вопросы решаются при проектировании статистического наблюдения?
11. Что понимают под критическим временем (моментом или интервалом) статистического наблюдения?
12. Назовите организационные формы статистического наблюдения.
13. Назовите разновидности специально организованного наблюдения.
14. В чем состоит отличие единовременных учетов от переписей?
15. Что такое мониторинг?
16. Что такое регистровое наблюдение?
17. Назовите способы статистического наблюдения.
18. В каких случаях используют такой способ статистического наблюдения, как опрос?
19. Назовите виды опросов.
20. Какие виды статистического наблюдения выделяют в зависимости от времени регистрации признаков?
21. Какие виды статистического наблюдения выделяют в зависимости от охвата единиц совокупности?
22. Назовите виды несплошного наблюдения.
23. По каким направлениям проводят контроль данных статистического наблюдения?
24. Что такое ошибка статистического наблюдения?
25. Назовите виды статистического наблюдения.
26. Какие ошибки наблюдения более опасны: систематические или случайные? Ответ объясните.
27. Для какого вида статистического наблюдения характерны ошибки репрезентативности?
28. Назовите формы контроля правильности регистрации значений признаков.

## Тест 2

1. Статистическое наблюдение ...
  - а) является заключительным этапом статистического исследования;
  - б) не является этапом статистического исследования;
  - в) является первым этапом статистического исследования.
2. Укажите, что характерно для статистического наблюдения:

- а) научно-организованный сбор данных;
- б) спонтанный сбор данных.

3. Проектирование статистического наблюдения предполагает ...

- а) разработку программно-методологического обеспечения;
- б) проведение статистической сводки;
- в) решение организационных вопросов наблюдения;
- г) контроль материалов статистического наблюдения.

4. Объект статистического наблюдения – это ...

- а) единица наблюдения;
- б) статистическая совокупность;
- в) отчетная единица.

5. Перечень признаков единицы наблюдения, подлежащих регистрации в процессе наблюдения называется ...

- а) статистическим формуляром;
- б) списком;
- в) программой наблюдения;
- г) статистическим инструментарием.

6. Статистический формуляр содержит ...

- а) описание процесса статистического наблюдения;
- б) программу статистического наблюдения;
- в) программу и результаты статистического наблюдения.

7. Критический момент (дата) статистического наблюдения устанавливается с целью ...

- а) получения сопоставимых данных;
- б) планирования загрузки статистиков;
- в) учета особенностей наблюдаемых совокупностей.

8. Срок наблюдения – это ...

- а) период, в течение которого происходит заполнение формуляров;
- б) момент, по состоянию на который должна быть проведена регистрация признаков по каждой единице наблюдения.

9. Статистическая отчетность – это ...

- а) вид статистического наблюдения;
- б) организационная форма статистического наблюдения;
- в) способ статистического наблюдения.

10. Статистическое наблюдение, которое основано на использовании в качестве источника информации данных различных документов идентифицируется как ...

- а) непосредственное наблюдение;
- б) документальное наблюдение;
- в) мониторинг.

11. Метод основного массива – это ...

- а) организационная форма статистического наблюдения;
- б) вид статистического наблюдения;
- в) способ статистического наблюдения.

12. Перепись населения страны – это ...

- а) единовременное, специально организованное, сплошное наблюдение;
- б) периодическое, специально организованное, сплошное наблюдение;
- в) периодическое, регистровое, несплошное наблюдение.

13. Величина расхождения между результатом наблюдения и истинным значением признака называется ...

- а) ошибкой репрезентативности;
- б) ошибкой наблюдения;
- в) расчетной ошибкой;
- г) случайной ошибкой.

14. Ошибки регистрации возникают ...

- а) только при сплошном наблюдении;
- б) только при выборочном наблюдении;
- в) при сплошном и несплошном наблюдении.

### **Задания по теме**

#### **Задание 2.1**

Какие бы вы наметили признаки, которые следовало зарегистрировать при проведении:

- а) обследования промышленного предприятия с целью изучения текучести кадров;
- б) обследования работы городского транспорта с целью изучения роли различных его видов в перевозках пассажиров;
- в) обследования студентов университета с целью изучения бюджета времени?

#### **Задание 2.2**

Определите цель, объект, единицу наблюдения и разработайте программу:

- а) статистического обследования автозаправочных станций;
- б) статистического обследования школ города;
- в) статистического обследования студентов (одного из курсов, факультета, одной из специальностей).

### **Задание 2.3**

Составьте анкету опроса студентов в целях выяснения:

- а) их состава по полу, возрасту, семейному положению, успеваемости, научным интересам;
- б) их оценки качества преподавания отдельных дисциплин;
- в) их бытовых условий и материального положения;
- г) бюджета их времени и использования свободного времени.

### **Задание 2.4**

Торговая сеть магазинов «Дионис» поручает вам разработать анкету опроса покупателей с целью изучения покупательского спроса и качества обслуживания. Укажите, к какому виду относится данное наблюдение по времени, охвату единиц и способу получения данных.

### **Задание 2.5**

На оптовую торговую базу поступила партия товара. Для проверки его качества была отобрана в случайном порядке десятая часть партии, и путем тщательного осмотра каждой единицы товара определялось и фиксировалось его качество. К какому виду наблюдения можно отнести это обследование?

### **Задание 2.6**

Производится статистическое наблюдение. Ответы на вопросы формуляра записываются на основании документов, содержащих необходимые сведения. Укажите способ проведения данного статистического наблюдения.

### **Задание 2.7**

С помощью логического контроля проверьте следующие ответы на вопросы переписного листа переписи населения:

- а) фамилия, имя, отчество – Иванова Ирина Петровна;
- б) пол – мужской;
- в) возраст – 5 лет;
- г) состоит ли в браке в настоящее время – да;
- д) образование – среднее специальное;
- е) место работы – детский сад;
- ж) занятие по этому месту работы – медицинская сестра.

В ответах на какие вопросы вероятнее всего произведены ошибочные записи?

### Тема 3

## СВОДКА И ГРУППИРОВКА СТАТИСТИЧЕСКИХ ДАННЫХ. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

- 3.1. Значение и содержание статистической сводки.
- 3.2. Сущность и виды статистических группировок.
- 3.3. Определение числа групп и величины интервалов при группировке по количественным признакам.
- 3.4. Виды и правила построения статистических таблиц.

Изучив данную тему, студенты должны:

- понимать сущность и значение статистической сводки в проведении статистического исследования;
- знать виды и правила построения статистических группировок;
- уметь представлять результаты статистической сводки и группировки в табличной форме.

### 3.1. Значение и содержание статистической сводки

В результате статистического наблюдения (первый этап статистического исследования) получают данные о каждой единице исследуемого объекта. Полученные данные не являются обобщающими показателями и на их основе нельзя сделать выводы в целом по совокупности. Эту задачу решает второй этап статистического исследования – *статистическая сводка*.

*Цель статистической сводки* – систематизация первичных данных и получение на этой основе сводной характеристики всего объекта при помощи обобщающих статистических показателей.

*Сводка это научно организованная обработка материалов наблюдения для выявления типичных черт и закономерностей, присущих изучаемой совокупности в целом.*

Таким образом, если при статистическом наблюдении собирают данные о каждой единице объекта, то результатом сводки являются обобщающие данные, характеризующие всю совокупность в целом.

Выделяют несколько видов статистической сводки на основе ее классификации по ряду признаков (рис. 3.1).

*По глубине обработки данных сводка может быть простой и сложной.*

*Простой* называется сводка, при которой совокупность не разбивается на группы и проводится только операция подсчета общих итогов по совокупности единиц или общего объема изучаемого признака.



Рис. 3.1. **Виды статистической сводки**  
 Источник: [1, с. 82]

Чаще простую сводку применяют в оперативной отчетности для осуществления оперативного управления процессами, неблагоприятное или недостаточно активное протекание которых требует вмешательства. Например, при уборке урожая, учете выработанной продукции и т.п.

*Сложной* называется сводка, в которой применяется статистическая группировка.

Проведению сложной сводки предшествует разработка ее программы, которая включает:

- выбор группировочных признаков;
- определение порядка формирования групп;
- разработку системы показателей для характеристики групп и объекта в целом;
- разработку системы макетов статистических таблиц, в которых должны быть представлены результаты сводки.

По форме обработки данных сводка может быть *централизованной* и *децентрализованной*.

При *децентрализованной сводке* обработка материала производится последовательными этапами. Так, отчеты предприятий по ряду форм статистической отчетности сводятся районными и городскими отделами статистики, затем их обработка осуществляется областными управлениями, а уже итоги по областям поступают в Национальный статистический комитет Республики Беларусь.

При *централизованной сводке* весь первичный материал поступает в одну организацию, где и подвергается обработке от начала до конца. Цен-



трализованная сводка обычно используется для обработки материалов единовременных статистических обследований.

По *технике исполнения* сводка может быть *компьютерной* и *ручной*.

При *компьютерной сводке* все операции по обработке первичных данных осуществляются с использованием компьютеров и программных продуктов, позволяющих обрабатывать любые объемы информации с различной степенью детализации. Операции сложной статистической сводки удобно проводить с помощью программного продукта SPSS (Statistical Package for Social Sciences) [5].

При *ручной сводке* все основные операции осуществляются вручную. Ручную сводку используют при малых объемах информации и незначительной глубине разработки статистического материала.

При проведении сводки составляется план, в котором определяются организационные вопросы:

- кто и в какие сроки осуществляет сводку;
- каким способом, куда поступают сводные данные;
- кто проводит их дальнейшую обработку;
- состав сведений, подлежащих публикации в периодической печати.

### **3.2. Сущность и виды статистических группировок**

#### **Сущность статистической группировки**

*Статистическая группировка – расчленение множества единиц изучаемой совокупности по одному или нескольким существенным признакам на однородные группы и характеристика их системой обобщающих показателей.*

Метод группировки позволяет представить данные, полученные в ходе статистического наблюдения в более сжатой форме и на этой основе выявить закономерности, присущие статистической совокупности.

В статистической методологии группировка выполняет две основные функции:

во-первых, группировка является *самостоятельным методом анализа* социально-экономических явлений;

во-вторых, она *предопределяет возможности использования других статистических методов*: метода средних величин, дисперсионного анализа, корреляционно-регрессионного анализа. Применение указанных методов предполагает наличие однородных совокупностей. Несоблюдение требования однородности делает применение этих методов некорректным и не позволяет получить надежные и достоверные результаты, пригодные для принятия обоснованных решений.

Применение метода статистической группировки основано на определении трех основных моментов:

- группировочного признака;
- числа групп;
- размера интервала (для количественных признаков).

*Группировочный признак – это существенный признак, в зависимости от значений которого происходит объединение единиц совокупности в однородные группы.* Иначе группировочный признак называют *основанием группировки.*

Группировка может быть выполнена по количественному (например, группировка, представленная в табл. 3.2) и качественному (см. табл. 3.1) признакам.

После определения основания группировки следует решить вопрос о числе групп, на которые надо разбить исследуемую совокупность. Число групп зависит от задач исследования и вида признака, положенного в основание группировки, численности совокупности, степени вариации признака.

При построении группировок по качественному признаку число групп соответствует количеству возможных значений признака. При группировке по количественному признаку все множество значений делится на интервалы. Более подробно порядок определения числа групп и размера интервалов рассматривается в подразделе 3.3 данной темы.

### **Виды статистических группировок**

Классификация статистических группировок может быть проведена по следующим признакам (рис. 3.2):

- задачи обобщения или систематизации статистических данных;
- число группировочных признаков;
- степень упорядоченности исходной информации.

Таблица 3.1

#### **Структура промышленного производства региона по формам собственности в 2011 г.**

	Удельный вес, %	
	в общем числе предприятий и производств	в общем объеме промышленной продукции
Вся промышленность,	100	100
в т.ч. по формам собственности:		
государственная	33,5	22,4
частная	64,8	76,7
иностранная	1,7	0,9

Таблица 3.2

**Распределение кредитного портфеля банка  
по группам риска на конец отчетного периода, млн руб.**

Количество дней просроченной задолженности	Обеспеченность ссуды	Группы ссуд по степени риска					Итого по группе
		стандартные	нестандартные	сомнительные	опасные	безнадежные	
Возврат ссуды в срок	обеспеченная	62					80
	недостаточно обеспеченная	11					
	необеспеченная	7					
До 30	обеспеченная	18					24
	недостаточно обеспеченная		4				
	необеспеченная			2			
30 – 60	обеспеченная		10				15
	недостаточно обеспеченная			3			
	необеспеченная				2		
60 – 90	обеспеченная			9			17
	недостаточно обеспеченная				7		
	необеспеченная					1	
Более 90	обеспеченная					1	4
	недостаточно обеспеченная					2	
	необеспеченная					1	
Итого по группе		98	14	14	9	5	140

Охарактеризуем виды статистических группировок в зависимости от задачи обобщения исходных данных. Как самостоятельный метод статистического исследования группировка решает три задачи:

- выделяет в совокупности качественно однородные группы (социально-экономические типы, классы);
- характеризует структуру совокупности и на этой основе позволяет оценивать структурные сдвиги, происходящие в ней;
- изучает связи и зависимости между признаками исследуемого явления.

В соответствии с этими задачами систематизации исходных данных группировки подразделяют на три вида:

- типологические;
- структурные;
- аналитические.

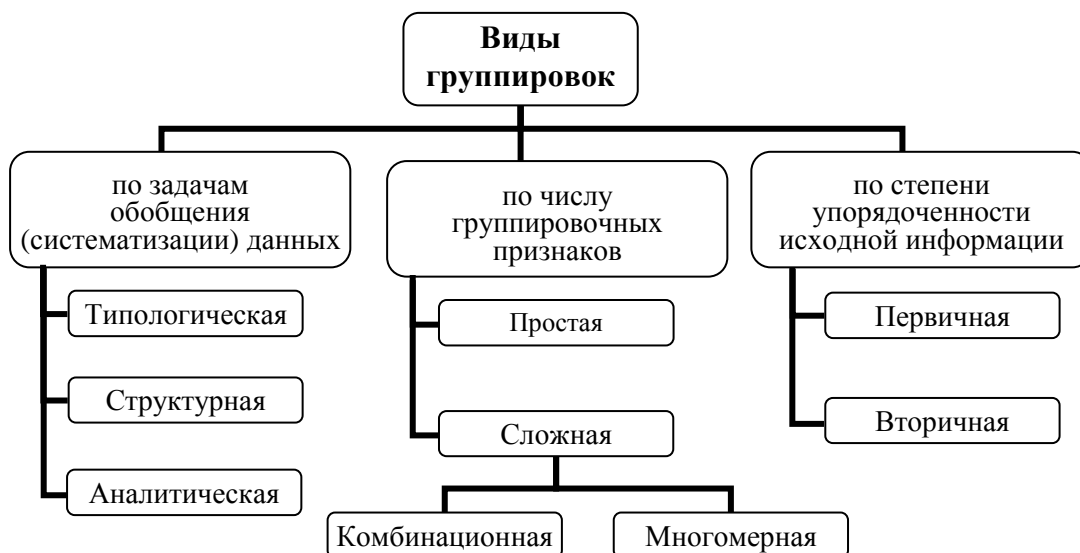


Рис. 3.2. Классификация видов статистических группировок  
Источник:[1, с. 85]

Рассмотрим их сущность.

### Типологическая группировка

*Типологическая группировка характеризуется выделением типов явлений, однородных по значениям существенных признаков и условиям развития, в которых действуют одни и те же закономерности влияния факторов.*

Схематично типологическую группировку представляет рис. 3.3.

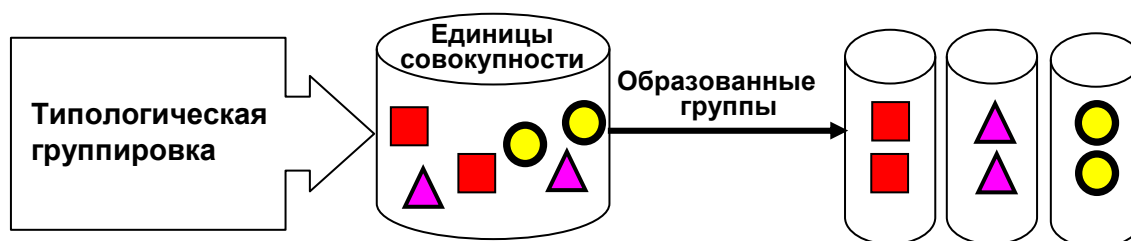


Рис. 3.3. Разделение разнородной совокупности на однородные группы  
Источник: [1, с. 85]

Примерами типологической группировки являются: группировка предприятий по формам собственности, финансовой устойчивости и др. К типологическим относятся группировки, представленные в табл. 3.1. и 3.2.

Этот вид группировки широко используется в маркетинге, в частности, при исследовании покупательского спроса, изучении товарного рынка, проведении его сегментации.

Типологическую группировку применяют, если нужно охарактеризовать качественные особенности отдельных групп совокупности.

При типологической группировке выбор группировочных признаков осуществляется на основе содержательного анализа изучаемой совокупности. Задачу построения типологической группировки трудно формализовать, здесь особенно важны теоретические и профессиональные знания исследователя.

Типы явлений могут быть выделены на основе градации значений одного признака (например, группировка предприятий по форме собственности, представленная в табл. 3.1). Но чаще своеобразие типов раскрывается с помощью нескольких признаков. Например, для выделения типов ссуд по степени риска используют два признака: продолжительность срочности возврата и обеспеченность ссуды (см. табл. 3.2).

Границы интервалов при группировке по количественному признаку устанавливаются таким образом, чтобы обозначить переход к новому качеству (типу). Чаще при типологической группировке применяют *неравные интервалы*, а также *открытые интервалы*.

Для типологической группировки характерно применение *специализированных интервалов* и *специализированных признаков*. Это означает, что в различных конкретных условиях места и времени для выделения одних и тех же типов используют разные признаки и разные интервалы.

Экспертный характер построения типологической группировки делает обязательным проверку ее правильности. Цель такой проверки состоит в том, чтобы доказать существенность различий выделенных типов явлений. Самым простым является неформальный критерий правильности построения типологической группировки. В соответствии с этим критерием качественные различия между выделенными типами признаются существенными, если им можно дать словесное (содержательное) наименование. Например, при группировке населения по уровню доходов можно выделить такие типы (группы): нищета, бедность, нуждаемость, относительный достаток, зажиточность, состоятельность. Другой пример: Ф. Котлер, характеризуя потребителей пива, выделил четыре их типа: пьющие в компании; пьющие для тонуса; пьющие много и пьющие бесконтрольно.

Формальная количественная оценка типологической группировки может быть проведена с помощью статистических критериев (*t*-критерия Стьюдента, *F*-критерия Фишера).

## **Классификации и классификаторы**

Типологическая группировка является основным приемом решения задачи классификации.

*Классификация – систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основании их сходства и различия.*

Отличительные черты классификации:

- в основе классификации лежит качественный признак;
- классификации стандартны: они устанавливаются органами государственной и международной статистики. Если в каждом конкретном исследовании строится своя группировка, то классификация едина для любого исследования независимо от того, проводят его органы государственной статистики или другие учреждения и ведомства (министерства, налоговые органы и т.п.);
- классификации устойчивы. Они остаются неизменными в течение длительного времени. Однако если появляются новые группы единиц, их классы, разряды, то в классификации вносятся соответствующие изменения и дополнения.

В классификации точно определены все возможные группы и имеются подробные указатели, которые помогают отнести любую единицу объекта в ту или иную группу в каждом конкретном случае.

Таким образом, *цель классификации* – однозначно идентифицировать единицы совокупности, обеспечить эффективный поиск информации и ее систематизацию, достичь сопоставимости с международными стандартами.

С категорией «классификация» тесно связана категория «классификатор». *Классификатор – классификация, в которой каждому значению группировочного признака присвоен код, т.е. условное цифровое обозначение.*

Использование кодов значительно упрощает процессы сбора, обработки и передачи информации, особенно в условиях применения ЭВМ.

Система классификаторов, используемых статистикой, включает:

- международные классификаторы, разработанные ООН;
- классификаторы ЕС, обязательные к исполнению странами-членами ЕС;
- национальные классификаторы, учитывающие особенности конкретной страны.

Примерами международных классификаторов являются: *Международная стандартная отраслевая классификация видов экономической деятельности* (МСОК, ISIC), *Международная Центральная классификация продукции* (СРС). В качестве национальных классификаторов Республики Беларусь можно указать *Общегосударственный классификатор про-*

фессий рабочих и должностей служащих; Общегосударственный классификатор форм собственности (ОКФС); Общегосударственный классификатор организационно-правовых форм (ОКОПФ); Общегосударственный классификатор видов экономической деятельности (ОКЭД) и др.

### **Структурная группировка**

*Структурной называется группировка, в которой однородная в качественном отношении совокупность расчленяется на группы, характеризующие ее структуру.*

Структурные группировки могут быть проведены на основе одного или нескольких атрибутивных или количественных признаков.

Примером структурной группировки по количественному признаку является группировка промышленных предприятий по среднесписочной численности промышленно-производственного персонала, представленная в табл. 3.3.

Таблица 3.3

**Группировка промышленных предприятий региона по среднесписочной численности промышленно-производственного персонала за 2011 г. (в процентах к итогу)**

	Число предприятий	Объем продукции (работ, услуг) в фактических ценах предприятий
Все предприятия, учитываемые в текущем порядке	100	100
в т.ч. с численностью промышленно-производственного персонала, чел.:		
до 50	17,0	0,6
50 – 100	21,8	2,3
100 – 200	26,5	6,5
200 – 300	10,2	4,8
300 – 500	8,7	7,3
500 – 1 000	7,6	13,0
1 000 – 2 000	4,5	11,3
2 000 – 3 000	1,8	10,4
3 000 – 5 000	0,9	8,7
5 000 – 10 000	0,7	16,7
10 000 и более	0,3	18,4

Если Вы обратите внимание на группировку промышленного производства региона по формам собственности (см. табл. 3.1), то увидите, что она наряду с выделением типов предприятий и производств по формам

собственности характеризует и структуру промышленного производства по этому признаку. Другими словами, группировка, представленная в табл. 3.1. сочетает черты типологической и структурной группировок.

На основе структурной группировки можно изучить динамику структуры совокупности, если известны структурные характеристики совокупности в разные периоды (моменты времени). Для оценки структурных сдвигов используют следующие показатели, оценивающие изменения структуры:

а) на основе формулы среднего линейного отклонения:

$$\bar{d}_{w_1-w_0} = \frac{\sum_{i=1}^k |w_{i_1} - w_{i_0}|}{k}, \quad (3.1)$$

где  $w_{i_0}$ ,  $w_{i_1}$  – доли  $i$ -группы в базисном и анализируемом периодах;

$k$  – число групп;

б) на основе формулы среднего квадратического отклонения:

$$\sigma_{w_1-w_0} = \sqrt{\frac{\sum_{i=1}^k (w_{i_1} - w_{i_0})^2}{k}}. \quad (3.2)$$

Если структурных сдвигов нет, то  $\bar{d}_{w_1-w_0}$  и  $\sigma_{w_1-w_0}$  равны нулю. Их величина будет тем больше, чем значительнее структурные сдвиги.

К структурным группировкам близко примыкают *ряды распределения* единиц совокупности по варьирующим признакам. Однако задачи анализа рядов распределения иные – раскрыть закономерности в характере распределения единиц совокупности. Методы анализа рядов распределения рассматриваются в теме 6.

### **Аналитическая группировка**

*Аналитическая группировка – это группировка, которая позволяет установить наличие и направленность связи между анализируемыми существенными признаками изучаемой статистической совокупности.*

Основу проведения аналитической группировки составляет содержательный анализ связей между анализируемыми признаками исследуемого явления. Результатом такого анализа должно стать установление *факторных* и *результативного признаков* (см. тему 1).

Особенностью аналитической группировки является то, что в основание группировки закладывается факторный признак, затем подсчи-



тывается количество единиц совокупности и общее суммарное значение результативного признака по каждой выделенной группе и дальше производится расчет среднего значения результативного признака по выделенным группам. Взаимосвязь проявляется в том, что с возрастанием (убыванием) значения факторного признака систематически возрастает (убывает) среднее значение результативного признака. Результаты группировки излагаются в статистической таблице. Пример аналитической группировки представлен в табл. 3.4.

Таблица 3.4

**Группировка коммерческих банков по сумме активов баланса**  
(данные условные)

Группы банков по сумме активов баланса, млн руб.	Количество банков, ед.	В среднем на один банк	
		численность занятых, чел.	прибыль, млрд руб.
До 20 000	19	184	22,5
20 000 – 30 000	8	313	31,6
30 000 – 40 000	7	374	36,0
40 000 – 50 000	9	468	69,2
50 000 и более	7	516	205,6

Данные табл. 3.4 показывают, что с ростом суммы активов банка численность занятых в среднем на один банк и средняя прибыль на один банк увеличиваются. Это указывает на то, что между анализируемыми признаками существует прямая связь, которая проявляется устойчиво, т.к. отмечается в каждой группе без исключения.

**Пример 3.1**

Рассмотрим пример проведения структурной и аналитической группировок.

Имеются следующие данные о рабочих производственного участка (табл. 3.5).

Построим ряд распределения рабочих по стажу, образовав пять групп с равными интервалами, с целью выявления зависимости выработки рабочего от стажа его работы.

Необходимо построить аналитическую группировку. *Факторный признак* – стаж работы, влияющий на значение выработки рабочего; *результативный* – выработка, значение которой зависит от стажа работы. Следовательно, группировку производим по стажу рабочих, для чего вычислим ширину равного интервала:

$$h = \frac{X_{\max} - X_{\min}}{k} = \frac{16 - 1}{5} = 3 \text{ (года)}.$$

Таблица 3.5

**Данные о рабочих малого предприятия**

№	Стаж работы, лет	Годовая выработка рабочего, тыс. шт.	№	Стаж работы, лет	Годовая выработка рабочего, тыс. шт.
1	1,0	200	16	10,5	276
2	1,0	202	17	1,0	234
3	3,0	205	18	9,0	270
4	6,5	290	19	9,0	264
5	9,2	298	20	6,5	252
6	4,4	250	21	5,0	241
7	6,9	280	22	6,0	256
8	2,5	230	23	10,1	262
9	2,7	223	24	5,5	245
10	16,0	310	25	2,5	240
11	13,2	284	26	5,0	244
12	14,0	320	27	5,3	252
13	11,0	295	28	7,5	253
14	12,0	279	29	7,0	252
15	4,5	222	30	8,0	262

Образует группы рабочих с равным интервалом  $h$ : от 1 до 4 лет; от 4 до 7 лет и т.д. В каждый интервал попадают рабочие со стажем, равным значению нижней границы «от», и не включаются рабочие со стажем, совпадающим с верхней границей интервала «до».

По каждой группе подсчитаем численность рабочих и оформим результаты в виде таблицы (табл. 3.6).

Таблица 3.6

**Группировка рабочих по стажу**

Номер группы	Группы рабочих по стажу, лет	Число рабочих, чел.	Число рабочих, % к итогу
I	1 – 4	7	23,3
II	4 – 7	10	33,3
III	7 – 10	6	20,0
IV	10 – 13	4	13,3
V	13 – 16	3	10,1
Итого	–	30	100,0

Имеем первичную простую структурную группировку, представленную вариационным рядом распределения с равными интервалами. Результаты группировки показывают, что более половины рабочих (53,3 %) име-

ют стаж работы от 4 до 10 лет. Равное число рабочих имеет стаж до 4 лет и свыше 10 лет, удельный вес которых равен 23,3 %. Для построения аналитической группировки заполним вспомогательную таблицу, указав рядом с факторным результативный признак (выработку) (табл. 3.7).

Таблица 3.7

**Вспомогательная таблица для построения аналитической группировки**

Номер группы	Группы рабочих по стажу, лет	Номер рабочего	Стаж, лет	Годовая выработка, тыс. шт.
I	1 – 4	1	1,0	200
		2	1,0	202
		3	3,0	205
		8	2,5	230
		9	2,7	223
		17	1,0	234
		25	2,5	240
	Итого	7	13,7	1 534
II	4 – 7	6	4,4	250
		4	6,5	290
		15	4,5	222
		20	6,5	252
		21	5,0	241
		22	6,0	256
		24	5,5	245
		26	5,0	244
		27	5,3	252
		7	6,9	280
	Итого	10	55,6	2 532
III	7 – 10	5	9,2	298
		18	9,0	270
		19	9,0	264
		28	7,5	253
		29	7,0	252
		30	8,0	262
	Итого	6	49,7	1599
IV	10 – 13	13	11,0	295
		14	12,0	279
		16	10,5	276
		23	10,1	262
	Итого	4	43,6	1 112
V	13 – 16	10	16,0	310
		11	13,2	284
		12	14,0	320
	Итого	3	43,2	914
	Всего	30	205,8	7 691

По итоговым данным определим средний стаж и выработку рабочих в каждой группе:

$$\text{Средний стаж} = \frac{\text{Стаж всех работающих}}{\text{Число рабочих в группе}}$$

$$\text{Средняя выработка} = \frac{\text{Выработка всех работающих}}{\text{Число рабочих в группе}}$$

$$\text{стаж}_1 = \frac{13,7}{7} = 1,96 \text{ (лет)}, \quad (\text{выработка}_1 = \frac{1534}{7} = 219,1 \text{ (тыс. руб.)},$$

$$\text{стаж}_2 = \frac{55,6}{10} = 5,56 \text{ (лет)}, \quad \text{выработка}_2 = \frac{2532}{10} = 253,2 \text{ (тыс. руб.)},$$

$$\text{стаж}_3 = \frac{49,7}{6} = 8,28 \text{ (лет)}, \quad \text{выработка}_3 = \frac{1599}{6} = 266,5 \text{ (тыс. руб.)},$$

$$\text{стаж}_4 = \frac{43,6}{4} = 10,9 \text{ (лет)}, \quad \text{выработка}_4 = \frac{1112}{4} = 278,0 \text{ (тыс. руб.)},$$

$$\text{стаж}_5 = \frac{43,2}{3} = 14,4 \text{ (лет)}, \quad \text{выработка}_5 = \frac{914}{3} = 304,7 \text{ (тыс. руб.)}.$$

Средний стаж и выработка рабочих по совокупности в целом:

$$\text{стаж} = \frac{205,8}{30} = 6,86 \text{ лет}, \quad \text{выработка} = \frac{7691}{30} = 256,4 \text{ тыс. руб.}$$

Строим аналитическую простую группировку (табл.3.8).

Таблица 3.8

#### Группировка рабочих по стажу работы

Номер группы	Группы рабочих по стажу, лет	Число рабочих, чел.	Средний стаж, лет	Годовая выработка, тыс. шт.	
				всего	на одного рабочего
А	В	1	2	3	4
I	1 – 4	7	1,96	1 534	219,1
II	4 – 7	10	5,56	2 532	253,2
III	7 – 10	6	8,28	1 599	266,5
IV	10 – 13	4	10,9	1 112	278,0
V	13 – 16	3	14,4	914	304,7
Итого	–	30	6,86	7 691	256,4

Сравнивая гр. 2 и гр. 4 табл. 3.8, видим, что с увеличением стажа рабочих растет годовая выработка продукции. Следовательно, между изучаемыми признаками имеется прямая связь, которая проявляется устойчиво, т.к. нет ни одного исключения в выявленной закономерности.

Деление рассмотренных группировок, в зависимости от решаемой задачи систематизации данных, на три вида носит, как уже отмечалось выше, условный характер, поскольку группировка, выделяя типы, может одновременно показывать структуру совокупности и отражать закономерности изменения значений одного признака в зависимости от другого.

### **Простые и сложные группировки**

Рассмотрим виды статистических группировок в зависимости от *числа группировочных признаков*. С этой точки зрения различают *простые* группировки (один признак) и *сложные* (два и более признака). Примером простых группировок являются группировки, приведенные в табл. 3.1, 3.3, 3.4.

Сложные группировки, в свою очередь, делятся на *комбинационные* (два – четыре признака) и *многомерные* (любое число признаков свыше четырех).

Принцип построения *комбинационной* группировки заключается в том, что сначала группы формируются по одному признаку, затем они делятся на подгруппы по другому признаку, эти, в свою очередь, делятся по третьему и т.д. Такая группировка позволяет изучить единицы совокупности одновременно по нескольким признакам. При построении комбинационной группировки возникает вопрос о последовательности разбивки единиц объекта по признакам. Как правило, рекомендуется сначала производить группировку по атрибутивным признакам, значения которых имеют ярко выраженные качественные различия, а затем дополнять ее группировкой по количественным признакам.

Комбинационные группировки применяются, как правило, при изучении сложных социально-экономических явлений и процессов. Необходимым и обязательным условием построения данного вида группировок является наличие достаточно большого числа наблюдений. Дело в том, что комбинация группировочных признаков приводит к резкому увеличению числа групп. Численность же единиц в каждой из них может оказаться недостаточной. В результате исследователь может прийти к недостаточно обоснованным выводам. Поэтому на практике строят комбинационные группировки не более чем по трем признакам.

Примером комбинационной группировки являются данные, приведенные в табл. 3.2.

Более сложным видом группировок являются *многомерные группировки*. В них группировка осуществляется не последовательно по группировочным признакам, а одновременно по всему комплексу признаков.

Существуют различные методы проведения многомерной группировки, основными из которых являются:

- расчет многомерной средней;
- кластерный анализ (от англ. the cluster – пучок, группа, куст).

Эти методы получили распространение благодаря использованию компьютеров и программных продуктов, позволяющих разрабатывать любые объемы информации с различной степенью детализации. Многомерные группировки позволяют решать целый ряд важных задач экономико-статистического исследования, таких, как формирование однородных совокупностей, выбор существенных признаков, выделение типичных групп объектов по множеству существенных признаков и др.

### **Первичная и вторичная группировки**

В зависимости от степени упорядоченности исходных данных группировки могут быть *первичными* и *вторичными*.

*Первичная группировка* строится на основе первичных данных, полученных в результате статистического наблюдения, т.е. формируется на основе неупорядоченной исходно информации.

Если группировка производится на основе систематизированных в результате первичной группировки данных, то она является *вторичной*.

Необходимость проведения вторичной группировки возникает в следующих случаях:

- когда первичная группировка не удовлетворяет целям исследования в отношении числа групп и их заполненности;
- когда необходимо провести сравнительный анализ группировок, имеющих разные интервалы и разное количество групп.

Применяют два основных способа проведения вторичной группировки. Первым, наиболее простым, способом является изменение, чаще укрупнение, первоначальных интервалов. Второй способ получил название долевой перегруппировки и состоит в образовании новых групп на основе закрепления за каждой группой определенной доли единиц совокупности.

### **Пример 3.2**

Рассмотрим методику проведения вторичной группировки. Имеются данные о распределении населенных пунктов по числу сельскохозяйственных дворов [4, с. 267 – 269] (табл. 3.9).

Сравним структуру населенных пунктов по числу дворов. Приведенные данные представляют первичную группировку. Они не позволяют провести сравнение распределения населенных пунктов в двух районах по чис-

лу дворов, т.к. в каждом из них имеется различное число групп населенных пунктов. Необходимо анализируемые группировки провести к сопоставимому виду. За основу сравнения возьмем группировку населенных пунктов первого района. Тогда по второму району надо произвести вторичную группировку или перегруппировку населенных пунктов, образовав такое же число групп с теми же интервалами, что и в первом районе. В результате перегруппировки получим следующие сопоставимые данные (табл. 3.10).

Таблица 3.9

**Распределение населенных пунктов по числу дворов**

I район			II район		
Группы пунктов	Число дворов	Удельный вес группы, %	Группы пунктов	Число дворов	Удельный вес группы, %
I	до 100	4,3	I	до 50	1,0
II	100 – 200	18,3	II	50 – 70	1,0
III	200 – 300	19,5	III	70 – 100	2,0
IV	300 – 500	28,2	IV	100 – 150	10,0
V	Свыше 500	29,5	V	150 – 250	18,0
–	–	–	VI	250 – 400	21,0
			VII	400 – 500	23,0
			VIII	Свыше 500	24,0
Итого	–	100,0	–	Итого	100,0

Таблица 3.10

**Распределение населенных пунктов по числу дворов в районах области**

Номер группы	Группы населенных пунктов, по числу дворов	Удельный вес населенных пунктов группы, % к итогу	
		I район	II район
I	До 100	4,3	4,0
II	100 – 200	18,3	19,0
III	200 – 300	19,5	16,0
IV	300 – 500	28,2	37,0
V	Свыше 500	29,7	24,0
Итого	–	100,0	100,0

Поясним расчеты. В первую, вновь образованную группу населенных пунктов второго района с числом дворов до 100, войдут первые три группы населенных пунктов, удельный вес которых равен 4 % (1 + 1 + 2). Теперь надо образовать вторую группу населенных пунктов с числом дворов от 100 до 200. В нее входит четвертая группа населенных пунктов с числом дворов от 100 до 150, составляющая 10 % общего числа населенных пунктов, а также часть пятой группы, из которой должны перейти 50 дворов. Для определения числа населенных пунктов, которые надо взять из

пятой группы во вновь образованную, условно принимается, что оно должно быть пропорционально доле отобранных дворов. Доля 50 дворов в пятой группе равна

$$\frac{50}{250 - 150} = \frac{50}{100} = 0,5 \text{ или } 50 \%.$$

Следовательно, в новую группу надо взять половину населенных пунктов из пятой группы:

$$\frac{50 \cdot 18}{100} = 9 \%.$$

Таким образом, удельный вес дворов новой группы с числом дворов 100 – 200 составит 19 % (10 + 9).

Формируем группы населенных пунктов с числом дворов 200 – 300. В нее войдет часть пятой группы с удельным весом населенных пунктов 9 % и часть шестой группы, из которой надо добавить в третью группу 50 дворов, пропорционально которым должно быть отобрано из этой группы 7 % населенных пунктов:

$$\frac{50}{400 - 250} \cdot 21 = 7 \%.$$

Тогда группа населенных пунктов с числом дворов 200 – 300 составит 16 % (9 + 7).

Аналогично проводят расчет при образовании других групп. Таким образом, сравнивая оба района по числу дворов в населенных пунктах, видим, что во втором районе оно было более дифференцировано, чем в первом.

### **3.3. Определение числа групп и величины интервалов при группировке по количественным признакам**

При группировке по качественному признаку число групп определяются числом значений группировочного признака.

При группировке по *количественному признаку* необходимо определить *число групп* и *размер интервалов*. Эти характеристики взаимосвязаны: чем больше будет образовано групп, тем меньше величина интервала и наоборот.

Вопрос о числе групп и величине интервала следует решать с учетом следующих факторов:

- цель исследования (группировки);
- характер и размах вариации изучаемого признака;
- число единиц изучаемой совокупности.

При определении числа групп следует учитывать особенности изучаемого явления.



Чем интенсивнее меняется признак и шире размах его вариации, чем больше статистическая совокупность, тем больше может быть образовано групп. Но это правило следует применять разумно, руководствуясь логикой и здравым смыслом. Если групп много, то в некоторых из них может оказаться очень мало единиц. Некоторые из групп могут оказаться вообще пустыми. Это делает характеристики групп ненадежными.

Однако и малое число групп сужает возможности анализа, т.к. увеличивает вероятность того, что в некоторые группы будут включены качественно различные единицы.

Поэтому количество групп должно быть оптимальным, в каждую группу должно входить достаточно большое число единиц совокупности, что отвечает требованию закона больших чисел. Однако в отдельных случаях интерес представляют и малочисленные группы, т.к. новое, передовое, пока оно не станет массовым, проявляется в незначительном числе фактов. Поэтому задача статистики – выявить эти факты, изучить их.

Таким образом, при решении вопроса о числе групп нужно руководствоваться не формальными принципами, а знанием сущности изучаемого явления.

При проведении *структурной* и *аналитической группировок* оптимальное число групп с равными интервалами можно определить по формуле Стерджесса:

$$k = 1 + 3,322 \lg N, \quad (3.3)$$

где  $k$  – число групп;

$N$  – объем совокупности.

На основе формулы Стерджесса может быть составлена следующая номограмма:

$N$	15 – 24	24 – 44	45 – 89	90 – 179	180 – 359
$k$	5	6	7	8	9

В соответствии с формулой (3.3) число групп зависит от численности совокупности.

Недостаток формулы (3.3) состоит в том, что ее применение дает хорошие результаты, если совокупность состоит из большого числа единиц и распределение единиц по признаку, положенному в основание группировки, близко к нормальному.

Другой способ определения числа групп основан на применении показателя среднего квадратического отклонения ( $\sigma$ ). При этом весь диапазон изменения показателя предполагается равным  $6\sigma$ . Если величина интерва-

ла равна  $0,5\sigma$ , то совокупность разбивается на 12 групп, а когда величина интервала равна  $2/3\sigma$  или  $\sigma$ , то совокупность делится соответственно на 9 и 6 групп.

Однако при определении числа групп данными методами существует большая вероятность получения «пустых» или малочисленных групп. Поэтому данным методом нельзя пользоваться механически.

При проведении *типологической группировки* используется другой подход к определению числа групп. Для *типологической группировки* определение числа групп трудно формализовать, т.е. выполнить на основе формулы, т.к. число групп и их интервалы определяются числом типов явления и границами признака для каждого типа.

Вторым существенным вопросом при группировке по количественным признакам является определение интервалов группировки.

*Интервал группировки* – это интервал значений варьирующего признака, лежащих в пределах определенной группы. Каждый интервал имеет свою ширину, верхнюю и нижнюю границы или хотя бы одну из них. *Нижней границей* интервала называется наименьшее значение признака в интервале, а *верхней границей* – наибольшее значение признака в нем. *Ширина интервала* (ее еще часто называют величиной интервала или интервальной разницей) представляет собой разность между верхней и нижней границами интервала:

$$h_i = x_{\max_i} - x_{\min_i}, \quad (3.4)$$

где  $h_i$  – размер интервала для  $i$ -той группы (размер  $i$ -того интервала);

$x_{\max_i}$ ,  $x_{\min_i}$  – максимальное и минимальное значение признака в  $i$ -той группе (интервале).

Интервалы могут быть равными и неравными. Равные интервалы применяются в тех случаях, когда изменение количественного признака внутри совокупности происходит равномерно и в относительно нешироких пределах.

Расчет величины интервала при равных интервалах производится по формуле

$$h = \frac{x_{\max} - x_{\min}}{\text{число групп}}, \quad (3.5)$$

где  $x_{\max}$ ,  $x_{\min}$  – максимальное и минимальное значение признака в совокупности.

Если результат деления, полученный по формуле (3.5), не является целым числом, рекомендуется использовать следующие правила округления.

Если шаг интервала, рассчитанный по формуле (3.5), представляет собой величину, имеющую один знак до запятой (например, 0,66; 1,372; 5,8), то в этом случае полученные значения целесообразно округлить до десятых долей. В приведенном выше примере шагом интервала будут соответственно значения 0,7; 1,4; 5,8.

Когда рассчитанный шаг интервала имеет две значащие цифры до запятой и несколько знаков после запятой, то это значение надо округлить до целого числа. Пусть величина интервала равна 12,785. Тогда это значение следует округлить до целого числа, т.е. 13.

В случае, когда рассчитанный шаг интервала представляет собой трехзначное, четырехзначное и т.д. число, эту величину необходимо округлить до ближайшего числа, кратного 100 или 50. Например, 248 следует округлить до 250.

Рассмотрим пример.

Пусть требуется произвести группировку с равными интервалами предприятий по стоимости основных средств. При этом максимальное значение признака равно 2 040 млн руб., а минимальное его значение – 290 млн руб. Совокупность включает 80 единиц. Согласно формуле (3.3) совокупность должна быть разбита на 7 групп. Сначала найдем размах вариации ( $R$ ):

$$R = 2\,040 - 290 = 1\,750 \text{ (млн руб.)}.$$

Затем определим величину интервала:

$$h = 1\,750 : 7 = 250 \text{ (млн руб.)}.$$

После этого построим интервалы групп (табл. 3.11).

Чтобы не писать каждый раз от ... до, границы групп обозначают следующим образом: 290 – 540; 540 – 790 и т.д.

Особенностью 1-го варианта построения групп является то, что у всех групп интервалы являются *закрытыми*. Во 2-м варианте первая и последняя группы – это группы с *открытыми интервалами*.

*Открытые* – это те интервалы, у которых указана только одна граница: верхняя или нижняя.

*Интервалы, в которых указаны обе границы, называются закрытыми.*

Открытые интервалы используют в случаях, когда признак в выделяемой группе единиц изменяется неравномерно и в широких пределах и когда *отсутствуют качественные различия отдельных единиц данной группы*.

При группировке по количественному признаку границы интервалов могут быть обозначены по-разному. Если основанием группировки

служит признак с непрерывной вариацией, то одно и то же значение признака выступает и верхней, и нижней границами у двух смежных интервалов. Таким образом, верхняя граница  $i$ -того интервала равна нижней границе  $(i + 1)$ -того интервала. Примером такой группировки служат интервалы, приведенные в табл. 3.11.

Таблица 3.11

**Варианты построения интервалов групп**

Группа предприятий	Интервал, млн руб.	
	1-ый вариант	2-ой вариант
I	от 290 до 540	до 540
II	от 540 до 790	540 – 790
III	от 790 до 1 040	790 – 1040
IV	от 1 040 до 1 290	1 040 – 1 290
V	от 1 290 до 1 540	1 290 – 1 540
VI	от 1 540 до 1 790	1 540 – 1 790
VII	от 1 790 до 2 040	1 790 и более

При таком обозначении границ может возникнуть вопрос, в какую группу включать единицы объекта, значения признака у которых совпадают с границами интервалов. Например, во вторую или третью группу должно войти предприятие со стоимостью фондов 790 млн руб. Если нижняя граница формируется по принципу «включительно», а верхняя – по принципу «исключительно», то предприятие должно быть отнесено к третьей группе, в противном случае – ко второй. Для того чтобы правильно отнести к той или иной группе единицу объекта, у которой значение признака совпадает с границами интервалов, можно ориентироваться на открытые интервалы.

Чтобы показать, по какому принципу формируется верхняя граница, иногда используют следующие обозначения:  $540^{(+)}$  или  $540^{(-)}$ . Первая запись означает, что верхняя граница формируется по принципу «включительно», а вторая запись означает, что использован вариант «исключительно».

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

Например, пусть совокупность состоит из 80 предприятий и ее надо разделить на группы по численности занятых. Минимальное и максимальное значения группировочного признака соответственно равны 290 и 2 040 человек. В этом случае возможны следующие варианты построения групп (табл. 3.12).

Таблица 3.12

**Варианты построения интервалов групп**

Группа	Интервал, чел.	
	1-ый вариант	2-ой вариант
I	290 – 540	до 540
II	541 – 790	541 – 790
III	791 – 1 040	791 – 1 040
IV	1 041 – 1 290	1 041 – 1 290
V	1 291 – 1 540	1 291 – 1 540
VI	1 541 – 1 790	1 541 – 1 790
VII	1 791 – 2 040	1 791 и более

*Неравные интервалы* применяются в статистике, когда значения признака варьируются неравномерно и в значительных размерах, что характерно для большинства социально-экономических явлений, особенно при анализе макроэкономических показателей.

Закрытые интервалы могут быть прогрессивно-возрастающими, прогрессивно-убывающими, произвольными. Часто при формировании неравных интервалов используют *принцип равнозаполненности* групп.

Например, если необходимо построить группировку предприятий по показателю выручки от реализации продукции, который варьирует от 500 млн руб. до 4 000 млн руб., то строить группировку с равными интервалами нецелесообразно, потому что, как правило, совокупность предприятий включает большое число малых предприятий, имеющих небольшую выручку. С ростом выручки от реализации продукции значительно уменьшается число предприятий. Таким образом, распределение числа предприятий по величине выручки является неравномерным. Поэтому следует построить группировку с неравными интервалами (табл. 3.13).

Таблица 3.13

**Построение групп с неравными интервалами**

Группа	Интервал, млн руб.
I	500 – 800
II	800 – 1 300
III	1 300 – 2 000
IV	2 000 – 2 900
V	2 900 – 4 000

Величина каждого последующего интервала у этой группировки больше предыдущего на 200 млн руб., т.е. увеличивается в арифметической прогрессии.

### 3.4. Виды и правила построения статистических таблиц

#### Понятие о статистической таблице и ее элементы

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

Отличительной чертой статистической таблицы является то, что она содержит результаты подсчета эмпирических данных и представляет итоги сводки первичных данных.

*Статистической таблицей называется таблица, которая содержит сводную числовую характеристику исследуемой совокупности по одному или нескольким существенным признакам, взаимосвязанным логикой экономического анализа.* Прежде чем переходить к рассмотрению видов и правил построения статистических таблиц, необходимо иметь представление об их основных элементах. Основные элементы статистической таблицы, составляющие ее основу, показаны на рис. 3.4.

Название таблицы

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

Рис. 3.4. Основные элементы статистической таблицы

Источник: [1, с. 129]

При *табличной форме* представления числовой информации значение статистического показателя располагается на пересечении вертикального столбца, называемого *графой* и горизонтальной полосы, являющейся в таблице *строкой*. И графы, и строки в таблице имеют соответствующие четко сформулированные *заголовки*.

Таким образом, внешне таблица представляет собой пересечение граф и строк, которые формируют ее остов. Каждое пересечение образует

клетку таблицы. Размер таблицы определяется произведением числа строк на число граф.

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

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

В случае необходимости таблицы могут сопровождаться примечанием, используемым с целью пояснения заголовка, методики расчета некоторых показателей, источников информации и т.д.

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

*Подлежащим статистической таблицы* называется *объект изучения, характеризующийся цифрами*. Это могут быть: статистическая совокупность в целом; отдельные единицы статистической совокупности в порядке их перечисления; группы единиц сгруппированные по одному или нескольким признакам; территориальные единицы; временные периоды. Обычно подлежащее таблицы располагается в левой ее части.

*Сказуемое статистической таблицы* образует *система показателей, которые характеризуют объект изучения, т.е. подлежащее таблицы*. Сказуемое формирует верхние заголовки и составляет содержание граф с логически последовательным расположением показателей слева направо.

Расположение подлежащего и сказуемого может меняться местами, что зависит от конкретной исследовательской задачи и определяется стремлением аналитика обеспечить наиболее лучшее прочтение и анализ исходной информации об исследуемой совокупности.

### **Виды таблиц**

Статистические таблицы классифицируются в зависимости от характера представления подлежащего и сказуемого.

В зависимости от структуры подлежащего и группировки в нем единиц объекта различают статистические таблицы *простые* и *сложные*. Последние, в свою очередь, подразделяются на *групповые* и *комбинационные*.

Виды статистических таблиц по характеру разработки подлежащего представлены на рис. 3.5.

В простой таблице в подлежащем дается простой перечень каких-либо объектов или территориальных единиц, т.е. в подлежащем нет группировки единиц совокупности. Такие таблицы называют *монографическими* и *перечневыми*. *Монографические таблицы* характеризуют не всю совокупность единиц изучаемого объекта, а только одну какую-либо группу из него, выделенную по определенному признаку. Примером монографической таблицы является таблица 3.14.

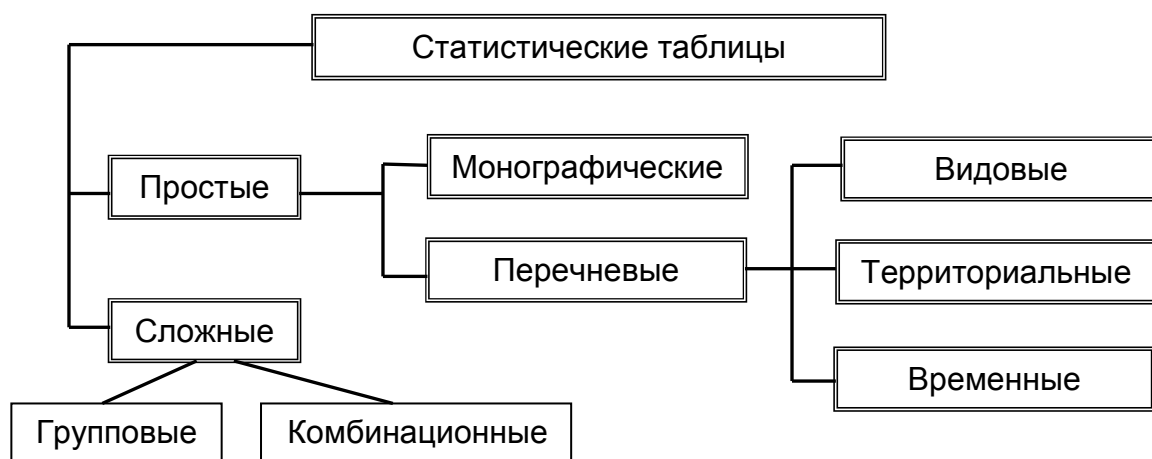


Рис. 3.5. Классификация статистических таблиц по характеру разработки подлежащего  
Источник: [1, с. 131]

Таблица 3.14

**Ввод в действие зданий нежилого назначения в регионе в 2010 г. (данные условные)**

Объект изучения	Показатели ввода		
	Количество зданий, ед.	Общий строительный объем, тыс. м <sup>3</sup>	Общая площадь зданий, тыс. м <sup>3</sup>
А	1	2	3
Здания нежилого назначения	700	35,07	7,2

*Простыми перечневыми называются таблицы, подлежащее которых содержит перечень единиц изучаемого объекта по различному признаку – видовому, территориальному, временному и т.д.*

Примером видовой таблицы может служить табл. 3.15 [4, с. 309], в подлежащем которой содержится перечень основных факторов, препятствующих управленческим новациям.



Таблица 3.15

**Структура факторов, препятствующих внедрению управленческих новаций  
на предприятиях, % к числу всех ответов**

Основные факторы	Распределение ответов
Отсутствие финансовых возможностей	76
Высокая загруженность руководителя	24
Отсутствие информации об эффективных нововведениях в сфере управления	9
Отсутствие возможности получить квалифицированную помощь и консультации	15
Недостаточная квалификация управленческих кадров	13
Отсутствие необходимости в новациях	12
Незначительное влияние новации на результаты деятельности предприятия	10
Отсутствие идей	3

Источник: [4, с. 309]

Примером перечневой таблицы по территориальному признаку является табл. 3.16, которая дает представление об общем коэффициенте рождаемости по регионам Республики Беларусь.

Таблица 3.16

**Общие коэффициенты рождаемости по регионам Республики Беларусь  
для городского и сельского населения (на 1 000 чел.) в 2008 г.**

Регион	Общий коэффициент рождаемости	
	Город	Село
Брестская область	12,4	11,1
Витебская область	10,0	8,8
Гомельская область	11,4	11,4
Гродненская область	12,7	8,2
Минская область	11,1	10,8
Могилевская область	11,1	10,2
г. Минск	11,4	-
Республика Беларусь	11,5	10,3

Источник: [6, с. 54]

Перечневую таблицу по временному признаку представляют данные о структуре работников занятых в сельском хозяйстве Республике Беларусь по уровню образования (табл. 3.17).

Простые таблицы содержат лишь *описательные сведения*. Они не дают возможность выявить социально-экономические типы изучаемых явлений, их структуру, а также взаимосвязи и взаимозависимости между характеризующими их признаками. Эти задачи решаются с помощью групповых, и особенно комбинационных, таблиц.

Таблица 3.17

**Распределение численности работников занятых в сельском хозяйстве  
по уровню образования (в % от общей численности)**

Год	Удельный вес работников, имеющих образование			
	высшее	среднее специальное	общее среднее, включая профессионально- техническое	общее базовое
2000	5,9	13,1	57,3	23,8
2001	6,1	13,3	58,2	22,4
2002	6,5	14,1	59,1	20,3
2003	6,6	13,8	62,5	17,1
2004	6,8	14,3	63,5	15,4
2005	7,2	14,8	63,9	14,1
2006	7,4	15,1	64,3	13,2
2007	7,7	15,4	64,4	12,5
2008	7,9	15,7	64,7	11,7

Источник: [7, с. 26]

*Групповыми* называются статистические таблицы, подлежащее которых содержит группы единиц совокупности по одному количественному или атрибутивному признаку.

Простейшим видом групповых таблиц являются *атрибутивные* и *вариационные ряды распределения*. Групповая таблица может быть более сложной, если в сказуемом приводится не только число единиц в каждой группе, но и ряд других важных показателей, количественно и качественно характеризующих группы подлежащего. Такие таблицы часто используются в целях сопоставления обобщающих показателей по группам, что позволяет делать определенные практические выводы.

Примером групповых таблиц являются табл. 3.1, 3.3, 3.4.

Для более полной характеристики сложных социально-экономических явлений бывает недостаточно провести группировку по одному признаку. Изучаемые объекты обычно характеризуются многими свойствами, многими признаками, часто взаимосвязанными. Для того чтобы раскрыть эти связи и полнее охарактеризовать типы явлений, прибегают к сложной группировке по двум и более признакам. При статистической сводке материалов на основе сложной группировки применяется комбинационная таблица.

*Комбинационными* называются статистические таблицы, подлежащее которых содержит группировку единиц совокупности одновременно по двум и более признакам: каждая из групп, построенная по одному признаку, разбита в свою очередь на подгруппы по какому-либо другому признаку и т.д.

Примером комбинационной является табл. 3.2, представляющая группировку ссуд банка по степени риска. Подлежащим в этой таблице и являются группы ссуд, выделенные по степени риска, который определяется количеством дней просроченной задолженности и обеспеченностью ссуды.

В зависимости от характера разработки сказуемого различают статистические таблицы с простой и сложной разработкой сказуемого. Классификация статистических таблиц по разработке сказуемого приведена на рис. 3.6.

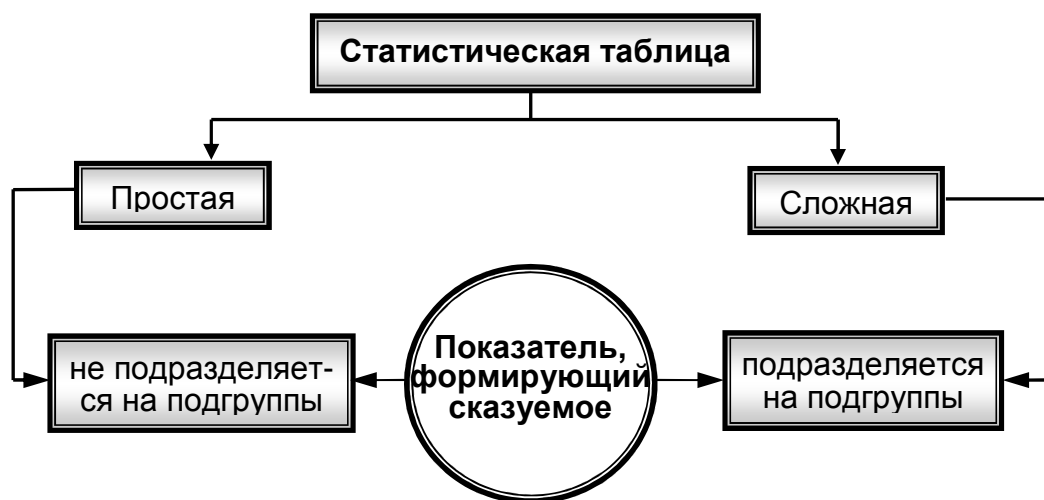


Рис. 3.6. Классификация статистических таблиц по разработке сказуемого  
Источник: [1, с. 136]

При *простой разработке* сказуемого показатель, представляющий его, не подразделяется на подгруппы, и итоговые значения получаются путем простого суммирования значений по каждому признаку отдельно независимо друг от друга. Примером простой разработки сказуемого могут служить табл. 3.1, 3.2, 3.3, 3.4, 3.15, 3.16.

Сложная разработка сказуемого предполагает в сказуемом комбинацию одного признака с другим. Ниже приведены примеры макетов таблиц с простой (табл. 3.18) и сложной (табл. 3.19) разработкой сказуемого.

Таблица 3.18

**Распределение рабочих по стажу и образованию**

Предприятия	Число рабочих, всего	В т.ч. имеющие стаж работы		В т.ч. имеющие образование	
		до 5 лет	свыше 5 лет	среднее	специальное

**Распределение рабочих по стажу и образованию**

Предприятия	Число рабочих, имеющих образование					
	среднее			специальное		
	всего	в т.ч. со стажем		всего	в т.ч. со стажем	
		до 5 лет	свыше 5 лет		до 5 лет	свыше 5 лет

**Правила построения таблиц**

Составление таблицы начинается с построения макета, т.е. незаполненной таблицы, в которой определяется подлежащее, сказуемое, состав и последовательность расположения показателей.

При составлении таблиц следует руководствоваться следующими правилами.

1. Таблица должна быть компактной и содержать только те исходные данные, которые непосредственно характеризуют исследуемые объекты, явления, процессы. Следует избегать второстепенной информации, а также информации, не имеющей отношения к данному объекту исследования. Цифровой материал следует располагать таким образом, чтобы при анализе таблицы сущность явления раскрывалась чтением строк слева направо и сверху вниз.

2. Таблица должна иметь четкое, лаконичное название, где обязательно указывается объект наблюдения, его территориальные и временные границы. Если название таблицы состоит из двух и более предложений, ставится точка, чтобы отделить предложения друг от друга; после последнего предложения точка не ставится.

3. Заголовки граф и заголовки строк также должны быть четкими и краткими. При их написании точки допускаются только при необходимых сокращениях.

4. Если названия отдельных граф повторяются между собой, содержат повторяющиеся термины или несут единую смысловую нагрузку, то им необходимо присвоить общий объединяющий заголовок.

5. Графы и строки полезно нумеровать. Графы, слева заполненные названием строк, принято обозначать заглавными буквами алфавита (А), (Б) и т.д., а все последующие графы – номерами в порядке возрастания.

6. В зависимости от содержания таблицы она может иметь итоговые строки и графы. Для части совокупности (промежуточный результат) пишется ИТОГО, для всей совокупности – ВСЕГО.

7. В таблице должны быть правильно проставлены единицы измерения приводимых показателей. Если они одинаковые для всех показателей,

то их указывают в заголовке таблицы. Если разные, то в каждой графе или строке. При этом используются общепринятые сокращения единиц измерения (чел., руб., кВт-ч и т.д.).

8. Взаимосвязанные и взаимозависимые данные, характеризующие одну из сторон анализируемого явления (например, число предприятий и их удельный вес по группам; темп роста и темп прироста и т.д.), целесообразно располагать в соседних друг с другом графах.

9. Для удобства работы числа в таблицах следует представлять в середине граф, одно под другим: единицы под единицами, запятая под запятой, четко соблюдая при этом их разрядность.

10. По возможности числа целесообразно округлять. Округление чисел в пределах одной и той же графы или строки следует проводить с одинаковой степенью точности (до целого знака или до десятичных и т.д.).

11. Если ряд чисел одной и той же графы или строки имеют один десятичный знак, а несколько чисел имеют два и более знака после запятой, то числа с одним знаком после запятой следует дополнять нулем, обозначая, таким образом, одинаковую, принятую в таблице точность расчетов.

12. Все клетки таблицы должны быть заполнены. При этом в каждой клетке должно быть только одно число. Незаполненность клетки может быть обусловлена разными причинами, что по-разному отмечается в таблице:

- если данная позиция (на пересечении соответствующих графы и строки) вообще не подлежит заполнению, то ставится знак «х»;
- когда по какой-либо причине отсутствуют сведения, то ставится многоточие «...» или «Нет свед.», или «Н. св.»;
- при отсутствии явления клетка заполняется тире «—»;
- для отображения очень малых чисел используют обозначения (0,0) или (0,00), предполагающие, что численное значение показателя имеется, но оно меньше принятой точности расчета.

13. В случае необходимости приведения дополнительной информации таблица может иметь примечания. Обязательно указывается источник данных.

14. Таблица должна быть небольшой. Лучше несколько небольших таблиц, чем одна большая.

15. К таблице должны быть даны объяснения, касающиеся методики расчета показателей, особенно в случаях, если используются оригинальные методики.

16. После таблицы должны быть сделаны выводы.

### Анализ статистических таблиц

Анализу статистических таблиц предшествует этап их чтения. Чтение предполагает, что исследователь, прочитав название, заголовки и числа таблицы, усвоил ее содержание, сформулировал первые суждения об объекте, уяснил назначение таблицы, понял ее содержание в целом, дал оценку явлению или процессу, описанному в таблице.

Анализ таблицы как метод научного исследования включает этапы, представленные на рис. 3.7.

Соблюдение правил и последовательности работы со статистическими таблицами помогает исследователю осуществить научно обоснованный экономико-статистический анализ объектов и процессов.

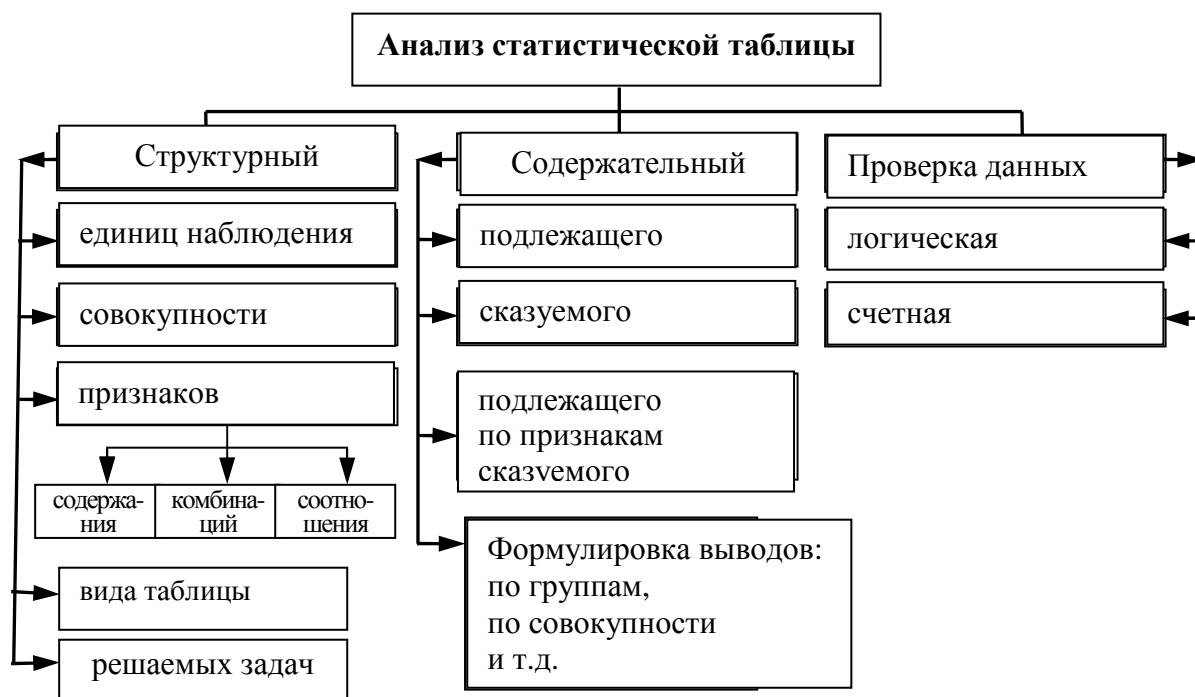


Рис. 3.7. Этапы анализа статистических таблиц

Источник: [1, с. 141]

Помимо табличной формы представления данных в статистике широко применяются и графики. Отдельные виды наиболее широко используемых графиков рассматриваются в соответствующих темах.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Какова цель статистической сводки как второго этапа статистического исследования?
2. Дайте определение статистической сводки.
3. По каким признакам может быть проведена классификация видов статистической сводки?
4. В чем отличие сложной статистической сводки от простой сводки?
5. Как проводится простая статистическая сводка, и какова сфера ее применения?
6. Что включает программа проведения сложной статистической сводки?
7. Что такое централизованная и децентрализованная статистическая сводка?
8. Какие организационные вопросы должны быть решены при проведении статистической сводки?
9. В чем заключается сущность статистической группировки и каково ее значение в статистике?
10. Какие исследовательские задачи решает статистическая группировка как самостоятельный метод статистического исследования?
11. Как связан метод группировки с другими методами статистики?
12. Что такое группировочный признак?
13. Какие виды группировок выделяют в зависимости от решаемых ими задач обобщения (систематизации) данных?
14. В чем состоит сущность типологической группировки?
15. В чем заключается цель проверки правильности проведения типологической группировки?
16. Дайте определение классификации и укажите ее отличительные черты.
17. Каково значение классификаторов в статистике?
18. Какую статистическую группировку называют структурной?
19. Какие показатели используют для оценки структурных сдвигов?
20. В чем заключается сущность аналитической группировки?
21. Какой признак выбирается в качестве группировочного при проведении аналитической группировки?

22. Когда на основе аналитической группировки делают вывод о том, что связь между факторным и результативным признаком проявляется устойчиво?
23. Какая группировка является простой?
24. Каковы разновидности сложной группировки?
25. Какие группировки называют комбинационными?
26. В чем состоит отличие многомерной группировки от комбинационной?
27. Укажите, в каких случаях проводят вторичную группировку.
28. Какие существуют способы проведения вторичной группировки?
29. Какая связь существует между числом групп и величиной интервалов при проведении статистической группировки?
30. В каких случаях целесообразно применение равных интервалов?
31. В каких случаях применяют неравные интервалы?
32. Как выполняется группировка, если группировочный признак является дискретным?
33. Что такое специализированный интервал и специализированный признак?
34. В каких случаях применяют открытые интервалы?
35. Что такое статистическая таблица?
36. Укажите основные элементы статистической таблицы.
37. Что такое подлежащее статистической таблицы?
38. Что является сказуемым статистической таблицы?
39. Какие виды статистических таблиц выделяют в зависимости от характера разработки подлежащего?
40. Какие виды статистических таблиц выделяют в зависимости от характера разработки сказуемого?
41. Назовите основные правила построения статистических таблиц.

### Тест 3

1. Статистическая сводка – это ...
  - а) завершающий этап статистического исследования;
  - б) первый этап статистического исследования;
  - в) этап, следующий за статистическим наблюдением;
  - г) этап, предшествующий статистическому наблюдению.
2. Целью статистической сводки является ...
  - а) систематизация первичных статистических данных;
  - б) получение объективной и полной информации об объекте наблюдения;



- в) контроль данных статистического наблюдения;
- г) публикация статистических данных.

3. По глубине обработки статистических данных сводка может быть ...

- а) простой и сложной;
- б) децентрализованной и централизованной;
- в) ручной и компьютерной.

4. По форме обработки данных сводка может быть ...

- а) простой и сложной;
- б) централизованной и децентрализованной;
- в) компьютерной и ручной.

5. При проведении сложной статистической сводки применяют ...

- а) метод статистической группировки;
- б) метод массового статистического наблюдения;
- в) индексный анализ.

6. В качестве группировочного признака может выступать ...

- а) только качественный признак;
- б) только количественный признак;
- в) как атрибутивный, так и количественный признак.

7. Основанием аналитической группировки является ...

- а) результативный признак;
- б) факторный признак.

8. Укажите, что характерно для типологической группировки:

- а) разделение разнородной совокупности на однородные группы;
- б) изучение наличия и направленности связи между анализируемыми признаками;
- в) изучение структуры совокупности.

9. Укажите отличительные черты классификации:

- а) в основе лежит количественный признак;
- б) в основе лежит качественный признак;
- в) она едина для всех пользователей;
- г) она относительно устойчива.

10. Если величина интервала равна  $0,5\sigma$ , то совокупность разбивается на ...

- а) 6 групп;
- б) 9 групп;
- в) 12 групп.

11. Если две группировки несопоставимы по величине интервалов и числу выделенных групп, их приводят к сопоставимому виду ...

- а) с помощью метода вторичной группировки;
- б) путем построения сложной группировки.

12. Статистическая таблица представляет собой ...

- а) форму рационального изложения и обобщения статистических данных при помощи цифр, расположенных в определенном порядке;
- б) числовые характеристики, расположенные в колонках таблицы;
- в) систему показателей, которые характеризуют объект изучения.

13. К статистическим таблицам можно отнести ...

- а) таблицу умножения;
- б) таблицу, характеризующую структуру населения по полу и возрасту;
- в) опросный лист социологического обследования.

14. По характеру разработки подлежащего различают статистические таблицы:

- а) простые;
- б) групповые;
- в) перечневые;
- г) комбинационные.

15. По характеру разработки сказуемого различают статистические таблицы:

- а) монографические;
- б) перечневые;
- в) простые;
- г) сложные.

16. Сказуемым статистической таблицы является ...

- а) исследуемый объект;
- б) показатели, характеризующие исследуемый объект;
- в) сведения, расположенные в заголовках таблицы.

17. Монографическая статистическая таблица характеризует ...

- а) совокупность всех единиц изучаемого объекта;
- б) группу единиц совокупности, выделенных по определенному признаку;
- в) каждую единицу совокупности.

### **Задания и задачи по теме**

#### **Задание 3.1**

Укажите, какие из указанных ниже группировок, являются типологическими:

- 1) группировка населения по полу;

- 2) группировка населения, занятого в экономике по регионам;
- 3) группировка предприятий промышленности по объему производства продукции (работ, услуг);
- 4) группировка студентов по полученным ими на экзамене баллам;
- 5) группировка предприятий торговли по формам собственности.

### Задание 3.2

Пользуясь формулой Стерджесса, определите, сколько групп должно быть создано при проведении группировки, а также величину интервала, если принято решение о формировании равных интервалов.

Исходные данные: исследуется совокупность из 20 сотрудников одного из филиалов организации. Наименьшая величина заработной платы в этой совокупности составляет 1 050 тыс. руб., максимальная – 3 250 тыс. руб.

### Задание 3.3

Определите, к какому виду группировки относится статистическая табл. 3.20, представляющая группировку промышленных предприятий по размеру основных производственных средств. Укажите:

- 1) группировочный признак, дайте его характеристику на основе классификации;
- 2) подлежащее и сказуемое таблицы;
- 3) вид статистической таблицы, представляющей результаты группировки.

Таблица 3.20

#### Группировка предприятий по размеру основных производственных средств

Группы предприятий по размеру основных производственных средств	Число предприятий	Объем продукции, млн руб.		Среднесписочная численность работников, чел.	
		Всего	В среднем на одном предприятии	Всего	В среднем на одном предприятии
Малые	20	1 500	75	2 000	100
Средние	20	2 000	100	3 000	150
Крупные	10	4 500	450	5 000	500
Итого	50	8 000	160	10 000	200

### Задача 3.4

Имеются следующие данные о стаже работы и месячной выработке рабочих производственного участка:

Таблица 3.21

Данные о стаже и месячной выработке рабочих

Табельный номер рабочего	Стаж работы рабочего, лет	Месячная выработка рабочего, шт. изделий
1	2	326
2	4	368
3	5	360
4	3	340
5	2	310
6	1	320
7	6	380
8	3	350
9	8	370
10	9	380
11	4	360
12	6	368
13	7	370
14	2	320
15	3	335
16	5	364
17	6	374
18	1	324
19	8	390
20	9	386

Для выявления зависимости уровня производительности труда рабочих от стажа их работы проведите аналитическую группировку. С этой целью сформируйте четыре группы рабочих по стажу работы с равными интервалами. По каждой группе и совокупности в целом укажите численность рабочих, средний стаж работы, а также среднюю месячную выработку. Результаты работы представьте в табличной форме и сформулируйте выводы.

### Задача 3.5

Имеются следующие данные о распределении промышленных предприятий двух регионов по среднесписочной численности промышленно-производственного персонала (ППП) (табл. 3.22).

Проведите вторичную группировку данных о распределении промышленных предприятий, пересчитав данные региона II в соответствии с группировкой региона I.

Таблица 3.22

**Группировка предприятий регионов по среднесписочной численности ППП**

Регион I			Регион II		
Группы предприятий по среднесписочной численности ППП, чел.	Удельный вес числа предприятий, %	Удельный вес среднесписочной численности рабочих, %	Группы предприятий по среднесписочной численности ППП, чел.	Удельный вес числа предприятий, %	Удельный вес среднесписочной численности рабочих, %
До 100	32,0	1,0	До 300	34,0	1,0
100 – 500	38,0	4,0	300 – 600	28,0	6,0
500 – 1 000	17,0	10,0	600 – 1 000	20,0	10,0
1 000 – 2 000	9,0	15,0	1 000 – 2 000	13,0	15,0
2 000 – 5 000	3,0	32,0	2 000 – 4 000	4,0	43,0
5 000 и более	1,0	38,0	4 000 и более	1,0	25,0
Итого	100,0	100,0	Итого	100,0	100,0

**Задача 3.6**

Используя данные табл. 3.22, проведите вторичную группировку промышленных предприятий по среднесписочной численности ППП, пересчитав данные региона I в соответствии с группировкой региона II.

**Задача 3.7**

Используя данные табл. 3.22, проведите вторичную группировку промышленных предприятий по среднесписочной численности ППП, пересчитав данные регионов I и II таким образом, чтобы образовать следующие группы промышленных предприятий по среднесписочной численности ППП: до 500; 500 – 1 000; 1 000 – 2 000; 2 000 – 3 000; 3 000 – 4 000; 4 000 – 5 000; 5 000 и более.

**Задание 3.8**

Назовите подлежащее и сказуемое табл. 3.23, укажите ее вид по характеру разработки подлежащего и сказуемого.

**Задание 3.9**

По данным периодических изданий за текущий год, материалам сайта Национального статистического комитета Республики Беларусь подбе-

рите примеры следующих видов статистических таблиц: а) монографической; б) перечневой; в) групповой; комбинационной. По каждой таблице укажите подлежащее и сказуемое.

Таблица 3.23

**Характеристика основных сегментов рынка бизнес-услуг в Республике Беларусь**

Сегмент рынка бизнес-услуг	Емкость сегментов, тыс. долл. США	Прогнозируемый удельный вес в общем объеме рынка на 2007 г., %	Ожидаемая емкость сегментов в 2007г., тыс. долл. США
Аудитинг	750	4,53	923,790
Юридический консалтинг	1 150	9,45	1 927,04
Маркетинг и PR	2 100	16,86	3 439,51
IT-консалтинг	1 050	7,27	1 482,95
Управленческое консультирование	2 250	20,36	4 153,32
Инвестиционный консалтинг	1 650	13,14	2 680,72
Кадровый консалтинг	1 350	11,06	2 255,66
Обучение	2 100	16,61	3 388,29
Другое	2 400	0,73	148,72
Итого	14 800	100,00	20 400,00

Источник: [8, с. 62]

**Задание 3.10**

По данным периодических изданий за текущий год, материалам сайта Национального статистического комитета Республики Беларусь подберите примеры перечневых статистических таблиц, в которых разработка подлежащего была бы произведена по а) видовому; б) территориальному; в) временному принципам.

**Задание 3.11**

По данным периодических изданий за текущий год, материалам сайта Национального статистического комитета Республики Беларусь под-

берите примеры статистических таблиц а) с простой разработкой сказуемого; б) со сложной разработкой сказуемого.

### **Задание 3.12**

Составьте макет таблицы, характеризующей зависимость успеваемости студентов вашей группы от посещаемости учебных занятий. Сформулируйте название таблицы. Укажите, к какому виду таблиц относится макет. Назовите подлежащее и сказуемое.

### **Задание 3.13**

Составьте макет таблицы, характеризующей распределение работников предприятия по уровню образования за период, равный трем годам. Укажите вид таблицы, ее подлежащее и сказуемое.

### **Задание 3.14**

Имеются следующие данные об объеме производства продукции предприятия по годам (в млрд руб.): 2007 г. – 123,0; 2008 г. – 142,5; 2009 г. – 164,7; 2010 г. – 176,0. Из общего объема продукции было предназначено на экспорт (млрд руб.): 2007 г. – 41,5; 2008 г. – 53,0; 2009 г. – 57,6; 2010 г. – 64,1.

Представьте приведенные данные в табличной форме. Укажите вид таблицы.

## Тема 4

# СИСТЕМА СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ

- 4.1. Сущность и классификация статистических показателей.
- 4.2. Сущность абсолютных величин и их значение в статистике.
- 4.3. Сущность, виды и порядок расчета относительных величин.

Изучив данную тему, студенты должны:

- понимать сущность и значение абсолютных величин;
- знать виды и единицы измерения абсолютных величин;
- знать виды относительных величин и формы их выражения;
- знать аналитический потенциал относительных величин и уметь грамотно применять их в анализе социально-экономических явлений и процессов.

### 4.1. Сущность и классификация статистических показателей

*Статистический показатель* – обобщенная количественная характеристика социально-экономических явлений и процессов в их качественной определенности в конкретных условиях места и времени.

Статистические показатели – это язык статистики. Статистический показатель это не просто число, он обязательно должен иметь следующие атрибуты:

- название показателя (что характеризует?);
- численное значение с указанием единицы измерения (если показатель именованный);
- указание объекта, к которому относится показатель;
- указание о времени, к которому относится показатель.

Например, рассмотрим статистический показатель: «начисленная среднемесячная заработная плата работников Республики Беларусь в I квартале 2011 года составила 1463, 7 тыс. рублей». Укажем его атрибуты:

- *наименование*: начисленная среднемесячная заработная плата работников;
- *численное значение*: 1463, 7 тыс. руб.;
- *объект*, к которому относится показатель, – Республика Беларусь;
- *время*, к которому относится показатель – I квартал 2011 года.

Выделяют следующие виды статистических показателей:



1) по *форме обобщения*:

- *абсолютные величины*;
- *относительные величины*;
- *средние величины*.

Аналитические возможности и способы расчета этих показателей рассматриваются в данной и последующей темах;

2) по *содержанию*:

- *количественные (объемные) показатели*. К этой группе относят показатели, которые характеризуют либо размер совокупности, либо размер того или иного признака в совокупности. Например, списочная численность персонала предприятия на дату, фонд заработной платы по предприятию;
- *качественные (интенсивные) показатели*. Показатели этой группы характеризуют уровень развития явлений. Например, средняя заработная плата, фондоотдача, производительность труда и др. Деление показателей на количественные и качественные потребуются нам, в частности, при изучении темы «Индексный метод в статистических исследованиях»;

3) в зависимости от *объема объекта статистического изучения*:

- *индивидуальные показатели* (характеризуют отдельные единицы совокупности);
- *обобщающие показатели* (характеризуют всю совокупность или отдельные группы).

В отличие от признака статистический показатель получается расчетным путем. Это может быть простой подсчет единиц совокупности, суммирование значений признака, деление одной величины на другую или более сложные расчеты.

Как правило, изучаемые статистикой процессы и явления достаточно сложны, и их сущность не может быть отражена посредством одного отдельно взятого показателя. Поэтому обычно в статистико-экономическом анализе используют *системы статистических показателей* и задача статистики формировать такие системы и уметь их интерпретировать.

## **4.2. Сущность абсолютных величин и их значение в статистике**

### **Сущность и виды абсолютных величин**

В процессе статистического наблюдения получают данные о значениях анализируемых признаков по каждой единице исследуемой совокупно-

сти. Для характеристики совокупности в целом или ее частей данные по отдельным единицам совокупности подвергают сводке и получают обобщающие показатели. Обобщающие показатели, как уже отмечалось, могут быть представлены *абсолютными, относительными и средними величинами*.

*Абсолютные величины – статистические показатели, которые характеризуют либо размер совокупности, либо размер изучаемых признаков в совокупности в конкретных условиях места и времени.*

Примерами абсолютных величин являются: списочная численность работников на определенную дату (например, на конец года); прибыль предприятия; фонд заработной платы работников предприятия и т.п.

Различают три вида абсолютных величин: *индивидуальные, групповые и общие*.

*Индивидуальными называют такие абсолютные величины, которые выражают размеры количественных признаков у отдельных единиц изучаемой совокупности.* Индивидуальные абсолютные величины устанавливаются непосредственно в процессе статистического наблюдения и регистрируются в формулярах наблюдения.

*Групповые и общие абсолютные статистические величины выражают размер того или иного признака у всех единиц данной совокупности, или у отдельных их групп, либо численность единиц всей совокупности или отдельных ее частей (групп).* Из определения следует, что групповые и общие абсолютные величины получают в результате суммирования индивидуальных абсолютных величин, т.е. значений признака у отдельных единиц совокупности, либо путем подсчета числа единиц, входящих в отдельные группы или в совокупность в целом.

### **Единицы измерения абсолютных величин**

Абсолютные статистические показатели всегда являются *именованными*, т.е. имеют какую-либо единицу измерения. В зависимости от сущности исследуемых явлений, их физических свойств, абсолютные величины могут быть выражены в натуральных, стоимостных и трудовых единицах измерения.

*Натуральные единицы* измерения отражают конкретные свойства единиц совокупности или ее объем. Если исследуется одно свойство единиц совокупности или ее объем, то мы будем иметь *простую натуральную единицу измерения*. Например, количество переработанной нефти выражается в тоннах, количество произведенных изделий – в штуках, численность населения – в человеках.

Если одновременного изучается два свойства единиц совокупности, используют *составные натуральные единицы измерения*. Объем перевозок на транспорте, например, задается в тонно-километрах, т.е. измеряется и протяженность пути, и масса перевозимых грузов. Количество энергии, выработанной электростанциями, представляется в киловатт-часах (учитывается мощность электростанций и продолжительность работы).

В группу натуральных также входят *условно-натуральные измерители*, используемые в тех случаях, когда какой-либо продукт имеет несколько разновидностей, различающихся степенью проявления некоторого общего для них потребительского свойства. Например, все виды органического топлива переводятся в условное топливо, с теплотой сгорания 29,3 МДж/кг (7 000 Ккал/кг); мыло разных сортов – в мыло с 40 %-ным содержанием жирных кислот; консервы различного объема переводят в условные консервные банки объемом 353,4 см<sup>3</sup> и др.

Перевод в условные натуральные единицы измерения осуществляется на основе специальных коэффициентов, рассчитываемых как отношение значений потребительского свойства у отдельных разновидностей продукта к эталонному значению.

**Пример 4.1.** Имеются данные о поставках молочной продукции в торговую сеть города (табл. 4.1).

Таблица 4.1

**Объем поставки молочной продукции в торговую сеть города**  
(в натуральных единицах измерения)

Наименование продукции	Объем поставки, т
Молоко, 3,2 %	10,8
Молоко, 6,0 %	8,5
Молоко, 1,5 %	4,6
Кефир, 3,2 %	2,7
Ряженка, 4 %	0,6
Сметана, 30 %	5,8
Творог, 9 %	3,4

Определим общий объем поставок молочной продукции в торговую сеть города на основе пересчета объема поставок в натуральных единицах измерения в условно-натуральные единицы. За эталонное значение примем жирность молока 3,2 %, все остальные виды продукции с помощью специальных коэффициентов приведем к этому значению, т.е. выразим в условно-натуральных единицах измерения. Так, например, 8,5 т молока 6 %-ной жирности, будут эквивалентны 15,94 т условных натуральных единиц измерения, соответствующих жирности 3,2 % ( $8,5 \cdot (0,06 / 0,032)$ ). Анало-

гично расчет выполнен по остальным видам молочной продукции и результаты расчетов представлены в табл. 4.2.

Таблица 4.2

**Объем поставок молочной продукции**  
(в условных натуральных единицах измерения)

Наименование продукции	Объем поставки в пересчете на жирность 3,2 %, т
Молоко, 3,2 %	10,80
Молоко, 6,0 %	15,94
Молоко, 1,5 %	2,16
Кефир, 3,2 %	2,70
Ряженка, 4 %	0,75
Сметана, 30 %	54,38
Творог, 9 %	9,56
Итого	96,29

Таким образом, с учетом жирности молочной продукции в торговую сеть города в отчетном периоде поставлено 96,29 условных т продукции.

*Стоимостные* или *денежные единицы измерения* используются для характеристики в стоимостном выражении многих статистических показателей. Например, объема произведенной продукции, товарооборота, объема реализованной продукции, прибыли предприятия, доходов населения и др. Широкое распространение измерения различных показателей в денежных единицах объясняется тем, что с переходом от натуральных и других единиц измерения к денежным, обеспечивается возможность соизмерения, обобщения (суммирования) и сопоставления разнородных величин.

При анализе и сопоставлении стоимостных показателей, относящихся к разным периодам, необходимо иметь в виду их несопоставимость вследствие инфляции. Для того чтобы произвести подобные сравнения, осуществляют пересчет в сопоставимые цены.

К *трудовым единицам измерения* относят человеко-час, человеко-день, которые используют для измерения как общих затрат труда на производство продукции, так и трудоемкости отдельных операций производственного процесса или отдельных видов продукции.

В качестве своеобразных единиц измерения выступают сами единицы изучаемой совокупности (вещи, предметы, события и др.), когда производится их подсчет для определения объема (численности) совокупности в целом или отдельных ее частей (групп). Каждый отдельный предмет, случай, будучи единицей счета, в то же время является и единицей измерения.

### **Значение абсолютных величин в статистике**

Характеризуя значение абсолютных величин в статистическом анализе, следует отметить, что их аналитический потенциал ограничен, т.к. абсолютные величины представляют только либо размер совокупности, либо размер какого-либо признака совокупности. Другими словами, они не дают полного представления об изучаемом явлении (объекте), не характеризуют его структуру, развитие во времени, взаимосвязи с другими явлениями. С другой стороны, абсолютные величины играют очень важную роль в статистических методах исследования, т.к. на их основе строятся другие виды статистических показателей, например, средние и относительные величины.

### **Способы получения абсолютных величин**

Значительная часть абсолютных величин получается как результат непосредственных замеров, подсчета, взвешиваний и т.п. в процессе статистического наблюдения. Для получения других абсолютных величин необходимо использовать расчетный способ. Например, чтобы определить объем производства продукции в стоимостном выражении, необходимо количество продукции каждого вида умножить на цену и просуммировать полученные произведения по всем видам выпускаемой продукции. Расчетный метод определения абсолютных величин может опираться также на балансовый метод. Например, наличие оборудования на конец года определяется путем прибавления к количеству оборудования на начало года числа единиц введенного оборудования и вычитания числа единиц вышедшего оборудования. Более простым расчетным способом получения абсолютных величин является суммирование значений признаков у единиц совокупности. Например, суммируя размер начисленной заработной платы работников предприятия, получаем фонд заработной платы.

## **4.3. Сущность, виды и порядок расчета относительных величин**

### **Сущность и принципы построения относительных величин**

Относительные величины в статистических исследованиях применяются очень широко. Относительные величины используют в анализе группировок, анализе динамики, в индексном анализе.

*Относительная величина – величина, являющаяся мерой количественного соотношения статистических показателей и отображающая относительные размеры социально-экономических явлений.* Это может

быть соотношение численностей разных совокупностей, их отдельных признаков, размеров разных признаков одной и той же совокупности, соотношение плановой и фактической величин показателей, или величин показателя, относящегося к разным периодам.

Построение относительных величин предполагает сочетание знания свойств исследуемых объектов и общих принципов статистической методологии.

Рассмотрим принцип построения относительных величин [16, с. 50 – 52].

*Первый принцип.* Сравнимые в относительной величине показатели должны быть чем-то связаны в реальной жизни объективно, независимо от нашего желания. Необходимо добиваться как можно большего соответствия по смыслу сравниваемых показателей. Например, необходимо построить относительный показатель, характеризующий степень (уровень) грамотности населения. Можно разделить численность грамотного населения на общую численность населения, но это не лучший из показателей. Следует принять во внимание, что дети дошкольного возраста, некоторые категории инвалидов с детства не могут наравне со здоровыми и достигшими школьного возраста людьми быть обучены грамоте. Из всего населения эти категории лиц правильнее исключить при построении относительного показателя грамотности.

*Второй принцип.* При построении относительной статистической величины сравниваемые показатели могут различаться только одним атрибутом: или видом признака (при одинаковом объекте, периоде времени, плановом или фактическом значении показателей), или временем (при одном и том же признаке, объекте), или только фактическим, плановым или нормативным характером показателей (при том же объекте, признаке, времени). Нельзя сопоставлять показатели, различные по двум и более атрибутам, например, сравнивать объем производства продукции молокозавода в натуральных единицах измерения за январь 2010 г. со стоимостью продукции хлебозавода за 2010 г.

*Третий принцип.* Необходимо знать возможные границы применения относительных величин. Например, если исходные показатели в текущем и базисном периодах имеют разные знаки, то теряет смысл и не может применяться относительная величина динамики – темп роста. Например, если предприятие имело в 2009 г. убыток 150 млн руб., а в 2010 г. получило прибыль 300 млн руб., неверно ни то, что «финансовый результат вырос вдвое» (если отбросить знаки), ни то, что он «вырос в минус 2 раза», если делить + 300 млн на (-150 млн).

Как следует из определения относительных величин, получают их путем деления одной величины на другую. Чаще всего относительные величины являются результатом деления одной абсолютной величины на другую, но возможно и сопоставление средних и относительных величин. Таким образом, принципиально, любая относительная величина (ОВ) определяется следующим образом:

$$ОВ = \frac{\text{Сравниваемая(текущая,отчетная) величина}}{\text{Основание относительной величины}}. \quad (4.1)$$

Величина, с которой производится сравнение (знаменатель дроби), обычно называется *основанием относительной величины, базой сравнения* или *базисной величиной*, а та, которая сравнивается, называется *текущей, сравниваемой* или *отчетной величиной*.

### **Форма выражения относительных величин**

Относительные величины получают в результате сопоставления *одноименных* и *разноименных* показателей.

В результате сопоставления *одноименных показателей* получают относительные величины, как правило, не имеющие размерности, т.е. не именованные. Если база сравнения принимается за 1, то относительный показатель выражается в коэффициентах, если база принимается за 100, 1 000 или 10 000, то относительный показатель соответственно выражается в процентах (%), промилле (‰) и продецимилле (‱) (табл. 4.3).

Таблица 4.3

**Форма выражения относительных величин  
в зависимости от численного значения базы сравнения**

Численное значение базы сравнения	Форма выражения полученных относительных величин	Обозначение полученных величин
1	Коэффициент, доля	Коэффициент, десятичная дробь
100	Процент	%
1 000	Промилле	‰
10 000	Продецимилле	‱

Проценты, как правило, используют в тех случаях, когда сравниваемый показатель превосходит базисный не более чем в 2 – 3 раза. Проценты же свыше 200 – 300 обычно заменяют кратным отношением, т.е. коэффициентом. Так, вместо 530 % говорят, что сравниваемый показатель больше базисного в 5,3 раза.

Дадим пояснения по поводу статистического термина «промилле». Промилле применяется в случаях, когда сравниваемая величина намного меньше базы сравнения. Примером сферы широкого использования промилле является демографическая статистика, где уровень рождаемости, смертности, естественной прибыли (убыли) населения и другие показатели рассчитываются на 1 000 человек населения. Например, коэффициент рождаемости  $K_p$  определяется делением числа родившихся за период (обычно год)  $N_{род.}$  на среднегодовую численность населения ( $\bar{N}$ ) по формуле

$$K_p = \frac{N_{род.}}{\bar{N}} \cdot 1\,000. \quad (4.2)$$

Если при среднегодовой численности населения города 105 тыс. человек в течение года родилось 892 младенца, то коэффициент рождаемости составит

$$K_p = \frac{892}{100500} \cdot 1\,000 = 8,5 \text{ ‰}.$$

В такой форме относительная величина намного лучше воспринимается по сравнению с вариантом, представляющим ее в виде десятичной дроби 0,0085.

Продецимилле применяется редко. Наиболее широко применяют проценты, которые рекомендуется исчислять с точностью до одной десятой. Промилле также принято определять с точностью до одного десятичного знака.

### Пример 4.2

Рассмотрим пример, показывающий применение разных форм выражения относительных величин на одном и том же статистическом материале (табл. 4.4).

Таблица 4.4

**Структура студентов заочного отделения по возрасту**

Группы студентов по возрасту, лет	Численность студентов, чел.	Относительная величина структуры		
		доли ед.	%	‰
до 20	72	0,116	11,6	116
20 – 24	298	0,481	48,1	481
25 – 29	171	0,276	27,6	276
30 – 34	74	0,119	11,9	119
35 и старше	25	0,040	4,0	40
Итого	620	1,000	100,0	1000



Различие между исчисленными в табл. 4.4 относительными величинами заключается только в месте расположения запятой, отделяющей целые числа от дробей.

В тех случаях, когда относительная величина получается в результате сопоставления разноименных показателей, в большинстве случаев она имеет наименование, образуемое от наименований сравниваемых величин, например, тыс. руб./чел; чел./кв. км; ц/га.

### **Виды относительных величин**

В зависимости от содержания сравниваемых показателей выделяют следующие виды относительных величин:

- планового задания;
- выполнения плана;
- динамики;
- структуры;
- координации;
- интенсивности и уровня экономического развития;
- сравнения.

Рассмотрим сущность названных выше относительных величин.

**Относительная величина планового задания (ОВП).** *Относительная величина планового задания представляет собой отношение значения показателя, устанавливаемого на планируемый период, к его значению, достигнутому в периоде, предшествующему плановому:*

$$\text{ОВП} = \frac{\text{Показатель, планируемый на } (i + 1) \text{ период}}{\text{Показатель, достигнутый в } i \text{ – том периоде}}. \quad (4.3)$$

Строго говоря, относительные величины этого вида не являются статистическими показателями, т.к. представляют собой результат выполнения такой функции менеджмента, как планирование. Однако их принято рассматривать в статистике, т.к. они связаны с таким видом относительных величин, как относительная величина выполнения плана.

**Относительная величина выполнения плана (ОВВП).** *Относительной величиной выполнения плана называется величина, выражающая соотношение между фактическим и плановым уровнями показателя:*

$$\text{ОВВП} = \frac{\text{Фактический показатель, достигнутый в } (i + 1) \text{ периоде}}{\text{Плановый показатель, установленный на } (i + 1) \text{ период}}. \quad (4.4)$$

**Относительная величина динамики (ОВД).** Динамикой в статистике называется изменение явления во времени. *Относительная величина*

динамики представляет собой отношение уровня (значения) показателя за данный период (по состоянию на данный момент времени) к его уровню в прошлом:

$$\text{ОВД} = \frac{\text{Значение показателя в данном периоде}}{\text{Значение показателя в периоде, принятом за базу сравнения}}. \quad (4.5)$$

При исчислении относительных величин динамики возникает вопрос о выборе базы сравнения. Здесь возможны два варианта построения относительных величин динамики.

– относительные величины *с переменной базой*, или *цепные показатели*. В этом случае значение показателя в данном периоде сравнивается с его значением в предшествующем периоде;

– относительные величины динамики *с постоянной базой сравнения*, или *базисные показатели*. В этом случае каждый данный уровень показателя сравнивается с уровнем показателя, относящимся к начальному периоду или какому-либо другому периоду, принятому за базу сравнения. Более подробно порядок расчета показателей динамики будет рассмотрен в теме 8 «Статистическое изучение динамики социально-экономических явлений».

Следует обратить внимание, что при расчете относительных величин планового задания, выполнения плана и динамики необходимо обеспечить сопоставимость сравниваемых показателей по следующим параметрам:

– данные должны относиться к одному и тому же кругу объектов. Например, нельзя сопоставлять фонд заработной платы рабочих и фонд заработной платы всего персонала предприятия;

– данные должны быть сопоставимы в отношении конкретной сущности показателя, т.е. сравниваться должны одинаковые по содержанию показатели;

– сравниваемые показатели должны быть исчислены по единой методологии. Например, нельзя сопоставлять рентабельность продаж и рентабельность продукции;

– сравниваемые показатели должны быть выражены в одних и тех же единицах измерения;

– если сопоставляются стоимостные показатели, то они должны быть по возможности выражены в сопоставимых ценах. Если представление сравниваемых показателей в сопоставимых ценах по каким-либо причинам затруднено, необходимо оговаривать, что относительная величина динамики оценивает изменение анализируемого показателя в фактических ценах.

Рассмотрим пример расчета относительных величин планового задания, выполнения плана и динамики.

### Пример 4.3

Объем продаж предприятия в 2009 г. составил 10 млрд руб. Исходя из результатов анализа конъюнктуры товарного рынка и с учетом своих производственных возможностей предприятие запланировало довести объем продаж в 2010 г. до 12,7 млрд руб. Фактический объем продаж в 2010 г. составил 12 млрд руб.

В этом случае относительная величина планового задания составит:

$$\text{ОВП} = \frac{12,7}{10} = 1,27, \text{ или } 127\%, \text{ т.е. предприятие запланировало в 2010 г.}$$

увеличить объем продаж на 27 %, что соответствует 2,7 млрд руб.

Относительная величина выполнения плана составит:

$$\text{ОВВП} = \frac{12}{12,7} = 0,945, \text{ или } 94,5\%, \text{ т.е. фактический объем продаж в}$$

2010 г. составил 94,5 % от запланированного. Следовательно, невыполнение плана по объему продаж составило 5,5 % (94,5 – 100), или 0,7 млрд руб.

Определяем относительную величину динамики:

$$\text{ОВД} = \frac{12}{10} = 1,2, \text{ или } 120\%.$$

Таким образом, фактический объем продаж в 2010 г. по сравнению с 2009 г. увеличился в 1,2 раза.

Между относительными величинами планового задания, выполнения плана и динамики существует следующая зависимость:

$$\text{ОВД} = \text{ОВП} \cdot \text{ОВВП}. \quad (4.6)$$

В нашем примере:  $1,27 \cdot 0,945 = 1,2$ .

**Относительная величина структуры (ОВС).** *Относительная величина структуры представляет собой отношение численности единиц или размера изучаемого признака части совокупности к численности единиц или размеру признака всей совокупности:*

$$\text{ОВС} = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель, характеризующий всю совокупность в целом}}. \quad (4.7)$$

Относительные показатели структуры выражаются в долях единицы или в процентах и соответственно их называют долями или удельными весами. Обычно относительные показатели структуры обозначают  $d_i$ .

С помощью относительных величин структуры не только выявляется структура изучаемой совокупности, но и структурные сдвиги в ней, т.е. изменение соотношения ее частей за определенный период, а также направление и тенденции этих изменений. Для того чтобы оценить структурные сдвиги, нужно вычислить и проанализировать показатели структуры за несколько периодов или моментов времени. Порядок расчета структурных сдвигов рассматривается в теме 3.

Структурные показатели очень широко используются в статистике, например, при анализе структуры ВВП; инвестиций по источникам финансирования; структуры основных средств; структуры себестоимости продукции по элементам затрат и статьям калькуляции. Относительные показатели структуры определяются при изучении рынка труда, национального богатства, населения и уровня жизни людей.

#### **Пример 4.4**

Поступление иностранных инвестиций в экономику Республики Беларусь составило в 2006 г. 4 036 054 тыс. долл. США, в т.ч.:

- прямые инвестиции в реальный сектор – 748 594,1 тыс. долл. США;
- портфельные инвестиции – 3 224,2 тыс. долл. США;
- прочие инвестиции – 3 284 235,7 тыс. долл. США [9, с.16].

Необходимо представить структуру иностранных инвестиций в экономику Беларуси.

Рассчитаем удельный вес каждой формы инвестирования:

а) удельный вес прямых инвестиций в общем объеме иностранных инвестиций ( $d_1$ ) составил:

$$d_1 = \frac{748\,594,1}{4\,036\,054,0} \cdot 100 = 18,55 \%;$$

б) удельный вес портфельных инвестиций в общем объеме иностранных инвестиций ( $d_2$ ) был равен:

$$d_2 = \frac{3\,224,2}{4\,036\,054,0} \cdot 100 = 0,08 \%;$$

в) удельный вес прочих инвестиций в общем объеме иностранных инвестиций ( $d_3$ ) может быть определен двумя методами.

Во-первых, на основе использования формулы относительных величин структуры, как это было сделано в предыдущих расчетах:

$$d_3 = \frac{3\,284\,235,7}{4\,036\,054,0} \cdot 100 = 81,37 \%.$$

Во-вторых, как разность между 100 % и удельным весом первых двух элементов ( $d_1$  и  $d_2$ ):

$$d_3 = 100 - (d_1 + d_2) = 100 - (18,55 + 0,08) = 81,37 \%$$

Таким образом, наибольший удельный вес в общем объеме иностранных инвестиций в РБ в 2006 г. составили прочие инвестиции.

**Относительные величины координации.** *Относительной величиной координации называют показатель, который оценивает соотношение размеров частей некоторого целого между собой:*

$$\text{ОВК} = \frac{\text{Показатель, характеризующий } i \text{ – тую часть совокупности}}{\text{Показатель, характеризующий другую часть этой же совокупности}} \quad (4.8)$$

Относительная величина координации показывает, сколько единиц одной части целого приходится на 1; 100; 1 000 и т.д. единиц другой его части.

Примером относительных величин координации может служить соотношение между объемом экспорта и объемом импорта; соотношение между численностью мужского и женского, городского и сельского населения; соотношение собственного и заемного капитала; соотношение между численностью рабочих и служащих на предприятии; соотношение отечественных и импортных товаров на потребительском рынке и т.д.

При построении относительных величин координации, как правило, в качестве базы сравнения принимают меньшую часть совокупности, таким образом, показатель координации покажет, во сколько раз сравниваемая часть явления больше части, принятой за базу. В ряде случаев от этой рекомендации отступают, выбирая в качестве базы сравнения ту часть, которая является приоритетной с экономической, социальной или какой-либо другой точки зрения.

**Относительная величина интенсивности (ОВИ).** *Относительная величина интенсивности представляет собой соотношение размеров двух качественно различных явлений, находящихся в связи друг с другом. Одно из них – среда (ее размер), в которой происходит развитие какого-либо процесса, явления или которые ею порождаются, другое – изучаемый процесс, явление. Таким образом, относительная величина интенсивности характеризует степень развития (распространения) того или иного процесса, явления в определенной среде:*

$$\text{ОВИ} = \frac{\text{Показатель, характеризующий явление А}}{\text{Показатель, характеризующий среду распространения явления А}} \quad (4.9)$$

Поскольку при построении ОВИ сопоставляются разноименные показатели, этот вид относительных величин, как правило, является именованным, т.е. для облегчения понимания его содержания вводится единица измерения, отражающая смысл показателя. Например, при определении плотности населения численность населения делят на площадь территории страны (региона) и получают показатель интенсивности, имеющий единицу измерения (чел./км<sup>2</sup>).

В отдельных случаях ОВИ может быть выражена в виде коэффициента, в процентах или промилле. Например, при определении уровня рождаемости, измеряемого в ‰, рассчитывается число родившихся на 1 000 человек населения. В процентах (или в виде коэффициента) представляют, например, показатели рентабельности, которые по своей сути относятся к ОВИ.

Относительные величины интенсивности в статистике и экономике несут важную содержательную нагрузку. Они позволяют оценивать социально-экономическое развитие любого хозяйствующего субъекта, региона, страны. Примерами относительных величин интенсивности являются: показатели уровня производительности труда, показатели фондоотдачи, фондовооруженности труда, затраты на рубль произведенной продукции и др.

Разновидностью относительных величин интенсивности являются *относительные величины уровня экономического развития*, характеризующие соотношение величины важнейших экономических показателей страны, региона и численности населения (так называемые показатели на душу населения). Например, производство валового внутреннего продукта на душу населения, среднедушевое производство (или потребление) конкретных видов продукции в расчете на душу населения.

**Относительная величина сравнения (ОВСр).** *Относительная величина сравнения представляет собой соотношение одноименных показателей, характеризующих разные объекты (предприятия, регионы, страны и т.п.).*

$$\text{ОВСр} = \frac{\text{Показатель, характеризующий объект А}}{\text{Показатель, характеризующий объект Б}}. \quad (4.10)$$

Так, например, можно сравнивать среднюю заработную плату или средний размер пенсии по странам.

Итак, относительные величины – один из важнейших способов обобщения и анализа статистической информации. При этом важно отметить, что в процессе экономико-статистического анализа абсолютные и относительные величины должны рассматриваться во взаимосвязи, т.е. поль-

зоваться относительными величинами нужно не формально, а представлять, какая абсолютная величина скрывается за каждым относительным показателем. Особенно важно соблюдать это положение при расчете относительных величин динамики.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Охарактеризуйте сущность абсолютных величин.
2. Каково значение абсолютных величин в статистике?
3. Назовите виды абсолютных величин. Приведите примеры абсолютных величин каждого вида.
4. Укажите единицы измерения абсолютных величин.
5. Укажите способы получения абсолютных величин.
6. Определите сущность относительных величин.
7. Какие принципы необходимо соблюдать при построении относительных величин?
8. Какие существуют формы выражения относительных величин?
9. Какие виды относительных величин выделяют в зависимости от их познавательного значения?
10. Как связаны между собой относительные величины планового задания, выполнения плана и динамики?
11. Для чего рассчитывают относительные величины структуры?
12. В чем состоит различие между относительной величиной структуры и относительной величиной координации?
13. Каково значение относительных величин интенсивности? Приведите примеры относительных величин интенсивности, используемых в анализе социально-экономических процессов.
14. Почему важно анализировать абсолютные и относительные величины во взаимосвязи?

### **Тест 4**

1. К какому виду относится каждая из названных ниже абсолютных величин:
  - а) заработная плата рабочего Сидорова И.И. в мае 2010 г.;
  - б) фонд заработной платы работников предприятия за май 2010 г.;

в) фонд заработной платы работников одного из цехов данного предприятия за май 2010 г.

2. Укажите, какие единицы измерения объема продукции (работ, услуг) более универсальны:

- а) натуральные;
- б) денежные.

Ответ объясните.

3. Укажите причины несопоставимости следующих абсолютных величин:

- а) объем производства продукции предприятия А составил в 2010 г. 15 млрд. руб.;
- б) объем производства продукции того же предприятия составил в январе 2010 г. – 10 000 т.

4. Относительные величины выражают соотношение ...

- а) только между одноименными величинами;
- б) как между одноименными, так и между разноименными величинами;
- в) только между разноименными.

5. Для определения удельного веса рабочих в общей численности работников используют ...

- а) относительные величины структуры;
- б) относительные величины сравнения;
- в) относительные величины координации.

6. Число уволившихся по собственному желанию в 2010 г. составило на предприятии А – 80 человек, на предприятии Б – 160 человек. Можно ли сказать, что текучесть кадров на первом предприятии в 2 раза меньше? Почему?

7. Имеются следующие значения относительных величин:

- а) 132 промилле;
- б) 90 процентов;
- в) 16 продецимилле.

Запишите показатели в виде коэффициентов.

8. Какой вид относительных величин используется для оценки ...

- а) изменения явления во времени;
- б) соотношения частей одного целого.

9. Относительная величина выполнения плана по производству продукции в анализируемом периоде составила 110 %, при этом объем производства продукции в данном периоде по сравнению с предшествующим периодом вырос на 5 %. Что предусматривал план:



- а) снижение объема производства продукции;
- б) рост объема производства продукции.

10. Укажите, чему равен коэффициент изменения производительности труда, если в анализируемом периоде по сравнению с предшествующим периодом произошло снижение этого показателя на 17 %:

- а) 0,17;
- б) 1,17;
- в) 0,83;
- г)  $0,17^2$ .

11. Сумма относительных величин структуры, выраженных в процентах и рассчитанных по одной совокупности, должна быть ...

- а) равна 100 %;
- б) меньше 100 %;
- в) больше 100 %;
- г) больше или равна 100 %.

### Задачи по теме

#### Задача 4.1

По данным таблицы определите:

- общее потребление топлива по плану и фактически;
- процент выполнения плана по общему расходу топлива;
- структуру расхода топлива по плану и фактически.

Вид топлива	Ед. изм.	Расход	
		по плану	фактически
Нефть	т	500	520
Уголь	т	320	280

Теплота сгорания нефти – 45 МДж/кг; угля – 26,8 МДж/кг; условного топлива – 29,3 МДж/кг.

#### Задача 4.2

По данным табл. 4.5 определите:

- 1) удельный вес рабочих, руководителей, специалистов и других служащих в общей численности персонала;
- 2) удельный вес работников с высшим образованием, средним специальным, профессионально-техническим, общим средним и базовым образованием в общей численности работников предприятия.

Результаты расчетов оформите в таблице и проиллюстрируйте графически. Укажите вид использованных относительных величин.

Таблица 4.5

**Состав персонала предприятия по категориям, уровню образования и полу**

Наименование показателя	Списочная численность работников на конец года, чел.	В том числе				
		служащие	из них			рабочие
			руководители	специалисты	другие служащие	
1. Всего работников:	6 700	1 520	510	960	50	5 180
в т.ч. имеют образование						
высшее	1 590	990	375	603	12	600
среднее специальное	1 600	440	110	329	13	1 150
профессионально-техническое	2 160	35	12	8	10	2 190
общее среднее	1 250	54	13	20	14	1 158
базовое	100	1	–	–	1	82
Из строки 1 женщины	2 300	790	110	635	45	1 510

**Задача 4.3**

По данным табл. 4.5 определите:

- 1) удельный вес руководителей с высшим образованием в общей численности руководителей;
- 2) долю женщин в общей численности рабочих.

Укажите вид используемой относительной величины.

**Задача 4.4**

По данным табл. 4.5 определите:

- 1) соотношение численности рабочих и служащих;
- 2) соотношение численности рабочих и руководителей;
- 3) соотношение численности мужчин и женщин на предприятии.

Укажите вид используемой относительной величины.

**Задача 4.5**

Имеются следующие данные об объеме производства продукции на предприятии:

	2006	2007	2008	2009	2010
Объем производства продукции в сопоставимых ценах, млрд руб.	14,3	16,2	15,0	16,8	17,2

Рассчитайте относительные величины динамики с постоянной базой.

#### **Задача 4.6**

По данным предыдущей задачи рассчитайте относительные величины динамики с переменной базой сравнения.

#### **Задача 4.7**

Предприятие планировало в 2010 г. по сравнению с 2009 г. увеличить объем продаж на 12,5 %. В 2010 г. уровень выполнения плана по продажам составил 107,4 %. Определите относительный показатель динамики объема продаж.

#### **Задача 4.8**

Молочный комбинат в мае 2010 г. превысил плановое задание по объему производства продукции на 7,5 %, произведя продукции на 14,6 млн руб. продукции сверх плана. Определите общий объем производства продукции за май месяц.

#### **Задача 4.9**

Предприятие планировало снизить себестоимость продукции в 2010 г. по сравнению с 2009 г. на 3 %. Фактическая себестоимость в 2010 г. составила 96 % от прошлогоднего уровня. Определите относительную величину выполнения плана.

#### **Задача 4.10**

Определите относительную величину динамики доли продукции на экспорт в общем объеме производства, если известно, что в анализируемом периоде рост общего объема производства составил 113,7 %, а объем производства продукции на экспорт увеличился на 5,8 %.

#### **Задача 4.11**

В прошлом году объем грузооборота по грузовому автотранспортному предприятию составил 240 млн т-км. Планом текущего года было предусмотрено довести грузооборот до 250,5 млн т-км. Фактический объем грузооборота в текущем году составил 269,3 млн т-км.

Определите: 1) относительную величину планового задания по росту грузооборота; 2) относительную величину выполнения плана по грузообороту; 3) относительную величину динамики грузооборота.

#### **Задача 4.12**

В анализируемом году потребление электроэнергии по каждому из двух регионов составило: I регион – 7,3 млрд кВт-ч.; II регион – 9,79 млрд кВт-ч. Численность населения по регионам была равна: по первому региону на начало года – 1,52 млн чел., на конец года 1,46 млн чел.; по второму региону на начало года 1,86 млн чел.; на конец года – 1,91 млн чел.

Определите: 1) относительные показатели уровня экономического развития (по производству электроэнергии на душу населения) по регионам; 2) относительные показатели сравнения уровней экономического развития регионов (по производству электроэнергии на душу населения).

#### **Задача 4.13**

По двум промышленным предприятиям за отчетный год имеются следующие отчетные данные:

Предприятие	Объем производства продукции, млрд руб.	Среднесписочная численность ППП, чел.
Предприятие А	34,50	1 200
Предприятие Б	64,03	1 900

Определите: 1) среднегодовую выработку одного работника ППП на каждом предприятии; 2) относительную величину сравнения среднегодовой выработки одного работника ППП по анализируемым предприятиям.

## Тема 5

### СРЕДНИЕ ВЕЛИЧИНЫ В СТАТИСТИКЕ

- 5.1. Сущность средних величин, их виды и значение в статистическом анализе.
- 5.2. Порядок расчета и свойства средней арифметической.
- 5.3. Порядок расчета и сфера применения средней гармонической.
- 5.4. Порядок расчета и сфера применения средней геометрической и средней квадратической.
- 5.5. Структурные средние.

Изучив данную тему, студенты должны:

- понимать сущность метода средних величин и его значение в статистическом анализе;
- уметь правильно выбирать вид средней величины;
- знать методы расчета средних величин.

#### **5.1. Сущность средних величин, их виды и значение в статистическом анализе**

##### **Сущность средних величин**

Метод средних величин широко используется в статистическом анализе.

*Средняя величина – обобщенная количественная характеристика признака в статистической совокупности в конкретных условиях места и времени.*

Важный вклад в обоснование и развитие теории средних величин внес крупный бельгийский ученый XIX в. Адольф Кетле (1796 – 1874), член Бельгийской академии наук, член-корреспондент Петербургской академии наук. Согласно учению Кетле, закономерности и характеристики массовых процессов и явлений формируются под влиянием двух групп причин.

*Первая группа* включает общие для всех единиц совокупности причины, которые формируют типичный уровень признака для всех единиц данной совокупности и связаны с сущностью изучаемого явления.

*Вторая* – индивидуальные причины, формирующие специфические особенности отдельных единиц совокупности, а, следовательно, и их отклонения от типичного уровня.

Чем однороднее изучаемая совокупность, тем более случайный характер носят индивидуальные различия единиц совокупности. В средних

величинах эти индивидуальные различия случайного характера взаимопогашаются и, таким образом, обеспечивается проявление общих, типичных свойств, присущих совокупности.

Средние величины в статистическом анализе выполняют две основные функции:

- средние величины обобщают значение количественного признака по всем единицам совокупности. Эта функция выполняется всегда;

- средние величины представляют типичный уровень признака в совокупности. Для выполнения этой функции должны соблюдаться определенные условия применения метода средних величин.

Для того чтобы средние величины представляли типичный уровень признака в совокупности, необходимо соблюдать два основных условия:

- средние величины должны определяться по качественно однородной совокупности. При расчете средней величины для неоднородной совокупности будут погашаться не только случайные отклонения, но и существенные различия между отдельными единицами, которые могут свидетельствовать о зарождении новых тенденций;

- средние величины должны определяться для статистических совокупностей, значительных по численности. Формируя большие или значительные по численности единицы совокупности, мы обеспечиваем условия для проявления *закона больших чисел*.

Одним из методов статистики, который позволяет получить среднюю величину как типичную характеристику признака в совокупности, является *метод группировки* (см. тему 3). Если совокупность неоднородна, общие средние (по совокупности в целом) должны быть дополнены групповыми средними, т.е. средними, рассчитанными по качественно однородным группам.

Однако не следует сводить роль средних величин обязательно к характеристике типичных значений признаков в однородных по данному признаку совокупностях. На практике современная статистика довольно часто использует средние величины, обобщающие явно неоднородные явления. Например, средняя величина валового внутреннего продукта на душу населения в стране, среднее потребление различных продуктов – это характеристики государства как системы. В этом случае говорят о *системных средних*. Они могут характеризовать как пространственные системы, (государство, отрасль, регион), так и динамические системы, протяженные во времени (например, среднегодовая температура).

Сравнение системных средних по разным системам (объектам) придает им признаки типичности.

Например, среднегодовое потребление кофе на душу населения составляет в Финляндии – 11,6 кг, а в Великобритании – 2,4 кг [10, с. 34]. Это, конечно, не означает, что каждый житель в этих странах потребляет указанное количество кофе, но сравнение среднедушевого потребления кофе в Великобритании и Финляндии указывает на традиции или предпочтения жителей этих стран. Финляндия – страна, в которой предпочитают кофе, а Великобритания – страна с традициями употребления чая.

Следует иметь в виду, что даже типичная средняя не является раз и навсегда данной, неизменной характеристикой совокупности. С изменением условий существования совокупности изменяется и уровень средней величины. Поэтому типичность средней величины – понятие относительное, ограниченное как в пространстве, так и во времени.

Средняя величина характеризует изучаемую совокупность по одному признаку (исключение составляет *многомерная средняя*). Она отражает величину признака по отношению к единице совокупности, но не обязательно к какой-то конкретной единице.

Средняя величина сохраняет размерность (единицу измерения) осредняемого признака.

Для всесторонней характеристики совокупности целесообразно использовать систему средних величин по наиболее существенным признакам.

Например, совокупность рабочих предприятия можно охарактеризовать показателями средней выработки, средней заработной платы, среднего тарифного разряда, среднего стажа работы и др.

### **Значение средних величин в статистическом анализе**

Средние величины решают следующие статистические задачи:

- измеряют уровень развития явления;
- позволяют сравнивать уровень одних и тех же признаков по разным совокупностям;
- используются для расчета показателей вариации значений признака (среднего линейного отклонения, дисперсии, среднего квадратического отклонения, относительных показателей вариации);
- используются при проведении корреляционно-регрессионного анализа.

Средняя величина как обобщающая количественная оценка совокупности не дает исчерпывающей характеристики этой совокупности. Поэтому средние величины используют в сочетании с другими обобщающими показателями: показателями вариации, структурными средними.

## Виды средних величин

Средние величины делятся на две основные категории: *степенные* и *структурные средние* (рис. 5.1).

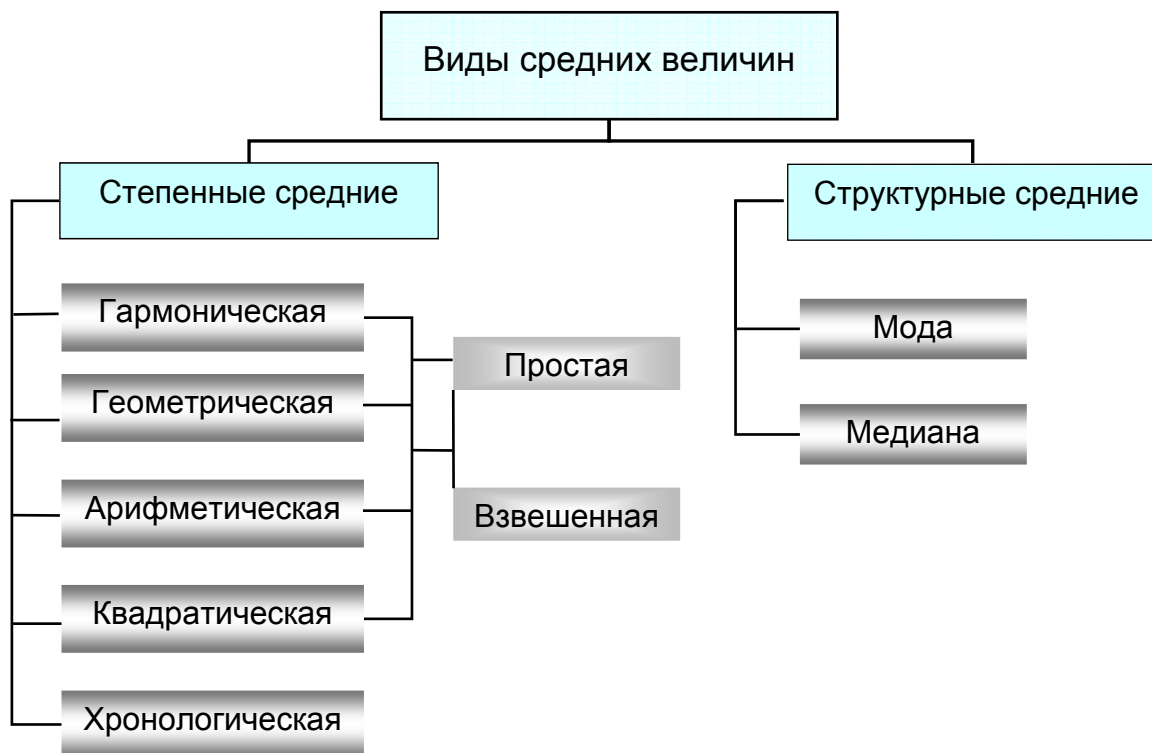


Рис. 5.1. Виды средних величин в статистике  
Источник: [4, с. 317]

В данной теме рассматриваются относящиеся к степенным средним такие их виды как средняя гармоническая, средняя геометрическая, средняя арифметическая и средняя квадратическая, а также *структурные средние*, представленные модой и медианой. Хронологические средние характеризуются в теме 8 «Статистическое изучение динамики социально-экономических явлений».

Выбор того или иного вида средних величин зависит от цели исследования, экономической сущности осредняемого показателя и характера имеющихся исходных данных.

Сущность степенных средних раскрывается через понятие «*определяющее свойство степенных средних*».

*Определяющим свойством степенных средних называется их способность сохранять неизменными некоторые обобщающие количественные характеристики статистической совокупности.*



Обобщающая количественная характеристика совокупности, которая не изменяется при замене всех индивидуальных значений признака на его среднюю величину, называется *определяющим показателем*.

Определяющий показатель, выраженный математически, называют определяющей функцией.

Каждый вид степенной средней имеет свой определяющий показатель и свою определяющую функцию.

На основе определяющей функции посредством замены индивидуальных значений признака средней величиной, составляется уравнение средней величины. Решая уравнение средней, получают формулу для расчета средней величины определенного вида.

## 5.2. Порядок расчета и свойства средней арифметической

### Определяющий показатель средней арифметической величины и порядок ее расчета

Средняя арифметическая величина – наиболее распространенный вид средних в статистике.

*Определяющим показателем средней арифметической является сумма значений признака у всех единиц совокупности.* В соответствии с этим, ее определяющая функция записывается следующим образом:

$$\omega = \sum_{i=1}^n x_i, \quad (5.1)$$

где  $\omega$  – определяющий показатель арифметической средней;

$x_i$  – индивидуальные значения признака;

$n$  – численность статистической совокупности.

На основе формулы (5.1) составляем уравнение средней арифметической величины ( $\bar{x}$ ):

$$n\bar{x} = \sum_{i=1}^n x_i. \quad (5.2)$$

Решая это уравнение относительно  $\bar{x}$ , получим формулу *средней арифметической простой*:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}. \quad (5.3)$$

Формула средней арифметической простой используется в тех случаях, когда каждое значение признака встречается один раз, либо когда при повторении значений признака они не сгруппированы.

Наряду с формулой средней арифметической простой используют *формулу средней арифметической взвешенной*. Эту формулу применяют в тех случаях, когда повторяющиеся по нескольку раз значения признака сгруппированы и для каждой группы определена частота ( $f_i$ ) или частость ( $\omega_i$ )  $i$ -того значения признака.

Как правило, такие данные представляются в виде ранжированного ряда распределения, в котором все значения осредняемого признака располагаются в порядке возрастания или убывания.

Формула средней арифметической взвешенной может быть записана как

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i}, \quad (5.4)$$

где  $x_i$  – отдельные значения признака (варианта);

$f_i$  – частота  $i$ -того значения признака ( $\overline{1, m}$ );

$m$  – количество групп.

Если вместо абсолютных частот известны *частости* ( $\omega_i$ ), выступающие в качестве весов, формула средней арифметической взвешенной имеет следующий вид:

$$\bar{x} = \sum_{i=1}^m x_i \omega_i, \quad (5.5),$$

если частости выражены в долях ( $\sum_{i=1}^m \omega_i = 1$ ), и

$$\bar{x} = \frac{\sum_{i=1}^m x_i \omega_i}{100}, \quad (5.6),$$

если частости выражены в процентах ( $\sum_{i=1}^m \omega_i = 100$ ).

### **Типичные случаи расчета средней арифметической**

В зависимости от содержания исходной информации, можно выделить следующие три типичных случая расчета средней арифметической.

#### **1. Расчет средней арифметической в дискретном ряду распределения.**

В этом случае средняя арифметическая рассчитывается как средняя арифметическая взвешенная, на основе формул (5.4), (5.5) или (5.6) (в зависимости от исходных данных).

### Пример 5.1

Имеются данные о распределении рабочих производственного участка по тарифным разрядам (табл. 5.1)

Таблица 5.1

Распределение рабочих участка по тарифным разрядам

Группы рабочих по тарифным разрядам $x_i$	Численность рабочих $f_i$ , чел.	$x_i f_i$
2	3	6
3	4	12
4	9	36
5	8	40
6	6	36
Итого	30	130

Определим средний тарифный разряд рабочих по участку. Так как рабочие сгруппированы по тарифным разрядам и известна численность рабочих каждого разряда, используем формулу (5.4). Для выполнения промежуточных расчетов в табл. 5.1 введена графа, представляющая произведения значений признака ( $x_i$ ) на его частоту ( $f_i$ ).

Выполним расчет среднего тарифного разряда рабочих:

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i} = \frac{130}{30} = 4,33 \text{ (разряда)}.$$

### 2. Расчет средней арифметической в интервальном ряду распределения.

Расчет средней арифметической в интервальном ряду распределения также выполняется на основе формулы средней арифметической взвешенной. Применению формулы предшествует операция замены интервала значений осредняемого признака в каждой группе на его среднее значение. Середина каждого интервала определяется простым расчетом:

$$x'_i = \frac{x_{H_i} + x_{B_i}}{2}, \quad (5.7)$$

где  $x'_i$  – середина  $i$ -того интервала (группы);

$x_{H_i}$ ,  $x_{B_i}$  – нижнее и верхнее значения признака в  $i$ -том интервале (группе).

### Пример 5.2

Имеются данные о распределении промышленных предприятий региона предприятий по среднегодовой стоимости основных производственных средств (табл. 5.2).

Таблица 5.2

**Распределение промышленных предприятий по среднегодовой стоимости основных производственных средств (ОПС)**

Группы предприятий по стоимости ОПС, млрд руб.	Число предприятий $f_i$	Удельный вес предприятий $\omega_i$ , % к итогу
3,6 – 5,0	2	10,0
5,0 – 6,4	4	20,0
6,4 – 7,8	6	30,0
7,8 – 9,2	5	25,0
9,2 и более	3	15,0
Итого	20	100,0

Необходимо определить среднюю стоимость ОПС по совокупности предприятий.

Сначала определим середину каждого интервала: середина первого интервала равна 4,3 ( $x'_1 = \frac{3,6+5,0}{2}$ ), середина второго интервала 5,7 ( $x'_2 = \frac{5,0+6,4}{2}$ ) и т.д. Пятый интервал (последняя группа) является открытым. Для преобразования этого интервала его величина условно принимается равной величине предшествующего интервала. В соответствии с этим, середина пятого интервала составит 9,9 ( $x'_5 = \frac{9,2+10,6}{2}$ ). Результаты определения середины интервалов представим в графе 2 табл. 5.3. Далее определяем для каждого интервала (группы) произведения  $x'_i f_i$  и  $x'_i \omega_i$  и результаты записываем в графы 5 и 6 соответственно.

Определим среднюю арифметическую взвешенную, используя в качестве весов частоту  $f_i$ :

$$\bar{x} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{i=1}^m f_i} = \frac{146,2}{20} = 7,31 \text{ (млрд руб.)}.$$

Выполним этот же расчет, используя в качестве весов частоты ( $\omega_i$ ):

$$\bar{x} = \frac{\sum_{i=1}^m x_i \omega_i}{100} = \frac{731,0}{100} = 7,31 \text{ (млрд руб.)}$$

Таблица 5.3

Расчетная таблица

Группы предприятий по стоимости ОПС, млрд руб.	Середина интервала, $x'_i$	Число предприятий $f_i$	Удельный вес предприятий $\omega_i$ , % к итогу,	$x'_i f_i$	$x'_i \omega_i$
1	2	3	4	5	6
3,6 – 5,0	4,3	2	10,0	8,6	43,0
5,0 – 6,4	5,7	4	20,0	22,8	114,0
6,4 – 7,8	7,1	6	30,0	42,6	213,0
7,8 – 9,2	8,5	5	25,0	42,5	212,5
9,2 и более	9,9	3	15,0	29,5	148,5
Итого	–	20	100,0	146,2	731,0

### 3. Расчет средней арифметической для совокупности в целом на основе групповых средних.

Если известны средние значения признака по группам анализируемой совокупности, то среднее значение признака по всей совокупности в целом определяется как средняя взвешенная из групповых средних:

$$\bar{X} = \frac{\sum_{j=1}^m \bar{x}_j n_j}{\sum_{j=1}^m n_j}, \quad (5.7)$$

где  $\bar{x}_j$  – среднее значение признака в  $j$ -той группе;  
 $n_j$  – численность  $j$ -той группы

#### Свойства средней арифметической

Средняя арифметическая обладает некоторыми математическими свойствами, более полно раскрывающими ее сущность и в ряде случаев используемыми при ее расчетах. Рассмотрим эти свойства.

1. Средняя от постоянной величины равна ей самой:

$$\bar{A} = A. \quad (5.8)$$

2. Произведение средней величины на сумму частот равно сумме произведений вариантов на их частоты:

$$\bar{x} \cdot \sum f_i = \sum x_i f_i. \quad (5.9)$$

Это свойство связано с определяющим показателем средней арифметической. Благодаря этому свойству средняя может быть использована для разного рода плановых и статистических расчетов как представитель или заместитель всех значений варьирующего признака.

3. Нулевое свойство средней. Оно устанавливает, что сумма отклонений индивидуальных значений признака от средней арифметической равна нулю:

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \quad (5.10)$$

и

$$\sum_{i=1}^m (x_i - \bar{x}) f_i = 0. \quad (5.11)$$

Доказательство:

$$\sum (x_i - \bar{x}) = \sum x_i - \sum \bar{x} = \sum x_i - n\bar{x} = \sum x_i - n \frac{\sum x_i}{n} = 0.$$

Это свойство может быть использовано для проверки правильности исчисления средней арифметической. Если

$$\sum (x_i - \bar{x}) \neq 0 \text{ и } \sum (x_i - \bar{x}) f_i \neq 0,$$

это значит, что средняя арифметическая рассчитана неверно.

4. Минимальное свойство средней арифметической. Оно устанавливает, что сумма квадратов отклонений индивидуальных значений признака от средней арифметической меньше суммы квадратов отклонений от любой другой произвольной величины  $a$ , если  $a \neq \bar{x}$ :

$$\sum (x_i - \bar{x})^2 < \sum (x_i - a)^2 \quad (5.12)$$

и

$$\sum (x_i - \bar{x})^2 < \sum (x_i - a)^2. \quad (5.13)$$

Доказательство:

$$\begin{aligned} \sum (x_i - a)^2 &= \sum (x_i - \bar{x} + \bar{x} - a)^2 = \sum [(x_i - \bar{x}) - (a - \bar{x})]^2 = \\ &= \sum [(x_i - \bar{x})^2 - 2(x_i - \bar{x})(a - \bar{x}) + (a - \bar{x})^2] = \end{aligned}$$

$$\begin{aligned} & \sum (x_i - \bar{x})^2 - 2(a - \bar{x}) \sum (x_i - \bar{x}) + \sum (a - \bar{x}) = \\ & \sum (x_i - \bar{x})^2 + \sum (a - \bar{x})^2. \end{aligned}$$

Следовательно,  $\sum (x_i - a)^2 > \sum (x_i - \bar{x})^2$  на величину  $\sum (a - \bar{x})^2$ .

5. Если каждое значение признака увеличить или уменьшить на одно и то же число  $A$ , то средняя арифметическая увеличится или уменьшится на то же число  $A$ :

$$\frac{\sum (x_i \pm A) f_i}{\sum f_i} = \frac{\sum x_i f_i}{\sum f_i} \pm \frac{A \sum f_i}{\sum f_i} = \bar{x} \pm A. \quad (5.14)$$

6. Изменение каждого варианта в одно и то же число раз изменяет среднюю во столько же раз:

$$\frac{\sum x_i h f_i}{\sum f_i} = \frac{h \sum x_i f_i}{\sum f_i} = h \cdot \bar{x}. \quad (5.15)$$

7. Если значение каждого веса разделить (или умножить) на одно и то же число  $d$ , то средняя от этого не изменится:

$$\frac{\sum x_i (d f_i)}{\sum d f_i} = \frac{d \sum x_i f_i}{d \sum f_i} = \bar{x}. \quad (5.16)$$

### 5.3. Порядок расчета и сфера применения средней гармонической

Средняя гармоническая применяется в тех случаях, когда известны значения осредняемого признака ( $x_i$ ), объем осредняемого признака ( $f_{hi}$ ), но не известна численность единиц совокупности.

*Определяющим показателем* средней гармонической является *сумма обратных значений признака*. Тогда определяющая функция средней гармонической может быть записана следующим образом:

$$\omega = \sum \frac{1}{x_i}. \quad (5.17)$$

Отсюда записываем уравнение средней гармонической:

$$n \cdot \frac{1}{x} = \sum \frac{1}{x_i}. \quad (5.18)$$

Исходя из формулы (5.18), получаем формулу средней гармонической простой:

$$\bar{x}_{\text{гарм}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}. \quad (5.19)$$

Формулу средней гармонической простой используют в тех случаях, когда веса средней гармонической ( $f_{hi}$ ), представляющие собой произведения значений признака на его частоту, равны между собой.

Если эти веса не равны между собой, то *используют формулу средней гармонической взвешенной*:

$$\bar{x}_{\text{гарм}} = \frac{\sum_{i=1}^n f_{hi}}{\sum_{i=1}^n \frac{f_{hi}}{x_i}}. \quad (5.20)$$

где  $f_{hi}$  – вес  $i$ -того значения признака (объем осредняемого признака для  $i$ -той группы).

Средняя гармоническая достаточно широко используется в экономических расчетах. Например, ее применяют, когда нужно определить среднюю заработную плату по группе предприятий (цехов) на основе данных о средней заработной плате и фонду заработной платы по каждому предприятию (цеху). Другой пример: зная трудоемкость каждого вида продукции и затраты рабочего времени на ее изготовление на основе средней гармонической определяют среднюю трудоемкость.

### Пример 5.3

Имеются данные о средних остатках по срочным вкладам за IV квартал 2010 г. (табл. 5.4). Необходимо определить средние остатки по трем филиалам сбербанка вместе.

Таблица 5.4

**Средние остатки средств по срочным вкладам по трем филиалам сбербанка**

№ филиала сбербанка	Средний остаток по срочным вкладам $x_i$ , млн руб.	Общая сумма остатков по срочным вкладам всех вкладчиков $f_{hi}$ , млн руб.
1	1,67	1897,8
2	2,80	5040,0
3	3,25	6987,5



В данном случае известны значения осредняемого признака ( $x_i$ ) – средние остатки по срочным вкладам по каждому филиалу, но неизвестно число вкладчиков (т.е. численность совокупности), поэтому используем среднюю гармоническую. Так как веса средней гармонической, относящиеся к отдельным группам, не равны между собой, используем формулу средней гармонической взвешенной:

$$\bar{x}_{\text{гарм}} = \frac{\sum_{i=1}^n f_{hi}}{\sum_{i=1}^n \frac{f_{hi}}{x_i}} = \frac{1\ 897,8 + 5\ 040,0 + 6\ 987,5}{\frac{1\ 897,8}{1,67} + \frac{5\ 040}{2,80} + \frac{6\ 987,5}{3,25}} = 2,74 \text{ (млн руб.)}$$

#### Пример 5.4

Три предприятия выпускают один и тот же бытовой прибор. Себестоимость изготовления прибора составляет на каждом предприятии: 50 тыс. руб.; 60 тыс. руб.; 30 тыс. руб.

Определить среднюю себестоимость прибора по трем предприятиям, если затраты на производство на каждом из предприятий составили по 600 000 тыс. руб.

Необходимость применения именно средней гармонической обусловлена содержанием исходных данных. Нам известны значения осредняемого признака, размер (объем) признака по каждому виду продукции, но неизвестно количество продукции. Так как в данном случае веса средней гармонической равны между собой, используем формулу средней гармонической простой:

$$\bar{x}_{\text{гарм}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{3}{\frac{1}{50} + \frac{1}{60} + \frac{1}{30}} = 42,86 \text{ (тыс. руб.)}$$

Таким образом, средняя себестоимость прибора по трем предприятиям составляет 42 860 руб.

### 5.4. Порядок расчета и сфера применения средней геометрической и средней квадратической

Средняя геометрическая ( $\bar{x}_{\text{геом.}}$ ) используется в анализе динамики для определения средних коэффициентов (темпов) роста.

Определяющим показателем средней геометрической является произведение значений признака у всех единиц совокупности. Тогда определяющая функция может быть записана следующим образом:

$$\omega = \prod_{i=1}^n x_i . \quad (5.21)$$

Отсюда представим уравнение средней геометрической:

$$\bar{x}^n = \prod_{i=1}^n x_i . \quad (5.22)$$

Решая его относительно средней величины, получаем формулу средней геометрической простой:

$$\bar{x}_{geom} = \sqrt[n]{\prod_{i=1}^n x_i} . \quad (5.23)$$

Средняя геометрическая может быть определена не только как простая, но и как взвешенная, по формуле

$$\bar{x}_{geom.} = \sqrt[n]{\prod_{i=1}^m x_i^{f_i}} , \quad (5.24)$$

где  $f_i$  – вес  $i$ -того значения признака.

### Пример 5.5

Анализируется изменение цены на продукт за два года. В первый год цена выросла в 3 раза, а во второй, по сравнению с первым, еще в 2 раза. Необходимо определить среднегодовой рост цены за этот период.

Используем формулу средней геометрической простой:

$$\bar{x}_{geom} = \sqrt[n]{\prod_{i=1}^n x_i} = \sqrt{3 \cdot 2} = 2,45 \text{ раза.}$$

Таким образом, при росте цены за два года в шесть раз, среднегодовой рост составил 2,45 раза.

*Средняя квадратическая* ( $\bar{x}_{kv.}$ ) используется при расчете показателей вариации (среднего квадратического отклонения), которые рассматриваются в следующей теме «Статистическое изучение вариации».

*Определяющим показателем средней квадратической является сумма квадратов значений признака у всех единиц совокупности.*

Определяющая функция средней квадратической имеет вид:

$$\omega = \sum_{i=1}^n x_i^2 . \quad (5.25)$$

Тогда уравнение средней квадратической выражается формулой

$$nx^{-2} = \sum_{i=1}^n x_i^2. \quad (5.26)$$

Отсюда формула средней квадратической простой записывается как

$$\bar{x}_{кв} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}. \quad (5.27)$$

Формула средней квадратической взвешенной имеет вид:

$$\bar{x}_{кв} = \sqrt{\frac{\sum_{i=1}^m x_i^2 f_i}{\sum_{i=1}^m f_i}}. \quad (5.28)$$

В табл. 5.5 представлены формулы расчета рассмотренных в данной теме степенных средних.

Таблица 5.5

**Формулы расчета степенных средних**

Виды средних величин	Простая	Взвешенная
Арифметическая ( $\bar{x}$ )	$\bar{x} = \frac{\sum x_i}{n}$	$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$
Квадратическая ( $\bar{x}_{кв}$ )	$\bar{x}_{кв} = \sqrt{\frac{\sum x_i^2}{n}}$	$\bar{x}_{кв} = \sqrt{\frac{\sum x_i^2 f_i}{\sum f_i}}$
Гармоническая ( $\bar{x}_{гарм}$ )	$\bar{x}_{гарм} = \frac{n}{\sum \frac{1}{x_i}}$	$\bar{x}_{гарм} = \frac{\sum f_{hi}}{\sum \frac{f_{hi}}{x_i}}$
Геометрическая ( $\bar{x}_{геом}$ )	$\bar{x}_{геом} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$	$\bar{x}_{геом} = \sqrt[n]{\sum f_i x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}$

Известно, что степенные средние различных видов, исчисленные по одной и той же совокупности, имеют различные количественные значения, устойчиво имеющие следующее соотношение:

$$\bar{x}_{гарм} < \bar{x}_{геом} < \bar{x} < \bar{x}_{кв}. \quad (5.27)$$

Это соотношение значений степенных средних называется *свойством мажорантности средних*.

## 5.5. Структурные средние

Важным направлением изучения статистических совокупностей является расчет и анализ структурных средних: *моды* и *медианы*. В отличие от средней арифметической структурные средние не являются обобщенной характеристикой совокупности и представляют значение признака у единиц, занимающих определенное положение в ранжированном ряду распределения.

### Мода

*Мода* – величина признака, имеющего наибольшую частоту (частотность).

Этот показатель используют в тех случаях, когда требуется охарактеризовать наиболее часто встречающуюся величину признака: наиболее распространенный уровень заработной платы на предприятии; модальную цену; модальный размер одежды, обуви и т.п.

Нахождение моды в атрибутивном ряду распределения, а также в дискретном вариационном ряду несложно. Нужно просто найти значение признака с наибольшей частотой (частотостью).

Например, для дискретного ряда распределения, представленного в табл. 5.1 мода равна 4 разряду. Именно это значение признака имеет наибольшую частоту, равную 9.

Расчет моды в интервальных вариационных рядах начинают с определения *модального интервала*. Затем интерполяцией между его границами определяется значение моды. Для интерполяции используется допущение о том, что в модальном и двух смежных с ним интервалах кривая распределения представляет собой квадратичную параболу, абсцисса вершины которой и является модой.

*Модальный интервал* – это интервал, в котором находится модальное значение признака.

Модальный интервал в интервальном ряду с равными интервалами находится по наибольшей частоте (частости).

В интервальном ряду с неравными интервалами модальным интервалом является интервал, имеющий наибольшую плотность распределения.

Расчет моды в интервальном ряду с равными интервалами производится по формуле

$$M_0 = x_{M_0} + h \frac{f_{M_0} - f_{M_0-1}}{(f_{M_0} - f_{M_0-1}) + (f_{M_0} - f_{M_0+1})}, \quad (5.28),$$

где  $x_{M_0}$  – нижняя граница модально интервала;

$h$  – величина модального интервала;

$f_{M_0}$  – частота модального интервала;

$f_{M_0-1}$ ,  $f_{M_0+1}$  – частота интервалов предшествующего и следующего за

модальным соответственно.

В интервальном ряду с неравными интервалами мода рассчитывается по формуле

$$M_0 = x_{M_0} + h \frac{\frac{f_{M_0}}{h} - \frac{f_{M_0-1}}{h_{M_0-1}}}{\left(\frac{f_{M_0}}{h} - \frac{f_{M_0-1}}{h_{M_0-1}}\right) + \left(\frac{f_{M_0}}{h} - \frac{f_{M_0+1}}{h_{M_0+1}}\right)}. \quad (5.29),$$

где  $h_{M_0-1}$  – величина интервала, предшествующего модальному;

$h_{M_0+1}$  – величина интервала, следующего за модальным.

Остальные обозначения в формуле (5.29) соответствуют обозначениям, принятым в формуле (5.28).

### Пример 5.6

Используя формулу (5.28), рассчитаем моду для вариационного ряда распределения с равными интервалами по данным табл. 5.2. Модальным является интервал с границами (6,4 – 7,8), т.к. он имеет наибольшую частоту ( $f = 6$ ).

Определим моду:

$$M_0 = 6,4 + 1,4 \frac{6 - 4}{(6 - 4) + (6 - 5)} = 7,33 \text{ (млрд руб.)}.$$

Таким образом, в данной совокупности предприятий самой распространенной величиной стоимости основных производственных средств является 7,33 млрд руб.

Моду можно определить и графически на основе гистограммы вариационного ряда.

В исследованиях встречаются распределения, где все варианты признака имеют одинаковую частоту, т.е. все значения признака одинаково модальны. Бывают распределения, в которых имеются две или более моды. В таких случаях говорят о бимодальном (две моды) или мультимодальном (более двух мод) распределении. Как правило, бимодальность (мультимодальность) указывает на качественную неоднородность совокупности по исследуемому признаку.

## Медиана

*Медиана – это значение признака у той единицы совокупности, которая стоит в середине ранжированного ряда распределения.*

Расчет медианы для дискретного и интервального рядов распределения имеет свои особенности.

Рассмотрим методику определения медианы в дискретном ряду.

Если число единиц совокупности ( $n$ ) является нечетным, то медиана равна значению признака у единицы с порядковым номером  $N_{Me}$ :

$$Me = X_{N_{Me}}, \quad (5.30)$$

где  $N_{Me} = \frac{n+1}{2}$ .

Если число единиц совокупности ( $n$ ) является четным, то медиана равна полусумме значений признака у единиц совокупности с порядковыми номерами  $N_1$  и  $N_2$ :

$$Me = \frac{X_{N_1} + X_{N_2}}{2}, \quad (5.31)$$

где  $N_1 = \frac{n}{2}$ ,  $N_2 = \frac{n+2}{2}$ .

## Пример 5.7

Рассмотрим определение медианы для дискретного ряда, представленного в табл. 5.1.

Так как  $n = 20$ , т.е. число рабочих является четным, медиану определяем по формуле (6.6):

$$Me = \frac{X_{N_1} + X_{N_2}}{2} = \frac{4 + 4}{2} = 4 \text{ (разряд)},$$

т.к.  $N_1 = \frac{n}{2} = \frac{20}{2} = 10$ ,  $N_2 = \frac{n+2}{2} = \frac{20+2}{2} = 11$ .

Таким образом, мы должны были определить разряд у десятого и одиннадцатого рабочего в ранжированном ряду распределения. Рабочие с такими номерами имеют 4 разряд.

Нахождение медианы в интервальных вариационных рядах требует предварительного определения *медианного интервала*.

*Медианный интервал – это интервал, для которого в графе накопленных частот мы впервые встречаем значение равное или большее половины численности совокупности.*

Расчетная формула медианы в случае интервального ряда имеет вид:

$$Me = X_{Me} + h_{Me} \frac{\sum f - S_{Me-1}}{f_{Me}}, \quad (5.32)$$

где  $X_{Me}$  – нижняя граница медианного интервала;

$h_{Me}$  – ширина медианного интервала;

$S_{Me-1}$  – накопленная частота интервала, предшествующего медианному;

$f_{Me}$  – частота медианного интервала;

$\sum f$  – сумма частот, или иначе – численность совокупности.

### Пример 5.7

По данным табл. 5.2 рассчитаем медиану. Для определения ее величины найдем кумулятивные частоты (табл. 5.6) и используем формулу (5.32).

Таблица 5.6

**Распределение предприятий  
по стоимости основных производственных средств**

Группы предприятий по стоимости ОПС, млрд руб.	Число предприятий $f_i$	Кумулятивная (накопленная) частота
3,6 – 5,0	2	2
5,0 – 6,4	4	6
6,4 – 7,8	6	12
7,8 – 9,2	5	17
9,2 и более	3	20
Итого	20	$x$

Сначала определим медианный интервал. Находим в графе накопленной частоты первое значение (при перемещении сверху вниз) равное или больше  $10 \left(\frac{20}{2}\right)$ . Таким значением является накопленная частота, равная 12. Ей соответствует интервал (6,4 – 7,8). Это и есть медианный интервал. Теперь определим значение медианы:

$$Me = 6,4 + 1,4 \cdot \frac{\frac{20}{2} - 6}{6} = 7,33 \text{ (млрд руб.)}$$

Таким образом, 50 % предприятий имеют стоимость ОПС менее 7,33 млрд руб., 50 % – более этой величины.

Медиана может быть определена и графически на основе кумуляты. Медиана, в отличие от средней, не является абстрактной величиной. Она находится точно в середине ряда, представляет собой реальное значение признака, соответствует определенному варианту и при этом наиболее точна в случае нечетного числа членов совокупности.

Медиана как характеристика ряда распределения не может заменить среднюю арифметическую. Медиана – это центр распределения численности единиц совокупности, а средняя арифметическая – это центр распределения отклонений значений признака от равнодействующей.

Медиана правильнее отражает типичный уровень признака в неоднородной совокупности, чем средняя арифметическая.

Медиана обладает свойством линейного минимума. Оно заключается в том, что сумма абсолютных значений отклонений признака у всех единиц совокупности от медианы есть величина минимальная:

$$\sum |x_i - Me| = \min \quad (5.33)$$

и

$$\sum |x_i - Me| \cdot f_i = \min . \quad (5.34)$$

Это свойство имеет важное значение для решения ряда практических задач. В частности его используют при проектировании расположения пунктов массового обслуживания: бензоколонок, магазинов, школ и т.п.

Моду и медиану особенно целесообразно использовать тогда, когда изучаемая совокупность содержит некоторое количество единиц с очень большим или очень малым значением варьирующего признака. Эти, не очень характерные для совокупности, значения вариантов, влияя на величину средней арифметической, не влияют на моду и медиану, что делает их очень ценными для экономико-статистического анализа.

Соотношение моды, медианы и средней арифметической указывает на характер распределения признака в совокупности, позволяет установить наличие асимметрии.

В симметричных рядах распределения все три показателя центра распределения совпадают, т.е. соблюдается равенство:  $\bar{x} = Me = M_0$ . Как известно, симметричным является распределение, в котором частоты любых двух вариантов, равностоящих в обе стороны от центра распределения равны между собой.

Чем больше расхождение между модой и средней арифметической, тем более асимметричен ряд. Притом, если имеет место соотношение  $\bar{x} < Me < M_0$ , делают вывод о наличии *левосторонней асимметрии*. Для *правосторонней асимметрии*, характерно следующее соотношение показателей центра распределения:  $\bar{x} > Me > M_0$ .



Наглядно асимметричные распределения с правосторонней и левосторонней асимметрией представляют кривые на рис. 5.1.

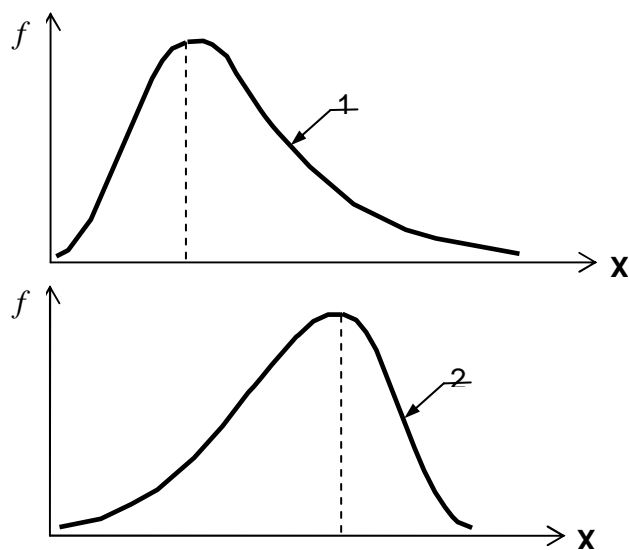


Рис. 5.1. Асимметричные ряды распределения:  
1 – с правосторонней асимметрией; 2 – с левосторонней асимметрией

Выявление характера асимметрии позволяет делать выводы о тенденциях в развитии исследуемых совокупностей. Так, наличие правосторонней асимметрии в распределении признаков, значения которых мы заинтересованы увеличивать (например, рентабельность, фондоотдачу, производительность труда, коэффициент оборачиваемости оборотных средств и др.), будет свидетельствовать о благоприятных тенденциях в развитии совокупности. Левосторонняя асимметрия в распределении значений таких признаков будет указывать на неблагоприятное положение.

В то же время, левосторонняя асимметрия в распределении признаков, значения которых мы заинтересованы уменьшать (например, трудоемкость продукции, фондоемкость, коэффициент текучести кадров и др.), будет указывать на благоприятное состояние дел.

## ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ

### Вопросы для самоподготовки и контроля знаний

1. Охарактеризуйте сущность средней величины.
2. Укажите основные функции средней величины в статистическом анализе.

3. Как Вы объясните понятие «типичная средняя величина»?
4. Какие виды средних величин применяются в статистике?
5. Что такое определяющее свойство степенных средних?
6. Какое значение имеет определяющее свойство и определяющая функции в применении метода средних величин?
7. Что является определяющим показателем средней арифметической величины?
8. При наличии каких исходных данных применяют среднюю арифметическую?
9. Объясните, в каких случаях используют формулу средней арифметической простой и формулу средней арифметической взвешенной.
10. В чем состоит нулевое свойство средней арифметической?
11. В чем заключается минимальное свойство средней арифметической?
12. При наличии каких исходных данных применяют среднюю гармоническую?
13. В каких случаях используют формулу средней гармонической простой?
14. Приведите конкретные примеры использования средней гармонической в экономических расчетах.
15. Какова сфера применения средней геометрической?
16. Что является определяющим показателем средней геометрической?
17. Какова сфера применения средней квадратической?
18. В чем состоит свойство мажорантности степенных средних?
19. Какие статистические показатели относятся к структурным средним?
20. Что такое мода?
21. Что такое модальный интервал и как он определяется в интервальном ряду с равными интервалами?
22. В чем особенность определения модального интервала в интервальном ряду с неравными интервалами?
23. Что такое бимодальное распределение?
24. Что такое медиана?
25. Как определяется медианный интервал в интервальном ряду?
26. В чем заключается свойство линейного минимума медианы?
27. Как соотносятся значения средней арифметической величины, медианы и моды в симметричном ряду распределения?
28. Какое соотношение значений средней арифметической величины, моды и медианы будет указывать на наличие правосторонней асимметрии в ряду распределения?

## Тест 5

1. Средняя величина осредняет значения ...
  - а) только количественных признаков;
  - б) только качественных признаков;
  - в) количественных и качественных признаков.
2. Средняя величина определяется на основе значений признака ...
  - а) у всех единиц изучаемой статистической совокупности;
  - б) только части единиц изучаемой статистической совокупности.
3. Средняя величина ... единицу измерения осредняемого признака.
  - а) сохраняет;
  - б) не сохраняет.
4. В статистической совокупности ... будут единицы, имеющие значение признака, равное средней величине.
  - а) обязательно;
  - б) не обязательно.
5. Типичность среднего значения признака свидетельствует ...
  - а) об однородности статистической совокупности;
  - б) о неоднородности совокупности.
6. Формулу средней арифметической взвешенной используют, если ...
  - а) значения признака не сгруппированы;
  - б) каждое значение признака встречается в совокупности один раз;
  - в) значения признака сгруппированы.
7. Если частоты осредняемого признака выразить в промилле, чему будет равен знаменатель при расчете средней арифметической?
  - а) 100;
  - б) 1 000;
  - в) 10 000.
8. Если все веса средней арифметической увеличить в одно и то же количество раз, то средняя арифметическая ...
  - а) не изменится;
  - б) уменьшится;
  - в) увеличится незначительно;
  - г) увеличится во столько же раз.
9. Если каждое значение признака уменьшить на некоторое постоянное число, то средняя арифметическая величина ...
  - а) останется неизменной;
  - б) уменьшится на эту постоянную величину.

10. Среднюю гармоническую величину используют, если ...

- а) известны значения осредняемого признака у единиц совокупности и их численность;
- б) известны значения осредняемого признака и объем осредняемого признака по группам и совокупности в целом.

11. Мода – это ...

- а) значение признака, которое имеет наибольшую частоту;
- б) наибольшее значение частоты в ряду распределения;
- в) наибольшее значение признака в совокупности.

12. Медиана – это ...

- а) значение признака у той единицы совокупности, которая стоит в середине ранжированного ряда;
- б) порядковый номер единицы, которая стоит в середине ранжированного ряда распределения;
- в) значение признака, имеющего наименьшую частоту.

13. Укажите соотношение между показателями центра распределения, которое характеризует правостороннюю асимметрию:

- а)  $\bar{X} > Me > M_0$ ;
- б)  $\bar{X} < Me < M_0$ ;
- в)  $\bar{X} = Me = M_0$ .

14. Левосторонняя асимметрия в распределении предприятий отрасли по уровню рентабельности свидетельствует ... тенденциях в распределении предприятий.

- а) о благоприятных;
- б) о неблагоприятных тенденциях.

15. Правосторонняя асимметрия в распределении студентов группы по уровню успеваемости свидетельствует ...

- а) о благоприятной ситуации с успеваемостью.
- б) о проблемах с успеваемостью в группе.

### Задания и задачи по теме

#### Задача 5.1

По данным табл. 5.7 определите средний стаж работы рабочих анализируемого участка. Укажите, что в данном случае является осредняемым признаком. Выбор вида средней величины, а также расчетной формулы обоснуйте.

Таблица 5.7

**Распределение рабочих участка по стажу работы**

Группы рабочих по стажу работы, лет	До 5 лет	5 – 10	10 – 15	15 и более
Численность рабочих, чел.	2	6	15	7

**Задача 5.2**

По данным за два месяца по трем цехам предприятия имеются следующие данные:

№ цеха	сентябрь		октябрь	
	Среднесписочная численность работников, чел.	Средняя месячная заработная плата, тыс. руб.	Средняя месячная заработная плата, тыс. руб.	Фонд заработной платы, млн руб.
Цех № 1	140	712	762	99,06
Цех № 2	200	654	710	149,10
Цех № 3	260	850	900	243,00

Определите:

- 1) среднемесячную заработную плату одного среднесписочного работника по трем цехам вместе в сентябре месяце;
- 2) среднемесячную заработную плату одного среднесписочного работника по трем цехам вместе в октябре месяце;
- 3) среднемесячную заработную плату одного среднесписочного работника по трем цехам вместе за два месяца;
- 4) относительную величину динамики средней месячной заработной платы одного среднесписочного работника по трем цехам в октябре по сравнению с сентябрем.

В каждом случае укажите вид используемой средней величины и обоснуйте ее выбор.

**Задача 5.3**

Имеются следующие данные об экспорте продукции металлургического комбината:

Вид продукции	Удельный вес продукции на экспорт, %	Стоимость продукции на экспорт, млн руб.
Сталь арматурная	40,0	32 100
Прокат листовой	32,0	42 500

Определите средний удельный вес продукции на экспорт по комбинату.

#### Задача 5.4

Определите среднюю себестоимость изделия по группе предприятий. Расчет выполните, используя два вида средних величин: среднюю арифметическую и среднюю гармоническую. Исходные данные представлены в таблице.

Группы предприятий по себестоимости одного изделия, тыс. руб.	Объем производства продукции, %	Затраты на производство продукции, %
110 – 115	9,0	8,2
115 – 120	18,0	17,2
120 – 125	24,0	23,9
125 и выше	49,0	50,7
Итого	100,0	100,0

#### Задача 5.5

По двум цехам имеются следующие данные о распределении рабочих по уровню месячной заработной платы за отчетный месяц:

Группы рабочих по уровню месячной заработной платы, тыс. руб.	Среднесписочная численность рабочих, чел.	
	Цех № 1	Цех № 2
500 – 550	32	17
550 – 600	36	40
600 – 650	150	220
650 – 700	70	110
700 – 750	32	83
Итого	320	470

Определите среднюю заработную плату по каждому из цехов и двум цехам вместе.

#### Задача 5.6

Работа автокомбината за месяц характеризуется следующими данными:

Номер автоколонны	Общие затраты на перевозку грузов, млн руб.	Средний месячный грузооборот автомашины, ткм	Себестоимость одного ткм, тыс. руб.
1	202,8	4 600	6,3
2	476,3	5 400	9,8
3	178,2	4 400	8,1

Определите:

- 1) среднюю себестоимость одного ткм;
- 2) среднее число автомашин в автоколонне;
- 3) средний месячный грузооборот автомашины по автокомбинату.

### Задача 5.7

По трем предприятиям, вырабатывающим один и тот же вид изделий, известны следующие данные за отчетный месяц:

Предприятие	Среднесписочная численность рабочих, чел.	Средняя месячная выработка одного рабочего, шт.	Себестоимость одного изделия, тыс. руб.
№ 1	120	500	30,0
№ 2	150	800	25,0
№ 3	300	850	20,8

Определите:

- 1) среднесписочную численность рабочих на одно предприятие;
- 2) среднюю выработку на одного рабочего по трем предприятиям;
- 3) среднюю себестоимость одного изделия по трем предприятиям.

### Задача 5.8

На основе данных о возрастной структуре производственного оборудования предприятия определите его средний возраст:

Группы оборудования по возрасту, лет	Удельный вес каждой группы, % к итогу
До 5	4,1
5 – 10	20,1
10 – 15	25,6
15 – 20	18,6
20 и более	31,6
Итого	100,0

### Задание 5.9

По периодическим изданиям за текущий год, материалам сайта Национального статистического комитета Республики Беларусь приведите примеры использования средних величин в статистике населения, уровня жизни населения, образования, здравоохранения, правонарушений, промышленности, сельского хозяйства, цен, строительства.

### Задание 5.10

По результатам сдачи последнего экзамена по высшей математике в вашей группе определите средний балл.

### Задача 5.11

Используя данные нижеприведенной таблицы, определите средний размер семьи, а также моду и медиану.

Число членов семьи, чел.	2	3	4	5	6	7	Итого
Число семей, % к итогу	15	34	25	16	8	2	100

### Задача 5.12

По данным нижеприведенной таблицы рассчитайте показатели центра распределения и квантили:

Группы групп торфа по влажности, %	Количество проб
20 – 22	18
22 – 24	26
24 – 26	34
26 – 28	20
28 – 30	12
30 – 32	16
Итого	116

Сделайте выводы о характере распределения на основе анализа соотношения среднего значения признака, его модального и медианного значений.



## Тема 6

### СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВАРИАЦИИ

- 6.1. Ряд распределения – основа изучения вариации.
- 6.2. Графическое представление вариационных рядов распределения.
- 6.3. Показатели вариации количественных признаков.
- 6.4. Измерение вариации альтернативного признака.
- 6.5. Правило сложения дисперсий и его применение в статистическом анализе.

Изучив данную тему, студенты должны:

- знать основные характеристики рядов распределения;
- уметь графически представлять вариационные ряды распределения;
- знать порядок расчета показателей вариации и уметь их интерпретировать;
- понимать сущность правила сложения дисперсий и уметь применять его при проведении исследований в области экономики и управления организацией.

#### 6.1. Ряд распределения – основа изучения вариации

Построение рядов распределения является составной частью сводной обработки данных статистического наблюдения. *Ряд распределения* – одна из важнейших характеристик статистической совокупности.

*Ряд распределения* – это группировка, в которой известна численность единиц в группах или удельный вес каждой группы в общей численности совокупности.

Ряд распределения включает два основных элемента:

- значения признака или варианты ( $x_i$ );
- частоты ( $f_i$ ) или частоты ( $\omega_i$ ).

В теме 5 понятия частота и частость уже использовались. Здесь только напомним, что частота – это численность единиц в каждой группе, а частость – это частота, представленная в виде относительной величины (доли единицы, удельного веса).

Сумма всех частот ряда распределения равна численности всей совокупности и иначе ее называют *объемом распределения*.

Сумма частостей  $\sum_{i=1}^n \omega_i$  равна 1, если они выражены в долях единицы, или 100 %, если они выражены в процентах.

Общая схема ряда распределения такова: в совокупности, состоящей из  $n$  единиц, исследуемый варьирующий признак  $x$  принимает различные значения  $x_1, x_2, x_3, \dots, x_n$ . Каждое из этих значений имеет частоту  $f_1, f_2, f_3, \dots, f_n$  (или частоту  $\omega_1, \omega_2, \omega_3, \dots, \omega_n$ ). Поэтому табличная форма представления ряда распределения будет иметь следующий вид (табл. 6.1).

Таблица 6.1

**Табличная форма представления ряда распределения (общий вид)**

Значение признака $x_i$	Частота $f_i$	Частость	
		в % к итогу	в долях единицы
$x_1$	$f_1$	$\omega_1$	$\omega_1$
$x_2$	$f_2$	$\omega_2$	$\omega_2$
$x_3$	$f_3$	$\omega_3$	$\omega_3$
...	...	...	...
...	...	...	...
$x_n$	$f_n$	$\omega_n$	$\omega_n$
Итого	$\Sigma f_i = n$	$\Sigma \omega_i = 100$	$\Sigma \omega_i = 1$

Ряды распределения могут быть образованы по качественному и количественному признакам.

Ряды распределения, построенные по качественным (атрибутивным) признакам, называют *атрибутивными*. Примером атрибутивных рядов может служить распределение персонала предприятия по полу, уровню образования, категориям.

Ряды распределения, построенные по количественному признаку, называют *вариационными*.

Обратите внимание, вариационные ряды распределения строятся как *ранжированные*. *Ранжированным* является ряд, в котором все значения признака расположены в возрастающем (или убывающем) порядке.

Вариационные ряды в зависимости от характера вариации исследуемого признака могут быть построены как *дискретные* и как *интервальные*.

*Дискретные* ряды распределения формируют в тех случаях, когда анализируемый признак варьирует дискретно и в относительно нешироких пределах, что делает возможным перечислить в таблице все встречающиеся значения признака (например, распределение рабочих по разрядам, распределение семей по числу детей).

*Интервальные вариационные ряды* строят, если признак варьирует непрерывно, или дискретно, но в широких пределах. Например, ряд распределения предприятий по численности работников строят как интервальный, т.к. анализируемый признак – численность работников – хотя и является дискретным, может варьировать в очень широких пределах: от нескольких де-

сятков до нескольких тысяч человек. В этом случае невозможно отразить в таблице все значения признака и ряд строят как интервальный.

Пример дискретного ряда распределения представлен в табл. 6.2, а примером интервальных рядов распределения являются ряды, приведенные в табл. 6.3 и 6.4.

Таблица 6.2

**Распределение рабочих по тарифным разрядам**

Тарифный разряд рабочего $x_i$	Численность рабочих, имеющих этот разряд $f_i$	Накопленная частота, $S_i$
2	3	3
3	4	7
4	9	16
5	8	24
6	6	30
Итого	30	–

Таблица 6.3

**Распределение предприятий по стоимости основных производственных средств  
(данные условные)**

Группы предприятий по стоимости ОПС, млрд руб.	Число предприятий $f_i$	Частость $\omega_i$	Накопленная частота $S_i$	Накопленная частость $\omega_i^s$
3,6 – 5,0	2	0,10	2	0,10
5,0 – 6,4	4	0,20	6	0,30
6,4 – 7,8	6	0,30	12	0,60
7,8 – 9,2	5	0,25	17	0,85
9,2 – 10,6	3	0,15	20	1,00
Итого	20	1,00	–	–

Правила и принципы построения интервальных рядов распределения аналогичны принципам и правилам построения статистических группировок (см. тему 3). В случаях, когда интервальный вариационный ряд распределения построен с равными интервалами, можно проводить непосредственное сравнение степени заполненности интервалов группировки. Например, в ряду, представленном в табл. 6.3, наиболее заполненной (насыщенной) является группа предприятий со стоимостью ОПС от 6,4 до 7,8 млрд руб.

Если вариационный ряд распределения имеет группы с неравными интервалами (см. табл. 6.4), то частоты в отдельных интервалах непосредственно несопоставимы, т.к. зависят от ширины интервала. Для того чтобы можно было сравнить заполненность групп в случае неравных интервалов, используют специальный показатель – *плотность распределения*.

*Плотность распределения – это число единиц совокупности, приходящееся в среднем на одну единицу ширины интервала.*

Если плотность распределения определяется отношением частоты на ширину интервала, она будет *абсолютной* ( $P_i^{abc.}$ ) (табл. 6.4):

$$P_i^{abc.} = \frac{f_i}{h_i}. \quad (6.1)$$

Таблица 6.4

**Распределение предприятий по величине прибыли**  
(данные условные)

Группы предприятий по величине прибыли, млрд руб.	Ширина интервала $h_i$	Число предприятий $f_i$	Абсолютная плотность распределения $P_i$
0 – 2	2	12	6,0
2 – 5	3	15	5,0
5 – 10	5	26	5,2
10 – 20	10	30	3,0
20 – 40	20	42	2,1
40 – 80	40	25	0,625
Итого	–	150	–

В случае, если находят отношение частоты к ширине интервала, получают *относительную плотность распределения* ( $P_i^{отн.}$ ):

$$P_i^{отн.} = \frac{\omega_i}{h_i}. \quad (6.2)$$

Важной характеристикой рядов распределения является *накопленная, или кумулятивная, частота* (частость).

*Накопленная частота* показывает, сколько единиц совокупности имеют значение признака не более интересующей исследователя величины.

*Накопленная частость* показывает, какая доля единиц совокупности не превышает данное значение признака.

В качестве примера обратимся к данным табл. 6.3. Накопленная частота для интервала (6,4 – 7,8) равна 12. Это значит, что 12 предприятий в данной совокупности имеют стоимость ОПС, не превышающую 7,8 млрд руб.

Для целей анализа и сравнительной характеристики рядов распределения применяют *обобщающие показатели вариационного ряда*. Система таких показателей может быть наглядно представлена при сравнении особенностей нескольких рядов распределения.

Так, на рис. 6.1 кривые распределения имеют одинаковый размах вариации и характер распределения частот, но отличаются величиной варь-

рующего признака, являющегося центром группирования (распределения), что отмечено на оси абсцисс.

Характеристики центра распределения составляют, таким образом, одну из групп обобщающих показателей. В качестве них используют среднюю арифметическую, моду, медиану.

Представленные на рис. 6.2 кривые распределения 3 и 4 имеют один и тот же *центр распределения* и симметричное расположение частот вокруг него, но отличаются *размахом вариации*. Следовательно, кроме показателей центра распределения необходимо иметь представление о степени колеблемости признака и знать показатели *меры вариации*. Эти две группы показателей – показатели центра распределения и вариации признака – имеют широкое применение при принятии управленческих решений.

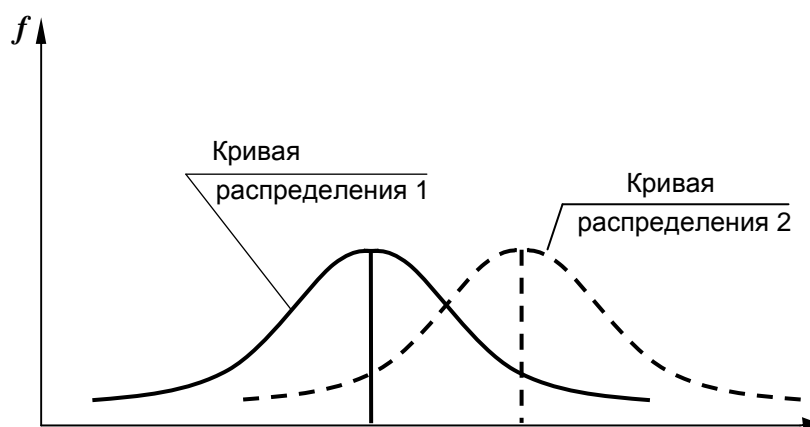


Рис. 6.1. Кривые с разными центрами группировки

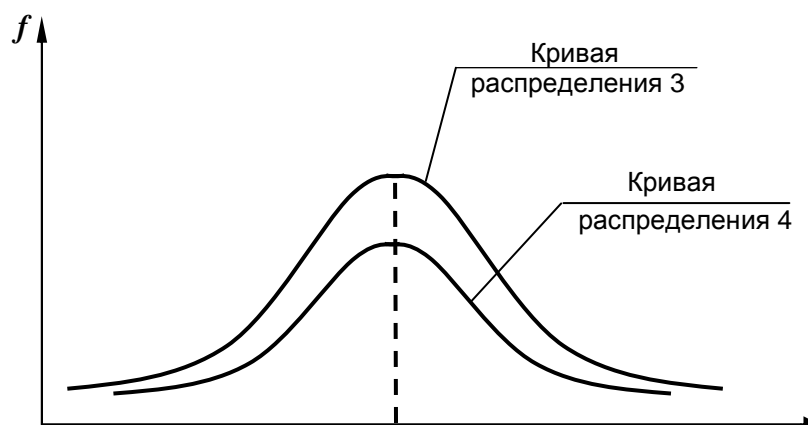


Рис. 6.2. Кривые распределения с различной вариацией признака

## 6.2. Графическое представление вариационных рядов распределения

Наряду с табличной формой представления рядов распределения используют также графическую форму их изображения. Графическое изображение рядов распределения облегчает их анализ и позволяет судить о форме распределения.

Для изображения вариационных рядов применяют линейные и плоскостные диаграммы, построенные в прямоугольной системе координат. Графически ряды распределения представляют с помощью следующих видов статистических графиков:

- *полигон распределения;*
- *гистограмма;*
- *кумулята;*
- *огива;*
- *кривая Лоренца.*

### Полигон распределения

*Полигон* используют для изображения дискретных вариационных рядов. Он представляет собой замкнутый многоугольник, абсциссами вершин которого являются значения варьирующего признака, а ординатами – соответствующие им частоты или частоты. Для замыкания полигона крайние вершины соединяются с точками на оси абсцисс, отстоящими на одно деление в принятом масштабе от  $x_{\min}$  и  $x_{\max}$ . Полигон, построенный по данным табл. 6.2, представлен на рис. 6.3.

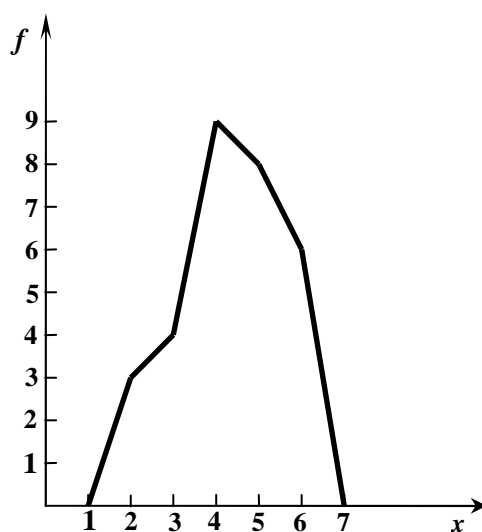


Рис. 6.3. Распределение рабочих участка по разрядам:  
 $x$  – тарифный разряд рабочего,  $f$  – число рабочих

## Гистограмма

*Гистограмма* используется для графического изображения интервальных вариационных рядов. При построении гистограммы для вариационного ряда с равными интервалами на оси абсцисс откладывают равные отрезки, которые в принятом масштабе соответствуют величине интервалов вариационного ряда, затем на отрезках строят прямоугольники, высота которых соответствует частотам (частостям) интервала. На рис. 6.4 изображена гистограмма ряда распределения предприятий по величине основных производственных средств.

Гистограмма может быть преобразована в полигон путем соединения середины верхних сторон прямоугольников отрезками прямых.

При построении гистограммы вариационного ряда с неравными интервалами по оси ординат следует наносить показатели плотности распределения (абсолютные или относительные). В этом случае высоты прямоугольников гистограммы будут соответствовать этим показателям.

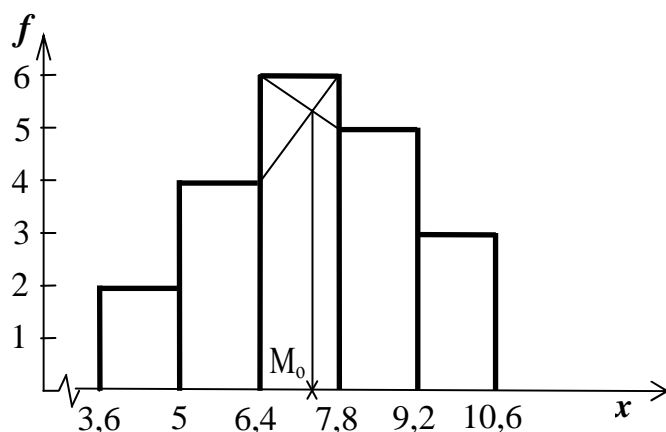


Рис. 6.4. Распределение предприятий по стоимости основных производственных средств:

$x$  – стоимость основных производственных средств, млрд руб.,  $f$  – число предприятий

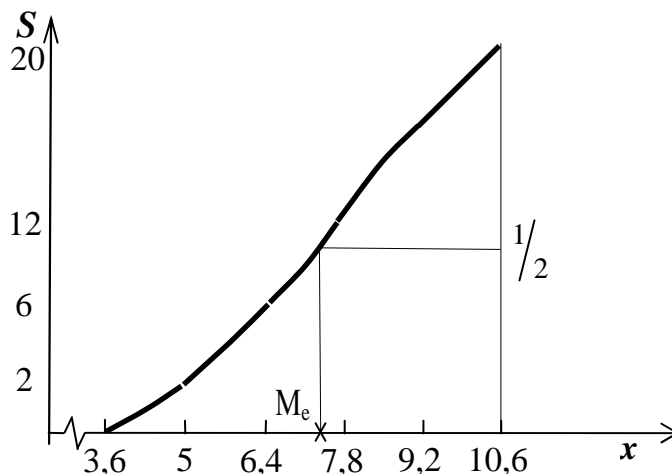
## Кумулята

В ряде случаев для изображения вариационных рядов используется кумулятивная кривая. Для ее построения необходимо рассчитать накопленные частоты.

При построении кумуляты интервального ряда распределения нижней границе первого интервала соответствует частота, равная 0, а верхней границе – вся частота данного интервала. Верхней границе второго интервала соответствует накопленная частота, равная сумме частот первых двух интервалов и т.д. На рис. 6.5 представлена кумулята, построенная на осно-

ве ряда распределения предприятий по среднегодовой стоимости ОПС (см. табл. 6.3).

Изображение вариационного ряда в виде кумуляты особенно удобно в экономических исследованиях, в частности для анализа концентрации производства, при проведении ABC-анализа.



**Рис. 6.5. Кумулята ряда распределения предприятий по стоимости основных производственных средств:**

$x$  – стоимость основных производственных средств, млрд руб.,  $S$  – накопленные частоты

### **Огива**

Если в прямоугольной системе координат по оси ординат откладывают значения признака по возрастанию, а по оси абсцисс – накопленные частоты, то получают графическое представление ряда распределения в виде *огивы*.

### **Кривая Лоренца**

Разновидностью кумулятивной кривой является *кривая Лоренца*. Этот вид графика используется для характеристики процессов концентрации, дифференциации, специализации и т.д.

Для построения кривой Лоренца, которую иначе называют *кривой концентрации*, необходимо иметь частотное распределение единиц исследуемой совокупности и взаимосвязанное с ним частотное распределение изучаемого признака. При этом для удобства вычисления и повышения аналитичности данных численность единиц совокупности обычно разбивается на равные группы: 10 групп по 10 % единиц в каждой или на 5 групп по 20 % единиц в каждой и т.д.



Кривая Лоренца строится в прямоугольной системе координат. На оси абсцисс откладываются накопленные частоты объема совокупности, а на оси ординат – накопленные значения объема признака. Полученная при соединении точек кривая линия и будет характеризовать степень концентрации.

Если распределение является строго равномерным, то первые 10 % единиц обладают 10 % объема признака, первые 20 % – соответственно 20 % объема признака и т.д. Такое распределение представляется прямой, проходящей из нижнего левого угла графика к верхнему правому углу, и называется *линией равномерного распределения*. Чем сильнее концентрация изучаемого признака (чем менее равномерно его распределение), тем заметнее кривая Лоренца отклоняется от линии равномерного распределения, и наоборот – чем слабее концентрация (чем равномернее распределение), тем ближе будет кривая распределения к прямой.

Рассмотрим пример анализа концентрации доходов населения условного региона в 2007 и 2010 гг. Для этого используем приведенную в табл. 6.5. группировку, в которой все население разбито на 10 равных групп таким образом, что первая группа объединяет 10 % населения с наименьшими доходами, вторая группа – следующие по уровню доходов 10 % населения и т.д. до последней, 10-ой группы, объединяющей 10 % населения с наибольшими доходами.

Таблица 6.5

**Распределение населения региона по уровню доходов в 2007 и 2010 гг. (в %)**

Группы населения, ранжированные по уровню среднедушевого дохода (по 10 % от общей численности населения)	2007		2010	
	Удельный вес в совокупном доходе	Накопленная частотность	Удельный вес в совокупном доходе	Накопленная частотность
1	4,3	4,3	3,2	3,2
2	6,1	10,4	4,8	8,0
3	7,1	17,5	6,1	14,1
4	8,1	25,6	7,2	21,3
5	9,1	34,7	8,4	29,7
6	10,1	44,8	9,7	39,4
7	11,2	56,0	11,3	50,7
8	12,6	68,6	13,2	63,9
9	14,3	82,9	15,8	79,7
10	17,1	100,0	20,3	100,0
Итого	100,0	–	100,0	–

Построенные по данным табл. 6.5 кривые концентрации доходов населения представлены на рис. 6.6.

Анализ данных табл. 6.5 позволяет заключить, что дифференциация населения по уровню доходов за рассматриваемый период усилилась. Так, если в 2007 г. 10 % беднейшего населения располагали 4,3 % совокупного дохода, а 10 % наиболее обеспеченных людей – 17,1 % дохода, то в 2010 г. эти показатели составили 3,2 % и 20,3 % соответственно. Если в соответственно 2007 г. доходы менее обеспеченной половины населения составляли 34,7 % общей величины, то в 2010 г. эта величина снизилась до 29,7 %.

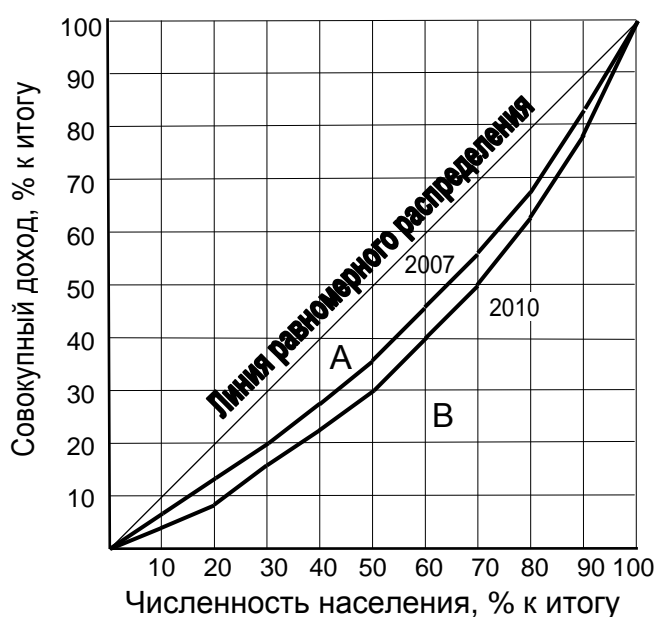


Рис. 6.6. Кривые концентрации доходов населения региона в 2007 и 2010 гг.

Этот вывод об усилении дифференциации населения по уровню доходов наглядно иллюстрирует график Лоренца, представленный на рис. 6.6. Кривая концентрации доходов в 2010 г. больше отклоняется от линии равномерного распределения, чем кривая концентрации доходов населения в 2007 г.

Широкое применение компьютеров в практической работе специалистов экономического профиля облегчает построение и анализ рядов распределения, а также их графическое представление.

### 6.3. Показатели вариации количественных признаков

Вариационные ряды распределения служат основой для анализа *вариации* изучаемого признака.

*Вариация характеризуется различием значений признака у разных единиц совокупности в один и тот же период или момент времени.*

Изменение значений признака у одного и того же объекта (единицы) в разные периоды или моменты времени называют динамикой и изучают с помощью показателей динамики, которые рассматриваются в теме 8.

Возникает вариация потому, что индивидуальные значения признака у единиц совокупности складываются под совокупным влиянием разных факторов, которые по-разному сочетаются и проявляются у разных единиц совокупности. В одних случаях индивидуальные значения признака весьма близки к средней арифметической. Это означает, что средняя хорошо представляет всю совокупность по величине анализируемого признака, т.е. является типичной (надежной) характеристикой совокупности. В этом случае может быть сделан вывод об однородности совокупности. В других случаях, наоборот, индивидуальные значения признака сильно отличаются от средней арифметической и тогда такая средняя не будет типичной (надежной) характеристикой совокупности, а сама совокупность будет считаться неоднородной.

Для количественной оценки меры вариации значений признака в теории статистики используют *показатели вариации*.

В статистическом анализе показатели вариации выполняют две функции:

- дают обобщенную оценку меры вариации значений признака в совокупности;
- позволяют сделать вывод о типичности (не типичности) средней величины и однородности (неоднородности) совокупности.

Состав показателей вариации представляет рис. 6.7.

Рассмотрим методику расчета указанных выше показателей вариации.

### **Абсолютные показатели вариации**

**Размах вариации.** Размах вариации  $R$  представляет собой разность между наибольшим и наименьшим значением анализируемого признака:

$$R = x_{\max} - x_{\min}, \quad (6.3),$$

где  $x_{\max}$ ,  $x_{\min}$  – соответственно максимальное и минимальное значения признака в совокупности.

Размах вариации – важный показатель колеблемости значений признака. Чем больше размах вариации, тем вероятнее то, что средняя окажется нетипичной.

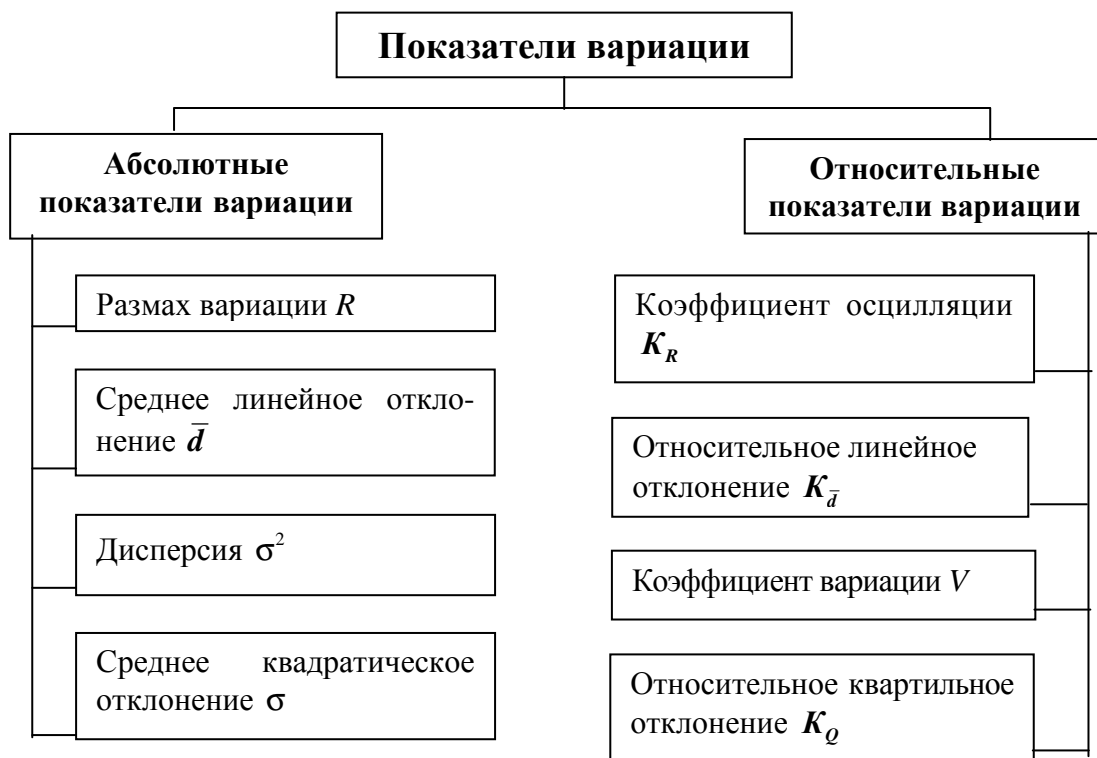


Рис. 6.7. Состав показателей вариации

Аналитический потенциал этого показателя невысок, т.к. он не учитывает характер и меру вариации значений признака вокруг его средней величины. Необходимы показатели, которые бы учитывали и измеряли вариацию признака по всем единицам совокупности, а не только по двум крайним значениям. Так как средняя арифметическая является обобщающей количественной характеристикой значений признака у всех единиц совокупности, большинство показателей вариации основано на измерении отклонений индивидуальных значений признака у всех единиц от средней.

**Среднее линейное отклонение.** Среднее линейное отклонение представляет собой среднюю арифметическую из абсолютных отклонений индивидуальных значений признака от его средней величины. В зависимости от исходных данных среднее линейное отклонение рассчитывается по следующим формулам:

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}, \quad (6.4),$$

$$\bar{d} = \frac{\sum_{i=1}^m |x_i - \bar{x}| f_i}{\sum_{i=1}^n f_i}, \quad (6.5)$$

где  $i = \overline{1, m}$ ,  $m$  – количество групп.

Формула (6.4) используется, если исходные данные не сгруппированы. Если построен вариационный ряд, применяют формулу (6.5).

Обратите внимание, что в этих формулах все отклонения индивидуальных значений от средней берутся по модулю. Если в этих расчетах не использовать знак модуля, то проявится нулевое свойство средней арифметической, в соответствии с которым сумма отклонений индивидуальных значений признака от его средней величины равна нулю.

### Пример 6.1

Рассмотрим пример расчета среднего линейного отклонения по данным табл. 6.6.

Таблица 6.6

**Распределение предприятий по среднегодовой стоимости основных производственных средств**

Группы предприятий по среднегодовой стоимости ОПС, млрд руб.	Число предприятий $f_i$	Середина интервала $x_i$	$ x'_i - \bar{x} $	$ x'_i - \bar{x} f_i$
А	1	2	3	4
3,6 – 5,0	2	4,3	3,01	6,02
5,0 – 6,4	4	5,7	1,61	6,44
6,4 – 7,8	6	7,1	0,21	1,26
7,8 – 9,2	5	8,5	1,19	5,95
9,2 – 10,6	3	9,9	2,59	7,77
Итого	20	–	–	27,44

Алгоритм расчета среднего линейного отклонения следующий:

1) для расчета среднего линейного отклонения найдем абсолютные отклонения середины интервалов, принимаемых в качестве вариантов признака ( $x'_i$ ) от средней величины ( $\bar{x}$ ). Средняя арифметическая величина для данного ряда распределения составляет 7,31 млрд руб. (см. пример 5.2 в теме 5). Результаты расчетов записываем в графу 3 табл. 6.6;

2) вычислим произведения отклонений  $|x'_i - \bar{x}|$  на их частоту ( $f_i$ ) и определим сумму этих произведений. Результаты записываем в графу 4 табл. 6.6;

3) определяем среднее линейное отклонение по формуле (6.5):

$$\bar{d} = \frac{27,44}{20} = 1,37 \text{ (млрд руб.)}.$$

Таким образом, в среднем величина абсолютных отклонений индивидуальных значений признака от его средней величины составляет 1,37 млрд рублей.

**Обратите внимание:** *среднее линейное отклонение сохраняет единицу измерения анализируемого признака.*

Наряду с использованием знака модуля, избежать проявления нулевого свойства средней арифметической можно, если отклонение каждого значения признака от его средней величины возвести во вторую степень. Этот подход и используется при расчете дисперсии.

**Дисперсия.** *Дисперсия ( $\sigma^2$ ) представляет собой средний квадрат отклонений индивидуальных значений признака от его средней величины и рассчитывается по следующим формулам:*

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}, \quad (6.6)$$

$$\sigma^2 = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 f_i}{\sum_{i=1}^m f_i}. \quad (6.7)$$

Формула (6.6) применяется, если каждое значение признака  $x_i$  встречается только один раз, или если данные не сгруппированы. Формула (6.7) используется для сгруппированных данных.

С учетом содержания исходных данных, дисперсия может быть рассчитана и по другой формуле:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2. \quad (6.8)$$

Получается эта формула путем преобразования формулы (6.6):

$$\begin{aligned} \sigma^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} - \bar{x}^2)}{n} = \frac{\sum_{i=1}^n x_i^2}{n} - \frac{2\bar{x} \sum_{i=1}^n x_i}{n} + \frac{n\bar{x}^2}{n} = \\ &= \overline{x^2} - 2\bar{x} \cdot \bar{x} + \bar{x}^2 = \overline{x^2} - (\bar{x})^2. \end{aligned}$$

Рассмотрим математические свойства дисперсии.

1. Дисперсия постоянной величины равна нулю:  $\sigma_A^2 = 0$ .
2. Если каждое из индивидуальных значений признака уменьшить или увеличить на одно и то же число  $A$ , то значение дисперсии не изменится:

$$\sigma_{x \pm A}^2 = \sigma_x^2. \quad (6.9)$$

3. Если каждое из индивидуальных значений признака уменьшить в одно то же число раз  $k$ , то дисперсия уменьшится в  $k^2$  раз:

$$\sigma_{x/k}^2 = \sigma_x^2 : k^2. \quad (6.10)$$

Среднее квадратическое отклонение при этом уменьшится в  $k$  раз.

4. Средний квадрат отклонений индивидуальных значений признака от любой величины  $A$ , будет всегда больше среднего квадрата отклонений от средней арифметической, если  $A \neq \bar{X}$ .

### Среднее квадратическое отклонение

Среднее квадратическое отклонение  $\sigma$  показывает, на сколько в среднем отклоняются индивидуальные значения признака от его средней величины. Оно выражается в тех же единицах измерения, что и признак и рассчитывается как корень квадратный из дисперсии:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}, \quad (6.11)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^m (x_i - \bar{x})^2 f_i}{\sum_{i=1}^m f_i}}. \quad (6.12)$$

Среднее квадратическое отклонение – наиболее распространенный и применяемый абсолютный показатель вариации. Его часто используют в качестве единицы измерения отклонений от средней арифметической. В зарубежной литературе этот показатель называют нормированным, или *стандартным, отклонением*.

По свойству мажорантности средних (см. тему 5) среднее квадратическое отклонение всегда больше среднего линейного отклонения. Если распределение признака близко к нормальному или симметричному распределению, то между  $\sigma$  и  $\bar{d}$  соблюдаются следующие соотношения:  $\sigma = 1,25\bar{d}$ ,  $\bar{d} = 0,8\sigma$ .

### Пример 6.2

Рассмотрим расчет дисперсии и среднего квадратического отклонения на примере данных, представленных в табл. 6.3. Результаты вспомогательных расчетов для определения показателей вариации приведены в табл. 6.7.

Таблица 6.7

**Распределение предприятий по среднегодовой стоимости ОПС**

Группы предприятий по среднегодовой стоимости ОПС, млрд руб.	Число предприятий, $f_i$	Середина интервала $x'_i$	$(x'_i - \bar{x})$	$(x'_i - \bar{x})^2$	$(x'_i - \bar{x})^2 f_i$
А	1	2	3	4	5
3,6 – 5,0	2	4,3	-3,01	9,06	18,12
5,0 – 6,4	4	5,7	-1,61	2,59	10,36
6,4 – 7,8	6	7,1	-0,21	0,04	0,24
7,8 – 9,2	5	8,5	1,19	1,42	7,10
9,2 – 10,6	3	9,9	2,59	6,71	20,13
Итого	20	–	–	–	55,95

Порядок расчета дисперсии и среднего квадратического отклонения следующий:

1) находим отклонение каждого варианта от средней арифметической  $(x'_i - \bar{x})$  и записываем их в графе 3. Затем эти разности возведем в квадрат и результаты заносим в графу 4;

2) квадраты разностей  $(x'_i - \bar{x})^2$  умножим на частоты  $f_i$  и запишем в графу 5 табл. 6.7. В итоговой строке получим сумму  $\sum (x'_i - \bar{x})^2 f_i$ ;

3) разделив указанную выше сумму произведений на сумму весов (частот или частостей), получаем дисперсию:  $\sigma^2 = \frac{55,95}{20} = 2,80$ ;

4) извлекая квадратный корень из дисперсии, получаем среднее квадратическое отклонение:  $\sigma = \sqrt{2,80} = 1,67$  млрд руб.

Таким образом, в среднем значения признака у единиц совокупности отклоняются от его средней величины на 1,67 млрд руб.

Следует отметить, что дисперсия, в отличие от  $R$ ,  $\bar{d}$ ,  $\sigma$ , не сохраняет единицу измерения признака, т.к. квадраты единиц измерения не имеют смысла.

**Относительные показатели вариации**

На основе абсолютных показателей вариации нельзя непосредственно сравнивать колеблемость признаков, если эти признаки имеют различные средние уровни.

Для целей сравнения изменчивости различных признаков в одной и той же совокупности, а также при сравнении вариации одного и того же признака в нескольких совокупностях (если средний уровень признака различен) используют *относительные показатели вариации*.



Принцип расчета относительных показателей вариации (ОПВ) отражает следующая формула:

$$ОПВ = \frac{\text{Абсолютный показатель вариации}}{\text{средняя арифметическая величина(или медиана)}}. \quad (6.13)$$

Рассмотрим состав и порядок расчета относительных показателей вариации.

*Коэффициент осцилляции ( $V_R$ ):*

$$V_R = \frac{R}{\bar{x}} \cdot 100 \% . \quad (6.14)$$

*Относительное линейное отклонение ( $V_{\bar{d}}$ ):*

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \cdot 100 \% . \quad (6.15)$$

*Коэффициент вариации ( $V$ ):*

$$V = \frac{\sigma}{\bar{x}} \cdot 100 \% . \quad (6.16)$$

*Относительное квартильное отклонение ( $V_Q$ ):*

$$V_Q = \frac{Q}{M_e} \cdot 100 \% \quad (6.17)$$

или

$$V_Q = \frac{Q_3 - Q_1}{2M_e} \cdot 100 \% . \quad (6.18)$$

Наиболее часто в практических расчетах применяется коэффициент вариации. Коэффициент вариации позволяет сравнить вариацию данного признака в разных совокупностях, а также вариацию разных признаков в данной совокупности. Если коэффициент вариации очень большой, то это означает, что значения признака сильно отличаются от его средней величины. Следовательно, такую среднюю нельзя считать надежной обобщающей характеристикой совокупности, а сама совокупность не является однородной. Совокупность считается однородной, а средняя величина типичной (надежной) ее характеристикой, если соблюдается условие:  $V \leq 33 \%$ .

### **Пример 6.3**

Определим коэффициент вариации для ряда распределения предприятий по среднегодовой стоимости ОПС (см. табл. 6.3). Ранее была определена средняя величина среднегодовой стоимости ОПС ( $\bar{x} = 7,31$  млрд руб.)

и среднее квадратическое отклонение ( $\sigma = 1,67$  млрд руб.) для этого ряда распределения. Тогда величина коэффициента вариации будет равна:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 \% = \frac{1,67}{7,31} \cdot 100 \% = 22,8 \% .$$

Таким образом, исходя из значения коэффициента вариации можно утверждать, что анализируемая совокупность предприятий однородна по величине среднегодовой стоимости ОПС, а среднее значение этого показателя (7,31 млрд руб.) типично для этой совокупности.

#### Пример 6.4

Рассмотрим пример использования коэффициента вариации в сравнительном анализе вариации взаимосвязанных показателей по разным совокупностям на основе данных, представленных в табл. 6.8.

Таблица 6.8

**Коэффициенты вариации выработки и заработной платы рабочих по двум цехам**

Наименование показателя	Значение показателя	
	Цех № 1	Цех № 2
1. Коэффициент вариации выработки рабочих, %	13	17
2. Коэффициент вариации заработной платы рабочих, %	14	39

Анализ представленных в табл. 6.8 показателей позволяет сделать вывод о том, эффективно ли организована оплата труда рабочих в каждом из цехов. Вывод о правильной организации оплаты труда возможен только в отношении цеха № 1, т.к. вариация выработки и заработной платы рабочих примерно одинаковы. А вот во втором цехе ситуация не столь благоприятна: при относительно невысокой вариации выработки рабочих (т.е. рабочие работают с относительно одинаковой производительностью), наблюдается высокая вариация их заработной платы (т.е. различие в оплате труда между рабочими значительно).

Показатели вариации находят широкое применение в экономических исследованиях, например, в финансовом менеджменте, инвестиционном анализе. Эти показатели используют для оценки хозяйственных рисков.

#### 6.4. Измерение вариации альтернативного признака

В предыдущем вопросе рассмотрен расчет показателей вариации для количественных признаков. В ряде случаев возникает необходимость

оценки вариации *качественных признаков с альтернативной вариацией*. Напомним, что признаками с альтернативной вариацией, или просто альтернативными признаками, называются признаки, которые имеют только два значения и каждая единица совокупности может принимать только одно из этих значений. Например, при изучении качества изготовленной продукции ее можно разделить на две группы: годную и брак; при изучении посещаемости студентами учебных занятий можно выделить студентов, не имеющих пропусков, и студентов, имеющих пропуски.

Для оценки вариации таких признаков принимают следующие допущения: наличие признака обозначается единицей, а его отсутствие – 0. Доля единиц, обладающих интересующим нас признаком, обозначается через  $p$ , доля остальных единиц составит:  $q = 1 - p$ . Определим для этих условий среднюю величину и дисперсию:

$$\bar{x}_p = \frac{\sum xf}{\sum f} = \frac{1 \cdot p + 0 \cdot q}{p + q} = p. \quad (6.19)$$

Тогда дисперсия альтернативного признака будет равна:

$$\sigma_p^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1 - p)^2 \cdot p + (0 - p)^2 \cdot q}{p + q} = \frac{q^2 p + p^2 q}{1} = pq. \quad (6.20)$$

Таким образом, дисперсия альтернативного признака равна произведению доли единиц, обладающих данным значением альтернативного признака, на долю единиц, не обладающих данным значением признака.

Среднее квадратическое отклонение альтернативного признака определяется по формуле

$$\sigma_p = \sqrt{pq}. \quad (6.21)$$

Максимальное значение дисперсии альтернативного признака равно 0,25 при  $p = 0,5$ . Любое другое значение  $p$  дает значение  $\sigma_p^2 < 0,25$ . Например, при  $p = 0,4$ ,  $\sigma_p^2 = 0,24$ ; при  $p = 0,7$   $\sigma_p^2 = 0,21$  и т.д.

Показатели вариации альтернативного признака широко используются в статистике, в частности, при проектировании выборочного наблюдения, обработке данных социологических обследований, статистическом контроле качества продукции, в ряде других случаев.

### Пример 6.5

В трех партиях готовой продукции, представленной для контроля качества, была обнаружена годная и бракованная продукция (табл. 6.9).

**Продукция, представленная на контроль качества**

Партия	Готовая продукция, шт.	Из них продукция	
		годная	бракованная
1	1 200	800	400
2	1 000	840	160
3	1 100	1 000	100

Определим в целом для всех партий продукции следующие показатели:

- 1) средний процент годной продукции и средний процент брака;
- 2) дисперсию, среднее квадратическое отклонение, коэффициент вариации годной продукции.

Произведем расчет данных показателей по данным табл. 6.9.

Средний процент годной продукции в трех партиях равен:

$$p = \frac{800 + 840 + 1000}{1\,200 + 1\,000 + 1\,100} = \frac{2\,640}{3\,300} = 0,8, \text{ или } 80\%.$$

Средний процент бракованной продукции составил:

$$q = 1 - 0,8 = 0,2, \text{ или } 20\%.$$

Дисперсия удельного веса годной продукции:

$$\sigma_p^2 = pq = 0,8 \cdot 0,2 = 0,16.$$

Среднее квадратическое отклонение удельного веса годной продукции:

$$\sigma_p = \sqrt{pq} = \sqrt{0,16} = 0,4.$$

Коэффициент вариации удельного веса годной продукции в общем выпуске продукции:

$$V = \frac{\sigma_p}{p} \cdot 100\% = \frac{0,4}{0,8} \cdot 100\% = 50\%.$$

### **6.5. Правило сложения дисперсий и его применение в статистическом анализе**

Из рассмотренных показателей вариации особое значение имеет дисперсия. На основе этого показателя можно решить задачу оценки меры влияния интересующего исследователя факторного признака ( $x$ ) на вариацию анализируемого результативного признака ( $y$ ). Для этого необходимо:

- провести аналитическую группировку, в которой изучаемая совокупность разбивается на однородные группы в зависимости от значений факторного признака  $x$ ;
- определить для каждой группы и совокупности в целом среднее значение результативного признака, а также его дисперсию;
- представить общую дисперсию результативного признака  $\sigma^2$  как сумму межгрупповой  $\delta^2$  и средней из внутригрупповых дисперсий  $\overline{\sigma_j^2}$ ;
- определить эмпирический коэффициент детерминации  $\eta_{\Theta}^2$ .

Как следует из приведенного выше алгоритма анализа, выполнение аналитической группировки позволяет разложить общую вариацию результативного признака на две дисперсии, одна из которых характеризует часть вариации, обусловленную влиянием группировочного факторного признака, а вторая – вариацию, происходящую под влиянием прочих факторов (кроме фактора, положенного в основу группировки).

Эта возможность разложения общей дисперсии определяется как *правило сложения дисперсий*:

$$\sigma^2 = \delta^2 + \overline{\sigma_j^2}, \quad (6.22)$$

где  $\sigma^2$  – общая дисперсия результативного признака  $y$ . Она измеряет вариацию результативного признака под влиянием всех факторов, обусловивших эту вариацию;

$\delta^2$  – межгрупповая дисперсия результативного признака  $y$ , иначе ее называют *факторной*. Она измеряет вариацию результативного признака, обусловленную влиянием факторного признака, положенного в основу группировки;

$\overline{\sigma_j^2}$  – средняя из внутригрупповых дисперсий. Оценивает вариацию результативного признака, обусловленную влиянием всех прочих факторов, исключая факторный признак, положенный в основу группировки. Иначе эту дисперсию называют *остаточной*.

Расчет общей дисперсии, межгрупповой и средней из внутригрупповых выполняется по следующим формулам:

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot f_i}{\sum_{i=1}^n f_i}, \quad (6.23)$$

где  $y_i$  –  $i$ -тое значение результативного признака;

$\bar{y}$  – общая средняя результативного признака по всей совокупности в целом;

$f_i$  – частота  $i$ -того значения результативного признака;

$$\delta^2 = \frac{\sum_{i=1}^m (\bar{y}_j - \bar{y})^2 \cdot n_j}{\sum_{i=1}^m n_j}, \quad (6.24)$$

где  $m$  – число групп;

$n_j$  – число единиц в  $j$ -той группе;

$\bar{y}_j$  – среднее значение результативного признака в  $j$ -той группе;

$$\bar{\sigma}_j^2 = \frac{\sum_{j=1}^m \sigma_j^2 \cdot n_j}{\sum_{j=1}^m i n_j}, \quad (6.25)$$

$$\sigma_j^2 = \frac{\sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2}{n_j}, \quad (6.26)$$

где  $\sigma_j^2$  – внутригрупповая дисперсия результативного признака в  $j$ -той группе.

Правило сложения дисперсии позволяет определить, какая доля вариации результативного признака обусловлена влиянием факторного признака, положенного в основу группировки. Показатель, представляющий эту долю, называют *эмпирическим коэффициентом детерминации*:

$$\eta_{\mathcal{D}}^2 = \frac{\delta^2}{\sigma^2}. \quad (6.27)$$

Корень квадратный из эмпирического коэффициента детерминации называют *эмпирическим корреляционным отношением*:

$$\eta_{\mathcal{D}} = \sqrt{\frac{\delta^2}{\sigma^2}}. \quad (6.28)$$

Эмпирическое корреляционное отношение используют для *оценки тесноты связи между факторным и результативным признаками*.

Эмпирический коэффициент детерминации изменяется в пределах от 0 до 1. Если  $\eta = 0$ , то факторный признак, положенный в основу группировки, не оказывает влияния на результативный. Если  $\eta = 1$ , то изменение результативного признака полностью объясняется влиянием факторного признака, положенного в основу группировки, а влияние прочих факторных признаков равно нулю.

### Пример 6.6

Рассмотрим правило сложения дисперсий. Имеются данные о стоимости недельного тура в Испанию в различные курортные города (табл. 6.10).

Таблица 6.10

**Группировка туристических фирм  
по географическим сегментам рынка в Испании**

Курорт	Число туристических фирм в сегменте $n_j$	Средняя цена недельного тура в группе $\bar{y}_j$ , долл.	Дисперсия цен в группе $\sigma_j^2$
Коста-Брава	7	951,4	6 755,6
Коста-дель-Соль	6	1 059,0	17 826,4
Итого	13	–	–

Вариация цен в обследованной группе туристических фирм, обусловленная различием в местоположении курорта, будет характеризоваться величиной межгрупповой дисперсии. Для ее определения сначала рассчитаем среднюю цену недельного тура по всем фирмам:

$$\bar{y} = \frac{\sum \bar{y}_j \cdot n_j}{\sum n_j} = \frac{951,4 \cdot 7 + 1\,059,0 \cdot 6}{7 + 6} = 1\,001,1 \text{ (долл.)}$$

Тогда межгрупповая дисперсия будет равна 2877,3:

$$\delta^2 = \frac{(951,4 - 1\,001,1)^2 \cdot 7 + (1\,059,0 - 1\,001,1)^2 \cdot 6}{7 + 6} = 2\,877,3$$

Вариация цен под влиянием всех прочих факторов, кроме местоположения курорта, будет характеризоваться средней из внутригрупповых дисперсий:

$$\overline{\sigma_j^2} = \frac{6\,755,6 \cdot 7 + 17\,826,4 \cdot 6}{7 + 6} = 11\,865,2$$

Вариация цен на недельные туры в Испанию, обусловленная влиянием всех факторов, формирующих их уровень в совокупности, состоящей из 13 фирм, определяется общей дисперсией, которую определяем по правилу сложения дисперсий:

$$\sigma^2 = 2\,877,3 + 11\,865,2 = 14\,742,5$$

Сопоставив межгрупповую дисперсию с общей дисперсией, определим эмпирический коэффициент детерминации:

$$\eta_{\Theta}^2 = \frac{2\,877,3}{14\,742,5} = 0,195, \text{ или } 19,5 \%$$

Таким образом, дисперсия цен на недельные туры в Испанию только на 19,5 % объясняется влиянием местоположения курорта, а на 80,5 % – влиянием других неучтенных факторов.

Определим эмпирическое корреляционное отношение:

$$\eta_{\mathcal{D}} = \sqrt{0,195} = 0,44.$$

Полученное значение эмпирического корреляционного отношения позволяет утверждать, что между местоположением курорта в Испании и ценой недельного тура существует слабая связь.

Для того, чтобы сделать более строгий вывод о существенности (несущественности) влияния того или иного факторного признака на результативный используют специальный статистический критерий: *F* – критерий (критерий Фишера). Если выполняется условие

$$F_{расч.} > F_{\alpha; k_1; k_2}, \quad (6.29),$$

делают вывод, о том, что влияние факторного признака на результативный существенно или статистически значимо с доверительной вероятностью  $p = (1 - \alpha)$ .

Охарактеризуем условные обозначения, принятые в формуле (6.29):

$F_{расч.}$  – расчетное значение *F*-критерия, которое определяют по следующим формулам (в зависимости от вида исходных данных):

$$F_{расч.} = \frac{\delta^2}{m-1} : \frac{\overline{\sigma_j^2}}{n-m}, \quad (6.30)$$

$$F_{расч.} = \frac{\eta_{\mathcal{D}}^2}{1-\eta_{\mathcal{D}}^2} \cdot \frac{n-m}{m-1}; \quad (6.31)$$

$F_{\alpha; k_1; k_2}$  – табличное значение *F*-критерия.

Параметры  $\alpha; k_1; k_2$  являются «входами» в таблицу, для определения табличного значения *F*-критерия (прил. 1);

$\alpha$  – уровень значимости. Уровень значимости задается исследователем самостоятельно и обычно принимается равным 0,05 или 0,01. Уровень значимости указывает вероятность ошибочности вывода о существенности (несущественности) влияния факторного признака на результативный;

$k_1 = m - 1$  и  $k_2 = n - m$  называют числом степеней свободы для межгрупповой и средней из внутригрупповых дисперсий соответственно.

Проверим существенность влияния факторного признака на результативный для примера, представленного в табл. 6.10.



Расчетное значение  $F$ -критерия равно:

$$F_{расч.} = \frac{0,195}{1-0,195} \cdot \frac{13-2}{2-1} = 2,66.$$

Табличное значение  $F$ -критерия при  $\alpha = 0,05$ ,  $k_1 = 1$ ,  $k_2 = 11$  составит:  
 $F_{0,05; 1; 11} = 4,84$ .

Так как  $2,66 < 4,84$ , с вероятностью  $p = 1 - \alpha = 1 - 0,05 = 0,95$  можно утверждать, что влияние местоположения курорта на цену недельного тура в Испанию несущественно. Вероятность ошибочности нашего вывода составляет 0,05.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Что такое ряд распределения?
2. Какие Вы знаете виды рядов распределения?
3. В каких случаях вариационный ряд распределения строят как дискретный?
4. В каких случаях ряд распределения строят как интервальный?
5. Что такое плотность распределения и как этот показатель используется в статистическом анализе?
6. Что такое кумулятивная частота?
7. С какой целью ряды распределения представляют графически?
8. В каких случаях для представления рядов распределения используют полигон?
9. В каких случаях для представления рядов распределения используют гистограмму?
10. Как строятся кумулята и огива?
11. Что представляет собой кривая Лоренца и каково ее значение в статистическом анализе?
12. Что такое вариация признака и чем она обусловлена?
13. Объясните различия между явлением вариации значений признака и динамикой значений признака.
14. Назовите абсолютные показатели вариации.
15. Какие показатели являются относительными показателями вариации и каков порядок их расчета?
16. Укажите математические свойства дисперсии.

17. Какие абсолютные показатели вариации сохраняют единицу измерения анализируемого признака?
18. Как используется коэффициент вариации в анализе однородности совокупности?
19. Как связаны между собой понятия однородность совокупности и типичность средней величины?
20. Как определяется дисперсия альтернативного признака?
21. Чему равно максимальное значение дисперсии альтернативного признака?
22. В чем состоит сущность правила сложения дисперсий?
23. Какие задачи статистико-экономического анализа решаются на основе правила сложения дисперсий?
24. Что характеризует межгрупповая дисперсия?
25. Что характеризует средняя из внутригрупповых дисперсий?
26. Как рассчитывается эмпирический коэффициент детерминации?
27. Что оценивает эмпирический коэффициент детерминации?
28. В каких пределах изменяется численное значение эмпирического коэффициента детерминации?
29. Что характеризует эмпирическое корреляционное отношение?
30. С помощью какого статистического критерия проверяют существование влияния факторного признака на результативный?

### Тест 6.1

1. Ряд распределения характеризует ...
  - а) взаимосвязь между значениями признаков;
  - б) изменение значений признака во времени;
  - в) взаимосвязь между значениями признака и частотой их проявления.
2. Ряд распределения, построенный по количественному признаку, называется ...
  - а) вариационным;
  - б) атрибутивным;
  - в) динамическим.
3. При непрерывной вариации количественного признака строят ... ряд распределения:
  - а) дискретный вариационный;
  - б) интервальный вариационный;
  - в) атрибутивный.
4. Плотность распределения – это ...

- а) число единиц совокупности, приходящееся на единицу ширины интервала;
- б) размер признака на единицу совокупности;
- в) наибольшее значение частоты признака.

5. При дискретной вариации количественного признака вариационный ряд распределения строят ...

- а) только как дискретный;
- б) только как интервальный;
- в) возможны оба из указанных выше вариантов.

6. Кумулятивная частота показывает, ...

- а) сколько единиц совокупности имеют значение признака не более средней его величины;
- б) сколько единиц совокупности имеют значение не более данной величины;
- в) сколько единиц совокупности имеют значение большее, чем медиана.

7. Для графического представления дискретного вариационного ряда используют ...

- а) гистограмму;
- б) полигон;
- в) структурную диаграмму.

8. Накопленные частоты используют при построении ...

- а) кумуляты;
- б) гистограммы;
- в) полигона.

9. Кривая Лоренца характеризует ...

- а) динамику явлений;
- б) процессы концентрации;
- в) асимметрию в распределении.

10. Какие статистические показатели не относят к показателям центра распределения:

- а) мода;
- б) коэффициент эксцесса;
- в) медиана;
- г) среднее значение признака.

11. Какие из абсолютных показателей вариации сохраняют единицу измерения признака:

- а) размах вариации;
- б) среднее линейное отклонение;
- в) дисперсия;
- г) среднее квадратическое отклонение.

## Тест 6.2

1. Дисперсия представляет собой ...

- а) средний размер отклонений индивидуальных значений признака от средней;
- б) средний квадрат этих отклонений;
- в) средний размер отклонений индивидуальных значений от произвольной постоянной величины.

2. Для сравнения вариации двух различных признаков необходимо использовать ...

- а) среднее линейное отклонение;
- б) среднее квадратическое отклонение;
- в) размах вариации;
- г) коэффициент вариации.

3. Для оценки степени однородности совокупности по величине анализируемого признака используют ...

- а) дисперсию;
- б) коэффициент вариации;
- в) моду;
- г) коэффициент эксцесса.

4. Если каждое значение признака увеличить на одну и ту же величину, то дисперсия ...

- а) не изменится;
- б) уменьшится на эту величину;
- в) увеличится на эту величину.

5. Если каждое значение признака увеличить в два раза, то дисперсия ...

- а) не изменится;
- б) увеличится в два раза;
- в) увеличится в четыре раза.

6. Средний стаж рабочих завода составляет 15 лет при дисперсии 9. Среднемесячная заработная плата рабочих составляет 650 тыс. руб., а среднее квадратическое отклонение равно 65 тыс. руб. Исходя из этих условий, выберите правильный вывод:

- а) вариация стажа больше вариации заработной платы рабочих;
- б) вариация стажа меньше вариации заработной платы;
- в) вариация заработной платы и стажа одинакова;
- г) сравнить вариацию указанных показателей невозможно.

7. Коэффициент вариации позволяет ...

- а) проверить типичность средней величины;

- б) сравнить вариацию разных признаков в данной совокупности;
- в) определить характер асимметрии;
- г) сравнить вариацию одного и того же признака в разных совокупностях.

8. Правило сложения дисперсий состоит в том, что ...

- а) общая дисперсия равна сумме групповых дисперсий;
- б) общая дисперсия равна сумме межгрупповой и средней из внутригрупповых дисперсий;
- в) средняя из внутригрупповых дисперсий равна общей дисперсии минус межгрупповая дисперсия.

9. Чему равен эмпирический коэффициент детерминации, если средняя из внутригрупповых дисперсий равна нулю:

- а) единице;
- б) нулю;
- в) колеблется от нуля до единицы.

10. Эмпирический коэффициент детерминации измеряет ...

- а) степень тесноты связи между анализируемыми признаками;
- б) вариацию, обусловленную влиянием всех факторов;
- в) долю вариации результативного признака, обусловленную влиянием изучаемого факторного признака.

### Задания и задачи по теме

#### Задача 6.1

Имеются следующие данные о величине межремонтного пробега автомобилей:

Величина межремонтного пробега, тыс. км	80 – 100	100 – 120	120 – 140	140 – 160	160 – 180
Количество автомобилей	10	60	100	26	14

Представьте данный вариационный ряд распределения графически в виде гистограммы и кумуляты. Используя графики, определите численное значение моды и медианы.

#### Задание 6.2

По результатам сдачи экзамена по высшей математике в зимнюю сессию постройте дискретный ряд распределения студентов вашей группы по полученным баллам. Дайте его графическое представление с помощью полигона распределения. Определите средний балл, моду и медиану.

### Задача 6.3

Акционерные общества области по среднесписочной численности работающих в отчетном году распределились следующим образом:

Группы ОАО по среднесписочной численности работающих, чел.	До 400	400 – 600	600 – 800	800 – 1000	1000 – 1200	1200 – 1400	1400 – 1600	1600 – 1800	Итого
Количество ОАО	11	23	36	42	28	17	9	4	170

Определите:

- среднее линейное отклонение;
- дисперсию двумя методами;
- среднее квадратическое отклонение;
- коэффициент вариации.

Сделайте выводы об однородности ОАО по среднесписочной численности работающих.

### Задача 6.4

Имеется следующее распределение рабочих двух производственных участков по стажу работы:

Группы рабочих по стажу работы, лет	Численность рабочих, чел.	
	Участок № 1	Участок № 2
0 – 5	2	7
5 – 10	15	25
10 – 15	20	12
15 – 20	3	8

Определите, на каком производственном участке состав рабочих по стажу работы более однороден.

### Задача 6.5

Средняя величина признака в совокупности равна 19, а средний квадрат индивидуальных значений этого признака – 397. Определите коэффициент вариации.

### Задача 6.6

Дисперсия признака равна 9, средний квадрат индивидуальных его значений – 130. Чему равно среднее значение признака?

### Задача 6.7

Средняя величина признака в совокупности равна 16, среднее квадратическое отклонение – 8. Определите средний квадрат индивидуальных значений этого признака.

### Задача 6.8

Удельный вес основных рабочих в общей численности рабочих в каждом из трех цехов составил: 80, 75 и 90 % соответственно. Определите дисперсию и среднее квадратическое отклонение доли основных рабочих по предприятию в целом, если численность всех рабочих составила: в первом цехе – 100 чел., во втором – 200 чел., в третьем – 150 чел.

### Задача 6.9

В таблице представлены данные об объеме производства продукции в поквартальном разрезе за год (млрд руб.):

Наименование показателя	Квартал			
	I	II	III	IV
Объем производства продукции, всего	150	120	160	180
В т.ч. продукция на экспорт	85	60	128	130

Определите за год дисперсию удельного веса продукции на экспорт.

### Задача 6.10

При изучении влияния стажа работы рабочих на производительность труда в цехе были получены данные, представленные в таблице:

Группы рабочих по стажу, лет	Численность рабочих в группах, чел.	Средняя часовая выработка рабочих в группах, изд.	Дисперсия выработки в каждой группе
До 5	30	20,0	3,0
5 и более	40	23,0	2,0

Определите:

- 1) среднюю из внутригрупповых дисперсий, межгрупповую и общую дисперсии;
- 2) эмпирический коэффициент детерминации и эмпирическое корреляционное отношение;
- 3) расчетное значение  $F$ -критерия. Сделайте выводы о существенности влияния стажа рабочих на уровень выработки.

## Тема 7

### ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

- 7.1. Сущность и значение выборочного наблюдения.
- 7.2. Основные способы формирования выборочной совокупности.
- 7.3. Средние и предельные ошибки выборочного наблюдения.
- 7.4. Определение необходимой численности выборки.
- 7.5. Малая выборка.
- 7.6. Распространение данных выборочного наблюдения на генеральную совокупность.

Изучив данную тему, студенты должны:

- понимать сущность и значение выборочного наблюдения;
- знать факторы, влияющие на точность результатов выборочного наблюдения;
- знать методику проведения выборочного наблюдения;
- уметь определять характеристики генеральной совокупности на основе данных выборочного наблюдения.

#### 7.1. Сущность и значение выборочного наблюдения

*Выборочное наблюдение – это несплошное наблюдение, при котором статистическому обследованию подвергаются не все единицы изучаемой совокупности, а лишь отобранные в соответствии с правилами выборочного метода.*

*Цель выборочного наблюдения – по характеристикам выборочной совокупности получить представление о характеристиках генеральной совокупности.*

При этом *генеральной совокупностью* называется статистическая совокупность, содержащая все явления или элементы, интересующие исследователя. *А совокупность единиц, отобранных по определенным правилам для непосредственного обследования, называется выборочной.*

Сущность выборочного наблюдения представляет рис. 7.1.

Для того чтобы по выборке сделать достоверные выводы о свойствах генеральной совокупности, выборочная совокупность должна быть **репрезентативной** (представительной), т.е. она должна полно и адекватно представлять свойства генеральной совокупности. Репрезентативность выборки достигается:

во-первых, *путем обеспечения объективности отбора единиц.* Нарушение принципа объективности отбора единиц, включение в выборку еди-



ниц, исходя из субъективного мнения исследователя, приводит к тому, что результаты такого исследования будут относиться не ко всей генеральной совокупности, а только к той ее части, которая была подвергнута наблюдению;

во-вторых, *в результате формирования достаточной по численности выборочной совокупности*. Соблюдение этого требования обеспечит проявление закона больших чисел.



Рис. 7.1. Сущность выборочного наблюдения

Выборочное наблюдение обладает рядом достоинств:

- оно обеспечивает экономию времени и ресурсов для проведения статистического исследования;
- позволяет оперативно получить результаты;
- заменяет сплошное наблюдение, когда его проведение невозможно (например, при проверке качества некоторых видов продукции, связанной с их разрушением);
- обеспечивает высокую достоверность результатов, т.к. уменьшение числа единиц наблюдения снижает ошибки регистрации, позволяет провести наблюдение более тщательно и квалифицированно;
- имеет широкую область применения. Выборочное наблюдение используется при опросах общественного мнения, при выявлении потребительских предпочтений, определении типичного потребителя, в управлении качеством продукции и др.

Проведение выборочного наблюдения предполагает выполнение следующих действий:

- 1) постановка цели наблюдения;
- 2) составление программы наблюдения;

- 3) определение границ генеральной совокупности;
- 4) правильный выбор способа формирования выборочной совокупности;
- 5) определение необходимой численности выборки;
- 6) регистрация признаков по программе у отобранных единиц;
- 7) обобщение данных наблюдения и расчет выборочных характеристик;
- 8) расчет ошибок выборки;
- 9) распространение выборочных характеристик на генеральную совокупность.

Теоретическую основу выборочного наблюдения составляют предельные теоремы закона больших чисел, прежде всего теоремы П.Л. Чебышева и А.М. Ляпунова. Они отражают основное свойство выборки: если численность выборки достаточно велика, то выборочные характеристики достаточно хорошо воспроизводят генеральные характеристики. При этом выборочные средние, а также их отклонения от генеральной средней имеют распределение, близкое к нормальному при любом распределении частот в генеральной совокупности. Указанное свойство *кривой выборочного распределения* используют для установления вероятности той или иной величины ошибки выборки.

Выборочный метод основан на использовании общепринятых условных обозначений характеристик генеральной и выборочной совокупностей, которые приведены в табл. 7.1.

Таблица 7.1

**Условные обозначения характеристик  
генеральной и выборочной совокупностей**

	Характеристики	Генеральная совокупность	Выборочная совокупность
1	Объем совокупности (численность единиц)	$N$	$n$
2	Численность единиц, обладающих данным значением альтернативного признака	$M$	$m$
3	Доля единиц, обладающих данным значением альтернативного признака	$p = \frac{M}{N}$	$\omega = \frac{m}{n}$
4	Доля единиц, не обладающих данным значением альтернативного признака	$q = 1 - p$	$1 - \omega$
5	Среднее значение количественного признака	$\bar{X} = \frac{\sum_{i=1}^N x_i}{N}$	$\tilde{x} = \frac{\sum_{i=1}^n x_i}{n}$
6	Дисперсия количественного признака	$\sigma^2 = \frac{\sum (x_i - \bar{X})^2}{N}$	$\sigma_B^2 = \frac{\sum (x_i - \tilde{x})^2}{n}$
7	Дисперсия доли альтернативного признака	$\sigma_p^2 = pq$	$\sigma_\omega^2 = \omega(1 - \omega)$

## 7.2. Основные способы формирования выборочной совокупности

Для того чтобы выборка была репрезентативной, отбор единиц из генеральной совокупности должен быть соответствующим образом организован. В каждом конкретном случае отбор единиц из генеральной совокупности будет зависеть от следующих факторов:

- сущности исследуемых явлений и процессов;
- размера и границ генеральной совокупности;
- необходимой численности выборки;
- меры вариации и характера распределения наблюдаемых признаков;
- наличия материальных и трудовых ресурсов для проведения выборочного наблюдения.

Разработка системы формирования выборочной совокупности включает решение следующих вопросов:

- 1) определение единицы отбора (*вида отбора*);
- 2) определение *метода отбора единиц*;
- 3) обоснование *способа организации выборки*.

Рассмотрим каждый из аспектов формирования выборочной совокупности более подробно.

В качестве *единиц отбора* могут выступать *непосредственно единицы генеральной совокупности*. Такой вид отбора называют *индивидуальным*. Возможен отбор *групп единиц совокупности* или, как говорят иначе, *гнезд или серий единиц*. Этот вид отбора называют *групповым*. В случае сочетания индивидуального и группового видов отбора говорят о *комбинированном виде отбора*.

Определение *метода отбора единиц* зависит от того, участвуют ли отобранные единицы в дальнейшей выборке или нет. В соответствии с этим принципом различают два метода отбора: *повторный* и *бесповторный*.

*Повторным* называется такой метод отбора, при котором отобранная однажды единица возвращается обратно в генеральную совокупность и снова участвует в выборке. При повторном отборе для всех единиц отбора сохраняется постоянная вероятность попасть в выборку.

*Бесповторным* называется такой метод отбора, при котором отобранная однажды единица обратно в генеральную совокупность не возвращается. При этом вероятность попадания отдельных единиц в выборку все время изменяется. Для остающихся в генеральной совокупности единиц она увеличивается. Бесповторный отбор дает более точные результаты по сравнению с повторным, т.к. при одном и том же объеме выборки наблюдению подвергаются все новые и новые единицы.

Обоснование *способа организации выборки* заключается в *установлении* правил образования выборочной совокупности в целом. В практике выборочных обследований используют следующие основные способы организации выборки (виды выборки):

- собственно-случайная (простая);
- механическая;
- типическая (стратифицированная);
- серийная;
- комбинированная.

Рассмотрим сущность и сферу применения каждого из указанных видов.

*Собственно-случайная выборка* заключается в отборе единиц из всей генеральной совокупности наугад без каких-либо элементов системности. Технически собственно-случайная выборка проводится через метод жеребьевки или по таблице случайных чисел, а также с помощью электронных таблиц.

Собственно-случайная выборка может быть повторной и бесповторной.

Она дает хорошие результаты, если между единицами исследуемой совокупности нет резких различий и есть исчерпывающий перечень всех единиц генеральной совокупности. Наиболее часто этот вид выборки используют в сочетании с другими видами выборки.

*Механическая выборка* в основном применяется в случаях, когда генеральная совокупность каким-либо образом упорядочена, т.е. имеется определенная последовательность в расположении единиц (список работников или студентов по алфавиту, телефонные номера респондентов, номера домов и квартир).

Механическая выборка удобна и в тех случаях, когда мы заранее не можем составить список единиц генеральной совокупности. Например, выборка берется из постепенно формирующейся во времени совокупности или из практически бесконечной совокупности. Так, при обследовании покупок можно наблюдать каждого десятого входящего в магазин покупателя; при контроле качества продукции проверять, например, каждую пятую сходящую со станка деталь и т.п.

При проведении механической выборки необходимо установить *шаг отсчета и начало отсчета единиц*. *Шаг отсчета* определяет расстояние между отбираемыми единицами. *Начало отсчета* указывает номер той единицы генеральной совокупности, которая должна быть обследована первой.

Например, если численность генеральной совокупности составляет 100 000 единиц и нужно получить 2%-ную выборку, то шаг отсчета составит:  $(100\ 000 : 2\ 000) = 50$ . Это означает, что из генеральной совокупности

должна выбираться каждая 50-ая единица. Генеральная совокупность разбивается на столько групп, сколько единиц нужно отобрать и из каждой группы, с установленным интервалом, выбирают единицы в выборочную совокупность.

Выбор начала отсчета зависит от способа расположения единиц в генеральной совокупности.

Если единицы не упорядочены по величине изучаемого признака, в каждой группе берут любую единицу и затем последовательно, через установленный шаг отсчета, из 2-ой, 3-ей и т.д. группы.

При упорядоченном расположении единиц генеральной совокупности по величине изучаемого признака, в каждой группе берут единицу, находящуюся в середине.

Обратите внимание, что механическая выборка всегда неповторная.

*Типическая (стратифицированная) выборка* применяется тогда, когда генеральная совокупность неоднородна по признакам, подлежащим изучению. До формирования типической выборки генеральная совокупность разбивается на типические группы. Типические группы могут содержать одинаковое или различное число единиц отбора. В первом случае выборочная совокупность формируется с одинаковой долей отбора из каждой группы, во втором – с долей, пропорциональной ее доле в генеральной совокупности. Отбор из каждой группы осуществляется в случайном (повторном или неповторном) либо механическом порядке. При типической выборке устраняется влияние межгрупповой вариации изучаемого признака на точность ее результатов, т.к. обеспечивается обязательное представительство в выборочной совокупности каждой из типических групп, что повышает точность результатов выборки.

*Серийная выборка* удобна в случаях, когда единицы совокупности объединены в небольшие группы или серии (гнезда). Например, это могут быть упаковки продукции; партии товара; студенческие группы; бригады и др.

Сущность серийной выборки заключается в собственно-случайном или механическом отборе серий, внутри которых производится сплошное обследование единиц. Серии могут быть равновеликими и неравновеликими.

При серийной выборке не обеспечивается представительство каждой из серий и не устраняется влияние межсерийной вариации изучаемого признака на результаты выборки. Если эта вариация значительна, она приведет к увеличению случайной ошибки выборки.

*Комбинированная выборка.* В практике статистических обследований помимо рассмотренных способов отбора применяется и их комбинация. Например, можно комбинировать типическую и серийную выборки, когда

серии выбираются в установленном порядке из нескольких типических групп. Возможна также комбинация серийного и собственно-случайного отборов, при которой отдельные единицы отбираются внутри серии в собственно-случайном порядке. Ошибка такой выборки определяется ступенчатостью отбора.

*Многоступенчатой* называется выборка, при которой из генеральной совокупности сначала избираются укрупненные группы, потом более мелкие и так до тех пор, пока не будут отобраны те единицы, которые подвергаются непосредственному обследованию (например, сначала выбирают факультеты, затем группы студентов, а уже внутри групп – студентов).

Многоступенчатый отбор дает, как правило, менее точные результаты по сравнению с одноступенчатым отбором, так как ошибки репрезентативности выборки возникают на каждой ступени.

В отличие от многоступенчатой *многофазная выборка* предполагает сохранение одной и той же единицы отбора на всех этапах его проведения. При этом отобранные на каждой стадии единицы подвергаются обследованию, на каждой последующей стадии отбора программа обследования расширяется. Расчет ошибки многофазной выборки производится для каждой фазы в отдельности.

Следует иметь в виду, что способ формирования выборочной совокупности влияет на выбор формул для оценки ошибок выборочного наблюдения.

### **7.3. Средние и предельные ошибки выборочного наблюдения**

*Ошибкой выборки называют расхождение между характеристиками выборочной совокупности и искомыми параметрами генеральной совокупности.*

Общая величина ошибки выборочного наблюдения формируется из ошибок двоякого рода: *ошибки регистрации* и *ошибки репрезентативности*. Наглядно состав ошибок выборочного наблюдения представляет рис. 7.2.

*Ошибки регистрации* свойственны любому статистическому наблюдению. Среди ошибок регистрации выделяются *систематические*, обусловленные причинами, действующими в каком-то одном направлении и искажающими результаты наблюдения (например, несовершенство измерительных приборов, неправильное понимание инструкций и т.д.), и *случайные*, проявляющиеся в различных направлениях и взаимопогашающие друг друга. Ошибки регистрации рассматривались в теме 2.



Рис. 7.2. Классификация ошибок выборочного наблюдения

*Ошибки репрезентативности* присущи только несплошному наблюдению. Они представляют собой расхождение между величиной полученных выборочных показателей и величиной этих показателей, если бы они были получены при проведении с одинаковой степенью точности сплошного наблюдения.

Ошибки репрезентативности могут быть:

- случайными;
- систематическими.

*Систематические ошибки репрезентативности* возникают в связи с нарушением установленных правил отбора. Оценить систематические ошибки очень сложно, поэтому необходимо знать причины этих ошибок и не допускать их появления. Систематические ошибки репрезентативности приводят к постоянному (тенденциозному) смещению, которое не уменьшается с увеличением объема выборки и вызвано недостатками и просчетами в системе отбора единиц. Основными причинами систематических ошибок репрезентативности являются следующие ситуации:

- подмена объективного отбора удобной выборкой. Например, использование добровольных респондентов из числа родственников, друзей, знакомых;
- неполный охват намеченных к наблюдению единиц или их типов;
- произвольная замена намеченных единиц другими;
- некачественно разработанные методики сбора данных, например, некорректное включение в программу вопросов, слишком болезненных для респондентов (о доходах, возрасте, об отношении к властям и др.) или слишком сложных.

*Случайные ошибки репрезентативности* возникают потому, что в выборочной совокупности недостаточно равномерно представлены раз-

личные типы единиц генеральной совокупности. Вследствие этого распределение отобранной совокупности единиц не вполне точно воспроизводит распределение единиц генеральной совокупности.

*Задачей выборочного метода является измерение случайных ошибок репрезентативности.* Величина случайной ошибки репрезентативности зависит:

- 1) от принятого способа формирования выборочной совокупности;
- 2) от объема выборки;
- 3) от степени вариации изучаемого признака в генеральной совокупности.

Ошибки выборки являются случайными величинами и могут принимать различные значения. Поэтому определяют среднюю из возможных ошибок выборки. В математической статистике доказывается, что *средняя ошибка выборки* для собственно-случайного отбора определяется по формулам:

- для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}}; \quad (7.1)$$

- для доли:

$$\mu_{\omega} = \sqrt{\frac{pq}{n}}, \quad (7.2)$$

где  $\mu_{\bar{x}}$  – средняя ошибка выборки для среднего значения количественного признака;

$\mu_{\omega}$  – средняя ошибка выборки для доли альтернативного признака;

$\sigma^2$  – дисперсия признака  $X$  в генеральной совокупности;

$pq$  – дисперсия доли альтернативного признака в генеральной совокупности.

В этих формулах  $\sigma^2$  и  $pq$  являются характеристиками генеральной совокупности, которые при выборочном наблюдении нам бывают неизвестны.

В математике доказывается, что соотношение между генеральной и выборочной дисперсиями описывает формула

$$\sigma^2 = \sigma_B^2 \frac{n}{n-1}. \quad (7.3)$$

Поскольку величина  $\frac{n}{n-1}$  при  $n \rightarrow \infty$  близка к 1, можно приближенно считать, что выборочная дисперсия равна генеральной дисперсии, т.е.  $\sigma^2 \approx \sigma_B^2$ .



Поэтому на практике в формулах средней ошибки выборки вместо генеральной дисперсии используют дисперсию выборочной совокупности.

Рассмотрим порядок расчета средней ошибки выборки для разных способов формирования выборочной совокупности.

1. *Собственно-случайная выборка.*

*Повторный отбор:*

а) для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_B^2}{n}}; \quad (7.4)$$

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}}. \quad (7.5)$$

*Бесповторный отбор:*

а) для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_B^2}{n} \left(1 - \frac{n}{N}\right)}; \quad (7.6)$$

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n} \left(1 - \frac{n}{N}\right)}. \quad (7.7)$$

2. При *механической выборке* средняя ошибка выборки оценивается по тем же формулам, что и при случайной бесповторной выборке.

3. *Типическая выборка.*

*Повторная типическая выборка:*

а) для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_{B_j}^2}{n}}; \quad (7.8)$$

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n}}. \quad (7.9)$$

*Бесповторная типическая выборка:*

а) для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma_{B_j}^2}{n} \left(1 - \frac{n}{N}\right)}, \quad (7.10)$$

где  $\overline{\sigma_{B_j}^2}$  – средняя из внутригрупповых дисперсий в выборочной совокупности (для количественного признака):

$$\overline{\sigma_{B_j}^2} = \frac{\sum_{j=1}^k \sigma_{B_j}^2 n_j}{\sum_{j=1}^k n_j}, \quad (7.11)$$

$$\sigma_{B_j}^2 = \frac{\sum_{i=1}^{n_j} (x_{ij} - \tilde{x}_j)^2}{n_j}; \quad (7.12)$$

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\omega(1-\omega)}{n} \left(1 - \frac{n}{N}\right)}, \quad (7.13)$$

где  $\overline{\omega(1-\omega)}$  – средняя из внутригрупповых дисперсий для альтернативного признака:

$$\overline{\omega(1-\omega)} = \frac{\sum_{j=1}^k \omega_j(1-\omega_j)n_j}{\sum_{j=1}^k n_j}. \quad (7.14)$$

#### 4. Серийная выборка.

*Серийная выборка при повторном отборе:*

а) для среднего значения признака:

$$\mu_{\tilde{x}} = \sqrt{\frac{\delta_B^2}{r}}, \quad (7.15)$$

где  $r$  – число серий в выборочной совокупности;

$\delta_B^2$  – межсерийная дисперсия выборочной средней:

$$\delta_B^2 = \frac{\sum_{i=1}^r (\tilde{x}_i - \tilde{x})^2}{r} \quad (7.16)$$

где  $\tilde{x}_i$  – средняя  $i$ -той серии;

$\tilde{x}$  – общая средняя по всей выборочной совокупности;

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\delta_{\omega}^2}{r}}, \quad (7.17)$$

где  $\delta_{\omega}^2$  – межсерийная дисперсия выборочной доли:

$$\delta_B^2 = \frac{\sum_{i=1}^r (\omega_i - \omega)^2}{r}, \quad (7.18)$$

где  $\omega_i$  – доля признака в  $i$ -той серии;

$\omega$  – доля признака во всей выборочной совокупности.

*Серийная выборка при бесповторном отборе:*

а) для среднего значения признака:

$$\mu_{\bar{x}} = \sqrt{\frac{\delta_B^2}{r} \left(1 - \frac{r}{R}\right)}; \quad (7.19)$$

б) для доли:

$$\mu_{\omega} = \sqrt{\frac{\delta_{\omega}^2}{r} \left(1 - \frac{r}{R}\right)}, \quad (7.20)$$

где  $R$  – число серий в генеральной совокупности.

Расчет средней ошибки выборки позволяет определить *предельную ошибку выборки*, которая связана с гарантирующей ее вероятностью.

Предельная ошибка выборки рассчитывается по формулам:

а) для среднего значения признака:

$$\Delta_{\bar{x}} = t\mu_{\bar{x}}; \quad (7.21)$$

б) для доли альтернативного признака:

$$\Delta_{\omega} = t\mu_{\omega}, \quad (7.22)$$

где  $t$  – коэффициент доверия, определяемый в зависимости от уровня вероятности, с которой можно гарантировать, что предельная ошибка выборки окажется в заданных пределах.

Значения  $t$  и соответствующие им вероятности определяют по таблицам закона нормального распределения (прил. 2). Чаще всего используют значения доверительной вероятности, приведенные в табл. 7.2.

Так, если  $t = 2$ , то с вероятностью 0,954 можно утверждать, что разность между выборочной и генеральной характеристиками не превысит двух средних ошибок.

Полученные значения  $\mu$  и  $\Delta$  являются точечными оценками интересующих нас характеристик.

Точечные оценки всегда приблизительны, т.к. их получают из отдельных выборок. Поэтому для оценки каждого параметра стремятся установить интервал, который с определенной вероятностью заключает в себе истинное значение параметра.

Фрагмент таблицы закона нормального распределения

$t$	Вероятность $\Phi(t)$
1	0,683
1,96	0,950
2	0,954
3	0,997

Зная выборочную среднюю признака (долю) и предельную ошибку выборки, можно определить пределы, в которых будет заключена генеральная средняя (генеральная доля):

$$\tilde{x} - \Delta_{\tilde{x}} \leq \bar{X} \leq \tilde{x} + \Delta_{\tilde{x}}, \quad (7.23)$$

$$\omega - \Delta_{\omega} \leq p \leq \omega + \Delta_{\omega}. \quad (7.24)$$

Пределы, в которых с данной вероятностью будет заключена характеристика генеральной совокупности, называются *доверительными*, а вероятность – *доверительной*.

Вероятность того, что величина генеральной средней (генеральной доли) выйдет за границы доверительного интервала, будет равна  $[1 - \Phi(t)]$ .

Следует иметь в виду, что оценка средней и предельной ошибок выборки, если ее объем менее 30 единиц (такие выборки называют *безусловно малыми*), имеет свою специфику. Поэтому оценка параметров генеральной совокупности по результатам малой выборки рассматривается отдельно.

#### 7.4. Определение необходимой численности выборки

Разрабатывая программу выборочного наблюдения, задают величину предельной ошибки выборки и доверительную вероятность. Незвестным остается тот минимальный объем выборки, который должен обеспечить требуемую точность. Формулы для определения численности выборки ( $n$ ) выводятся из формул предельной ошибки для разных способов отбора (табл. 7.3).

Значения предельных ошибок выборки  $\Delta_{\tilde{x}}$ ,  $\Delta_{\omega}$  и коэффициент доверия  $t$  определяются как спецификой задач, стоящих перед исследователями, так и природой изучаемого явления. Чем более достоверные результаты требуется получить, тем большую вероятность необходимо задать. С увеличением допустимой ошибки уменьшается необходимый объем выборки, и наоборот.

**Определение необходимого объема выборки  
для некоторых способов формирования выборочной совокупности**

Виды выборочного наблюдения	Повторный отбор	Бесповторный отбор
<i>Собственно-случайная выборка:</i>		
а) при определении среднего размера признака	$n = \frac{t^2 \sigma_B^2}{\Delta_{\bar{x}}^2}$	$n = \frac{t^2 \sigma_B^2 N}{\Delta_{\bar{x}}^2 N + t^2 \sigma_B^2}$
б) при определении доли признака	$n = \frac{t^2 \omega(1-\omega)}{\Delta_{\omega}^2}$	$n = \frac{t^2 \omega(1-\omega)N}{\Delta_{\omega}^2 N + t^2 \omega(1-\omega)}$
<i>Механическая выборка</i>	-	Как для собственно-случайной выборки
<i>Типическая выборка:</i>		
а) при определении среднего размера признака	$n = \frac{t^2 \overline{\sigma_{Bj}^2}}{\Delta_{\bar{x}}^2}$	$n = \frac{t^2 \overline{\sigma_{Bj}^2} N}{\Delta_{\bar{x}}^2 N + t^2 \overline{\sigma_{Bj}^2}}$
б) при определении доли признака	$n = \frac{t^2 \omega(1-\omega)}{\Delta_{\omega}^2}$	$n = \frac{t^2 \omega(1-\omega)N}{\Delta_{\omega}^2 N + t^2 \omega(1-\omega)}$
<i>Серийная выборка:</i>		
а) при определении среднего размера признака	$r = \frac{t^2 \delta_B^2}{\Delta_{\bar{x}}^2}$	$r = \frac{t^2 \delta_B^2 R}{\Delta_{\bar{x}}^2 R + t^2 \delta_B^2}$
б) при определении доли признака	$r = \frac{t^2 \delta_{\omega}^2}{\Delta_{\omega}^2}$	$r = \frac{t^2 \delta_{\omega}^2 R}{\Delta_{\omega}^2 R + t^2 \delta_{\omega}^2}$

Вариация признака существует объективно, независимо от исследователя, и к началу выборочного наблюдения она неизвестна. Приблизительно  $\sigma^2$  определяют следующими способами:

- 1) берут ее значение из предыдущих исследований;
- 2) проводят пробные обследования и по ним определяют дисперсию;
- 3) по правилу «трех сигм» общий размах вариации приблизительно укладывается в интервал, равный шести сигмам ( $R \cong 6\sigma$ , отсюда  $\sigma \cong R/6$ ). Для большей точности  $R$  делят на 5;

4) если хотя бы приблизительно известна средняя величина изучаемого признака, то  $\sigma \cong \bar{X}/3$ ;

5) при изучении альтернативного признака, если нет даже приблизительных сведений о доле единиц, обладающих заданным значением исследуемого признака, берется максимально возможная величина дисперсии, равная 0,25.

## 7.5. Малая выборка

К малым выборкам относят выборки объемом меньше 30 единиц. При оценке *генеральной совокупности* по результатам малой выборки формулы для большой выборки, основанные на *нормальном распределении вероятностей*, дают значительные неточности.

Оценка результатов малой выборки производится с учетом двух особенностей:

- необходимо использовать «исправленную» дисперсию;
- следует применять вместо закона нормального распределения закон распределения Стьюдента (*t*-распределение Стьюдента).

При малой выборке величина дисперсии значительно зависит от  $n$ , поэтому ее определяют с учетом числа степеней свободы, которое в данном случае равно  $(n - 1)$ :

$$\sigma_{мв}^2 = \frac{\sum (x - \tilde{x})^2}{n - 1}, \quad (7.25)$$

где  $\sigma_{мв}^2$  – дисперсия малой выборки;

$(n - 1)$  – число степеней свободы, т.е. количество значений признака, которое могут принимать произвольные значения, не меняющие величины средней.

Тогда средняя ошибка малой выборки может быть рассчитана по формуле

$$\mu_{мв} = \sqrt{\frac{\sigma_{мв}^2}{n}} = \sqrt{\frac{\sum (x - \tilde{x})^2}{n(n - 1)}}. \quad (7.26)$$

Среднюю ошибку малой выборки можно рассчитать и на основе «неисправленной» дисперсии выборки:

$$\mu_{мв} = \sqrt{\frac{\sigma_B^2}{n - 1}}. \quad (7.27)$$

Предельная ошибка малой выборки:

$$\Delta_{мв} = t \cdot \mu_{мв}, \quad (7.28)$$

где  $t$  – коэффициент доверия, определяемый по таблицам распределения Стьюдента, которые устанавливают вероятность, гарантирующую, что, предельная ошибка малой выборки не превзойдет  $t$ -кратную среднюю ошибку (прил. 3).

Стьюдент – псевдоним английского статистика В. Госсета, который разработал теорию малых выборок. В 1908 г. он открыл закон распределения  $t$ , который назвали *распределением Стьюдента*. Особенностью распределения Стьюдента является то, что вероятность ошибки выборки зависит не только от величины  $t$ , но и от объема выборки  $n$ . При увеличении объема выборки распределение Стьюдента приближается к нормальному. При  $n > 100$  таблицы Стьюдента дают те же результаты, что и таблицы интеграла вероятностей Лапласа; при  $30 \leq n \leq 100$  расхождения между ними незначительны; при  $n < 30$  (малые выборки) это расхождение велико.

Малые выборки в социально-экономических исследованиях следует применять с большой осторожностью. Выводы на основе малой выборки действительны лишь, если распределение признака в генеральной совокупности является нормальным или близким к нему. Точность результатов при малой выборке все же ниже, чем при выборке большого объема.

#### **7.6. Распространение данных выборочного наблюдения на генеральную совокупность**

Распространение выборочных оценок на генеральную совокупность состоит в определении характеристик генеральной совокупности на основе характеристик выборочной. Применяются два способа распространения выборочных данных:

- способ прямого пересчета;
- способ поправочных коэффициентов.

При первом способе средние величины и доли, полученные в результате исследования выборочной совокупности, переносятся на генеральную. Если известна численность единиц этой совокупности, то можно найти общий объем признака. Например, если средняя выборочная урожайность зерновых равна 20 ц/га, а предельная ошибка выборки  $\pm 1,5$  ц/га, при известной посевной площади в 20 000 га можно установить ожидаемые пределы валового сбора зерновых (от  $18,5 \times 20\ 000 = 37$  тыс. т до  $21,5 \times 20\ 000 = 43$  тыс. т) с вероятностью, принятой при расчете предельной ошибки.

Второй способ используется для уточнения данных сплошного наблюдения. Так, если выборочное наблюдение показало, что недоучет величины исследуемого явления составил 0,5 %, то эту последнюю величину (поправочный коэффициент) распространяют на результат, полученный при сплошном наблюдении, путем увеличения его на 0,5 %.

## ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ

### Вопросы для самоконтроля и проверки знаний

1. Что такое выборочное наблюдение?
2. Что такое генеральная совокупность и выборочная совокупность?
3. Укажите условия формирования репрезентативной выборки.
4. В чем заключаются достоинства выборочного метода?
5. Укажите основные этапы проведения выборочного наблюдения.
6. Каковы теоретические основы выборочного метода?
7. Какие виды отбора выделяют в зависимости от принятой единицы отбора?
8. Какие методы отбора Вы знаете?
9. В чем сущность повторного метода отбора?
10. Какой метод отбора дает более точные результаты: повторный или бесповторный?
11. В каких случаях целесообразно применение собственно-случайной выборки?
12. Какой способ организации выборки следует использовать, если генеральная совокупность не является однородной?
13. Что такое ошибка выборки?
14. Что представляют собой ошибки репрезентативности?
15. Назовите основные причины систематических ошибок репрезентативности.
16. Что влияет на величину случайной ошибки репрезентативности?
17. Как определяется предельная ошибка выборки для среднего значения и для доли?
18. От чего зависит выбор формул для расчета средней ошибки выборки?
19. Как определяется вероятность той или иной ошибки выборки?
20. Как определяются доверительные интервалы для генеральной средней и генеральной доли?
21. От чего зависит необходимый для исследования объем выборки?
22. Какие существуют способы приближенного определения дисперсии выборки?
23. Какие выборки являются безусловно малыми?
24. Чем отличается распределение ошибок простой случайной выборки при проведении больших и малых выборок?
25. В чем особенность определения ошибок выборки при проведении малых выборок?



26. Укажите основные способы распространения данных выборочного наблюдения на генеральную совокупность.

27. Укажите сферы практического применения выборочного метода.

### Тест 7.1

1. При выборочном наблюдении имеют место ...

- а) только ошибки регистрации;
- б) только ошибки репрезентативности;
- в) ошибки регистрации и репрезентативности;
- г) ошибки отсутствуют.

2. Какая категория шире?

- а) сплошное наблюдение;
- б) выборочное наблюдение.

3. Ошибки репрезентативности могут быть ...

- а) только систематическими;
- б) только случайными;
- в) систематическими и случайными.

4. Ошибки репрезентативности возникают в силу того, что ...

- а) выборочная совокупность не совсем точно воспроизводит свойства генеральной совокупности;
- б) производится регистрация неточных сведений;
- в) при заполнении формуляров возможны описки.

5. Применение теории выборочного метода позволяет измерить ...

- а) случайные ошибки репрезентативности;
- б) систематические ошибки репрезентативности;
- в) ошибки регистрации.

6. При формировании выборочной совокупности соблюдение принципа случайности отбора ...

- а) желательно;
- б) обязательно;
- в) необязательно.

7. При проведении отбора рабочих для обследования причин потерь рабочего времени на предприятии были заведомо исключены рабочие с сокращенным рабочим днем. Результаты выборочного наблюдения будут содержать ...

- а) систематические ошибки регистрации;
- б) систематические ошибки репрезентативности;
- в) случайные ошибки регистрации.

8. Укажите, при каких условиях целесообразно использовать собственно случайную выборку:

- а) совокупность неоднородна;
- б) совокупность однородна;
- в) известен исчерпывающий состав генеральной совокупности;
- г) не известен полный состав генеральной совокупности.

9. Какая выборка может быть проведена только на основе бесповторного отбора?

- а) собственно-случайная;
- б) типическая;
- в) механическая;
- г) серийная.

10. Между величиной ошибки выборки и объемом выборочной совокупности ...

- а) существует прямая зависимость;
- б) имеет место обратная зависимость;
- в) отсутствует какая-либо связь.

## Тест 7.2

1. Если генеральная совокупность неоднородна по величине анализируемого признака и в ней могут быть выделены однородные группы, то используют ...

- а) механическую выборку;
- б) типическую выборку;
- в) серийную выборку;
- г) собственно-случайную выборку.

2. Бесповторный отбор единиц из генеральной совокупности по сравнению с повторным ...

- а) не имеет преимуществ;
- б) дает менее точные результаты;
- в) дает более точные результаты.

3. Увеличение доверительной вероятности ...

- а) уменьшает предельную ошибку выборки;
- б) увеличивает предельную ошибку выборки;
- в) не влияет на предельную ошибку выборки.

4. Какой отбор при прочих равных условиях обеспечивает меньшую численность выборочной совокупности?

- а) повторный;
- б) бесповторный.

5. Средняя ошибка типической выборки при обоснованной типизации генеральной совокупности ...

- а) меньше средней ошибки собственно-случайной выборки;
- б) равна средней ошибке собственно-случайной выборки;
- в) больше средней ошибки собственно-случайной выборки.

6. Способ организации выборочной совокупности влияет на порядок определения ...

- а) средней ошибки выборки;
- б) предельной ошибки выборки.

7. Предельная ошибка выборки при уровне вероятности 0,75 ...

- а) меньше средней ошибки выборки;
- б) больше средней ошибки выборки;
- в) равна средней ошибке выборки.

8. Для каких способов организации выборочной совокупности необходимый объем выборки определяется по одним и тем же формулам?

- а) собственно-случайного бесповторного и механического;
- б) собственно-случайного и типического;
- в) типического и механического;
- г) собственно-случайного и серийного.

9. Для малых выборок коэффициент доверия определяется на основе ...

- а)  $t$ -распределения Стьюдента;
- б) значений функции Лапласа;
- в) распределения Фишера.

10. При серийной выборке в отобранных сериях обследуют ...

- а) все единицы;
- б) отобранные собственно-случайным бесповторным методом;
- в) отобранные механическим способом.

## **Задания и задачи**

### **Задание 7.1**

Объясните, почему следующие методы отбора данных могли бы привести к ошибкам репрезентативности в обследованиях, проводимых с целью:

1) получить представление о взглядах общества на смертную казнь с помощью опроса через журнал «Юрист»;

2) получить представление о реакции общества на закон о подоходном налоге проведением опроса при посещении домов в дневное время.

### **Задание 7.2**

Из приведенных выборочных обследований определите данные, которые содержат систематическую ошибку регистрации, а также данные, имеющие систематическую ошибку репрезентативности, в следующих случаях:

- 1) при изучении производительности труда из совокупности заведомо были исключены рабочие со стажем менее одного года;
- 2) при обследовании состояния животноводства в фермерских хозяйствах из-за недобросовестности счетчиков в некоторых хозяйствах не полностью был учтен молодняк скота;
- 3) при выборочном обследовании бюджета времени работающих, как оказалось в последствии, затраты времени на прачечную и химчистку вместо статьи «Покупка товаров и получение услуг» были включены в статью «Работа на дому».

### **Задача 7.3**

Из партии готовой продукции в порядке механической выборки проверено 400 изделий и установлено, что 80 % из них соответствует первому сорту.

С вероятностью 0,954 определите долю продукции первого сорта во всей партии.

### **Задача 7.4**

В городе проживает 250 тыс. семей. Для определения среднего числа детей в семье была организована 2%-ная случайная бесповторная выборка семей. По ее результатам было получено следующее распределение семей по числу детей:

Число детей в семье, чел.	0	1	2	3	4	5
Количество семей	1 000	2 000	1 200	400	200	200

С вероятностью 0,95 найдите пределы, в которых будет находиться среднее число детей в генеральной совокупности.

### **Задача 7.5**

Городским отделом статистики для изучения общественного мнения о работе РЭУ в порядке механического отбора было опрошено 6 400 человек или 1 % общей численности взрослого городского населения. Из числа опрошенных 3 840 человек положительно оценили работу РЭУ. С вероятностью 0,997 определите пределы, в которых находится доля лиц, положительно оценивающих работу РЭУ.

### Задача 7.6

На предприятии выборочно проверен стаж у 120 мужчин и 80 женщин. Результаты наблюдения представлены в табл. 7.4.

Таблица 7.4

Группы рабочих по полу	Численность работников, чел.	Средний стаж работы, лет	Среднеквадратическое отклонение стажа, лет
Мужчины	120	14	3
Женщины	80	11	2

Определите:

- 1) общий средний стаж работы для рабочих по выборочным данным;
- 2) с вероятностью 0,954 доверительные пределы среднего стажа работы рабочих в генеральной совокупности.

### Задача 7.7

Операция шлифования при обработке детали № 312 производится в цехе на трех станках. Для определения процента брака для всей партии продукции, выработанной за день, проведена типическая 10 %-ная выборка. Отбор деталей из выработки каждого станка – случайный бесповторный; объем выборки пропорционален размеру выпуска. На первом станке было обработано 1 700 деталей, на втором – 2 000, на третьем – 1 800. Число забракованных деталей в выборке: по первому станку – 2, по второму – 3, по третьему – 3.

Определить:

- 1) доверительные интервалы, в которых с вероятностью 0,95 заключен процент брака для всей партии продукции;
- 2) вероятность того, что процент брака для всей продукции отличается от полученного по выборке не более чем на 6 %.

### Задача 7.8

Для изучения текучести кадров на предприятиях города в течение года по специальной программе было выборочно опрошено 400 человек, или 20 % уволившихся по собственному желанию. Результаты наблюдения представлены в табл. 7.5.

Из числа опрошенных 176 человек были неудовлетворены жилищными условиями.

Определите с вероятностью 0,997:

- 1) доверительные интервалы среднего стажа работников, уволившихся по собственному желанию;

2) доверительные интервалы удельного веса работников, уволившихся по причине неудовлетворенности жилищными условиями.

Таблица 7.5

**Группировка уволившихся по стажу работы на предприятии**

Группы уволившихся по стажу работы, лет	Численность уволившихся, чел.
До 3	30
3 – 5	198
5 – 7	22
7 – 9	60
9 и более	20
Итого	400

**Задача 7.9**

В 100 туристических агентствах города предполагается провести обследование среднемесячного количества реализованных путевок методом механического отбора. Какова должна быть численность выборки, чтобы с вероятностью 0,683 предельная ошибка выборки не превышала 3 путевок, если по данным пробного обследования дисперсия составляет 225?

**Задача 7.10**

Какой должна быть необходимая численность выборочной совокупности при механическом отборе, если наблюдение проводится в целях определения доли единиц в генеральной совокупности, обладающих тем или иным значением альтернативного признака, чтобы с вероятностью 0,954 можно было гарантировать ошибку доли не более 3 %. Дисперсия доли неизвестна.

Результат определите для следующих условий:

- 1) объем генеральной совокупности составляет 1 000 единиц;
- 2) объем генеральной совокупности составляет 10 000 единиц;
- 3) объем генеральной совокупности неизвестен.

**Задача 7.11**

На основе тщательного изучения выборки объемом 868, извлеченной из 11 013 хранящихся на складе контейнеров, было установлено, что 3,6 % контейнеров не готовы к отгрузке. Отбор контейнеров проводился на основе механической выборки. Определите предельную ошибку выборки для доли. Доверительную вероятность установите самостоятельно.

Были бы Вы удивлены, узнав, что в действительности 4 % из 11 013 контейнеров не готовы к отгрузке? Почему?

Были бы Вы удивлены, узнав, что в действительности 10 % из 11 013 контейнеров не готовы к отгрузке? Почему?

## Тема 8

# СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ДИНАМИКИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

- 8.1. Сущность, виды и правила построения рядов динамики.
- 8.2. Показатели интенсивности изменения уровней ряда динамики.
- 8.3. Средние показатели ряда динамики.
- 8.4. Характеристика и методы выявления основной тенденции развития явлений.
- 8.5. Понятие и оценка сезонной неравномерности развития.
- 8.6. Методика оценки компонент динамики.

Изучив данную тему, студенты должны:

- иметь ясное представление о рядах динамики, значении и методах их анализа;
- знать правила построения рядов динамики и условия сопоставимости их уровней;
- уметь рассчитывать и интерпретировать показатели интенсивности изменения уровней ряда динамики;
- знать методику расчета средних показателей ряда динамики;
- иметь ясное представление о возможном составе и методах оценки компонент, формирующих уровни ряда динамики;
- знать и уметь применять методы выявления основной тенденции развития;
- уметь использовать результаты анализа динамики для целей прогнозирования;
- понимать значение анализа сезонной неравномерности в управлении организациями.

### **8.1. Сущность, виды и правила построения рядов динамики**

#### **Сущность рядов динамики**

Социально-экономические явления и объекты постоянно изменяются во времени. Для установления тенденций и закономерностей их развития статистика анализирует *ряды динамики*, иначе их называют *динамические*, или *временные, ряды*.

*Ряд динамики – ряд одноименных статистических показателей, расположенных в хронологической последовательности.*

Ряд динамики включает два элемента:

– *время*, к которому относятся приводимые статистические показатели, обычно обозначают *t*;

– *уровни ряда*, т.е. значения показателей, составляющих динамический ряд. Их принято обозначать через *y*.

Например, в табл. 8.1 время в динамическом ряду представлено годами, а уровни рядов динамики – соответствующими показателями работы промышленности Республики Беларусь.

Таблица 8.1

**Динамика основных показателей работы промышленности Республики Беларусь за 2005 – 2010 гг.**

Показатели	2005	2006	2007	2008	2009	2010
Объем продукции в фактически действовавших ценах, млрд руб.	62 545	77 267	95 515	130 830	127 316	161 934
Объем продукции в сопоставимых ценах в % к предыдущему году	110,5	111,4	108,7	111,5	98,0	111,3
Среднесписочная численность промышленно-производственного персонала, тыс. чел.	1 062	1 068	1 084	1104	1 068	1 059
Рентабельность реализованной продукции, работ, услуг, процентов	15,4	15,5	13,0	15,3	9,9	9,9

Источник: [11]

### **Виды рядов динамики**

Виды рядов динамики выделяют на основе их классификации по ряду признаков.

1. **В зависимости от вида статистических показателей, составляющих ряд динамики**, выделяют:

– *ряды динамики, представленные абсолютными величинами*. Например, объем валового внутреннего продукта за ряд лет, объем инвестиций за ряд периодов, численность населения страны (области, города) по состоянию на начало года за ряд лет и др.;

– *ряды динамики, представленные относительными величинами*. Примером динамических рядов этого вида являются ряды, представляющие индексы физического объема валового внутреннего продукта, индексы потребительских цен и др.;

– *ряды динамики, представленные средними величинами*. Например, ряды, содержащие показатели средней начисленной заработной платы, среднедушевые денежные доходы населения страны и др.



2. В зависимости от того, как представлено время в ряду динамики, выделяют:

- интервальные ряды;
- моментные ряды.

Интервальным называется ряд динамики, в котором уровни характеризуют результат, накопленный или полученный за определенный интервал времени (период). К интервальным относят ряды динамики, представляющие, например, объемы производства продукции, ВВП, показатели отработанного времени, объема инвестиций, уровня производительности труда, фондоотдачи, фондовооруженности, показатели рождаемости, смертности, естественного прироста (убыли) населения, миграции и др.

Моментным называется ряд динамики, в котором уровни характеризуют состояние явления на определенную дату или момент времени. Примерами моментных рядов является последовательность показателей численности населения на начало года, размера запаса какого-либо материала на начало периода. На определенный момент времени (начало и конец периода) приводятся данные в бухгалтерском балансе предприятия.

Обратите внимание, важное аналитическое отличие моментных рядов от интервальных состоит в том, что уровни моментных рядов нельзя суммировать, даже если они представлены абсолютными величинами, т.к. одна и та же величина будет учтена неоднократно. Разность уровней моментного ряда покажет абсолютное изменение показателя за определенный период.

Особенностью интервальных рядов динамики, представленных абсолютными величинами, является возможность суммирования их уровней. Примером интервальных рядов динамики являются ряды, представленные в табл. 8.1. Моментный ряд динамики представлен в табл. 8.2.

Таблица 8.2

**Списочная численность работников предприятия на конец года  
за 2003 – 2010 гг.**

	2003	2004	2005	2006	2007	2008	2009	2010
Списочная численность работников предприятия на конец года, чел.	4 050	4 105	4 080	4 100	4 120	4 125	4 140	4 140

**3. В зависимости от расстояния между уровнями рядов динамики** выделяют:

– *ряды с равноотстоящими уровнями*. Иначе их называют полными рядами. В таких рядах даты или периоды указываются через равный интервал времени (см. табл. 8.1, 8.2);

– *ряды с неравноотстоящими уровнями* или неполные ряды. В таких рядах расстояние между датами или периодами не равны (см. табл. 8.4).

**4. В зависимости от наличия основной тенденции развития изучаемого явления ряды динамики** подразделяют:

– на *стационарные*. Это ряды, в которых уровни с течением времени остаются неизменными;

– *нестационарные ряды*. Экономические процессы во времени обычно изменяются, т.е. имеют основную тенденцию развития и являются нестационарными.

**5. По числу показателей** можно выделить:

– *изолированные ряды динамики*. Если проводится анализ изменения во времени одного показателя, то ряд динамики является изолированным;

– *комплексные (многомерные) ряды*. В многомерном ряду представлено несколько взаимосвязанных показателей.

Наглядное представление о видах рядов динамики дает рис. 8.1.

### **Сопоставимость уровней и смыкание рядов динамики**

Важным условием качественного анализа рядов динамики является обеспечение сопоставимости их уровней. Соблюдение данного условия обеспечивается в процессе сбора и обработки данных, а также путем проведения специальных расчетов, обеспечивающих сопоставимость уровней.

Укажем основные условия сопоставимости уровней динамического ряда:

1) должна быть обеспечена сопоставимость уровней по кругу охватываемых объектов явления. Т.е. уровни динамического ряда за отдельные периоды или моменты времени должны характеризовать размер анализируемого явления по одному и тому же кругу входящих в него объектов. Например, если анализ динамики численности студентов предполагает анализ всего их состава, то нельзя в один год учитывать заочников, а в другой нет;

2) уровни динамического ряда должны быть рассчитаны по единой методологии (например, производительность труда на одного рабочего или работающего). Единство методологии расчета важно для средних и относительных показателей, составляющих ряд динамики;

3) для измерения уровней ряда динамики необходимо использовать одинаковые единицы измерения;



Рис. 8.1. Виды рядов динамики  
Источник: [4, с. 475]

4) сопоставимость уровней интервального ряда требует равенства периодов, за которые приведены данные; а для моментного ряда с сезонным характером изменения уровней должна быть обеспечена сопоставимость по критическому моменту регистрации;

5) уровни ряда должны быть сопоставимы по территории. В анализе динамики приходится сталкиваться с изменением территориальных границ областей, районов и т.п. Территориальная сопоставимость обеспечивается смыканием рядов динамики;

б) условием сопоставимости уровней временного ряда является периодизация динамики. Научный подход к изучению рядов динамики заключается в том, чтобы ряды, охватывающие большие периоды времени, расчленять на такие, которые бы объединяли лишь однокачественные пе-

риоды развития совокупности, характеризующиеся одной закономерностью развития.

Процесс выделения однородных этапов развития рядов динамики носит название *периодизации динамики*. Вопрос о выделении однородных этапов развития решается теорией той науки, к области которой относится изучаемая совокупность. Необходимость формирования рядов динамики по строго однородным периодам или этапам не означает отрицания возможности построения и изучения рядов динамики за длительные исторические периоды. Само понятие однородности относительно и зависит от уровня абстракции, принятой в исследовании.

Таким образом, прежде чем анализировать динамический ряд, надо, исходя из цели исследования, убедиться в сопоставимости его уровней. Если какие-то условия сопоставимости уровней ряда не соблюдаются, их приводят к сопоставимому виду.

Одним из основных приемов приведения уровней ряда к сопоставимому виду является *смыкание рядов динамики*.

*Под смыканием понимают объединение в один ряд (более длинный) двух или нескольких рядов динамики, уровни которых исчислены по разной методологии или разным территориальным границам.* Для проведения смыкания необходимо, чтобы для одного из периодов (переходного) имелись данные, исчисленные по разной методологии (или в разных границах).

Рассмотрим методику смыкания рядов динамики на следующем примере.

### Пример 8.1

По одному из районов области, территория которого изменилась в 2007 г., имеются данные о розничном товарообороте (табл. 8.3).

Таблица 8.3

**Динамика розничного товарооборота за 2005 – 2010 гг.**

Показатели	2005	2006	2007	2008	2009	2010
Розничный товарооборот, млрд руб.						
– в старых границах	1 400,0	1 700,0	1 900,0	–	–	–
– в новых границах	–	–	2 070,0	2 130,0	2 250,0	2 340,0
Сомкнутый (сопоставимый) ряд абсолютных величин, млрд руб.	1 524,6	1 851,3	2 070,0	2 130,0	2 250,0	2 340,0
Сопоставимый ряд, представленный относительными величинами, % к 2007 г.	80,2	89,5	100	102,9	108,7	113,0

Имея представленные в табл. 8.3 исходные данные, нельзя провести непосредственное сопоставление данных за 2008, 2009, 2010 гг. с данными за 2005 и 2006 гг., т.к. эти показатели относятся к разным территориальным границам. Для анализа динамики за весь период необходимо сомкнуть (объединить) исходные два ряда в один. Смыкание этих рядов возможно, т.к. данные за 2007 г. приведены в прежних и новых территориальных границах. Используя данные за 2007 г., рассчитаем коэффициент пересчета показателей, относящихся к старым территориальным границам в показатели, соответствующие новым территориальным границам:

$$K_{пер.} = \frac{\text{Показатель 2007 г., относящийся к новым территориальным границам}}{\text{Показатель 2007 г., относящийся к старым территориальным границам}} =$$

$$= \frac{2\,070}{1\,900} = 1,089.$$

Далее умножаем на полученный коэффициент данные за 2005 – 2006 гг. и приводим их, таким образом, в сопоставимый вид с данными за 2008 – 2010 гг. Уровни ряда в 2005 и 2006 гг. в новых границах составят:

$$y_{2005} = 1\,400 \cdot 1,089 = 1\,524,6 \text{ (млрд руб.)};$$

$$y_{2006} = 1\,700 \cdot 1,089 = 1\,851,3 \text{ (млрд руб.)}.$$

Сомкнутый (сопоставимый) ряд динамики показан в предпоследней строке табл. 8.3.

Другой способ смыкания рядов динамики состоит в том, что уровни года, в котором произошли изменения (в нашем примере уровни 2007 г., как до, так и после изменений, т.е. 1900 и 2070), принимаются за 100 %, а остальные уровни пересчитываются в процентах по отношению к этим уровням соответственно (в старых границах – по отношению к 1900, а в новых – 2070). В результате получаем сомкнутый ряд динамики, который представлен относительными величинами и показан в последней строке табл. 8.3.

Проблема приведения к сопоставимому виду возникает и при параллельном сравнении изменения во времени показателей, характеризующих развитие отдельных стран, регионов, хозяйственных комплексов и т.п. В таких случаях сравниваемые ряды динамики приводят к *одному основанию*, т.е. к одному и тому же периоду или моменту времени, который принимается за базу сравнения, а все остальные уровни выражают в виде коэффициентов или в темпах роста по отношению к нему.

## Пример 8.2

Имеются данные о динамике экспорта товаров по отдельным регионам (табл. 8.4).

Таблица 8.4

**Динамика экспорта товаров по регионам за 2005 – 2010 гг.**  
(млн долл. США)

Регион	2005	2007	2009	2010
Регион А	377,0	674,7	985,8	1 133,3
Регион Б	601,9	1 111,2	2 008,3	2 373,4
Регион В	1 084,5	1 659,5	3 437,2	4 060,2

Различные значения абсолютных уровней приведенных рядов динамики, затрудняют выявление особенностей роста объема экспорта в трех областях. Поэтому приведем абсолютные уровни рядов динамики к *общему основанию*, приняв за постоянную базу сравнения уровни 2005 г. Результаты расчета представим в табл. 8.5.

Таблица 8.5

**Динамика экспорта товаров по регионам за 2005 – 2010 гг.,**  
% к 2005 г.

Регион	2005	2007	2009	2010
Регион А	100,0	179,0	261,5	300,6
Регион Б	100,0	184,5	333,7	394,3
Регион В	100,0	153,0	316,9	374,4

В относительных величинах, представленных базисными темпами роста по каждой области, более наглядно видны различия в динамике экспорта товаров. В частности, данные табл. 8.5 показывают, что более высокими темпами увеличивался экспорт товаров в регионе Б.

## 8.2. Показатели интенсивности изменения уровней ряда динамики

### Состав и порядок расчета показателей интенсивности изменения уровней ряда динамики

Одним из важных направлений изучения динамики является оценка *интенсивности изменения уровней временного ряда*. Такую оценку проводят с помощью специальных статистических показателей, которые иначе называют *аналитическими показателями ряда динамики*.

Расчет показателей интенсивности изменения уровней ряда динамики позволяет установить направление и характер изменения явлений во

времени, в частности оценить меру снижения или увеличения уровней ряда, определить, является ли развитие равномерным или неравномерным, какова величина и характер ускорения.

К показателям интенсивности изменения уровней ряда динамики (аналитические показатели ряда динамики) относят:

- абсолютный прирост;
- коэффициент роста (темп роста);
- темп прироста;
- абсолютное значение 1 % прироста.

Расчет этих показателей основан на абсолютном или относительном сравнении уровней ряда динамики. Сравнимый уровень ряда динамики называется *текущим* (анализируемым) и обозначается в общем случае  $y_i$ , а уровень, с которым производится сравнение, называется *базисным*, или *базой сравнения* (обозначается  $y_0$ ,  $y_1$  или  $y_{i-1}$  в зависимости от принятой схемы расчета). При сравнении более двух последовательных уровней динамического ряда возможны две схемы их сопоставления:

- сравнение с *постоянной базой*, когда каждый уровень динамического ряда сравнивается с одним и тем же уровнем, принятым за базу. Обычно в качестве базы выбирают начальный уровень динамического ряда;
- сравнение с *переменной базой*, когда каждый уровень динамического ряда сравнивается с предшествующим ему уровнем.

*Показатели динамики с постоянной базой*, так называемые базисные показатели, характеризуют *окончательный* результат всех изменений в уровнях ряда от периода, к которому относится базисный уровень, до анализируемого  $i$ -того уровня.

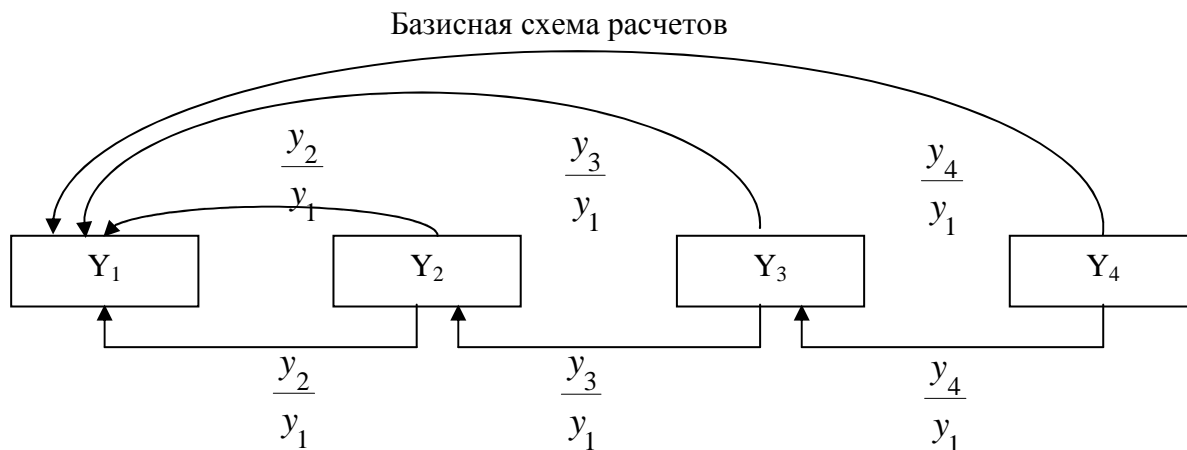
*Показатели динамики с переменной базой* (цепные показатели) характеризуют интенсивность изменения уровней от периода к периоду (или от даты к дате) в пределах изучаемого временного интервала.

Наглядно принципы построения цепных и базисных показателей динамики представлены на рис. 8.2 (на рисунке приведен вариант относительного сопоставления уровней).

Рассмотрим методику расчета показателей интенсивности изменения уровней ряда динамики.

### **Абсолютный прирост**

*Абсолютный прирост* ( $\Delta y$ ) показывает, на сколько данный уровень ряда изменился по сравнению с уровнем, принятым за базу сравнения. Определяется абсолютный прирост как разность между двумя уровнями ряда, т.е. в основе расчета этого показателя лежит абсолютное сравнение уровней.



Цепная схема расчетов (при этом  $\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot \frac{y_4}{y_3} = \frac{y_4}{y_1}$ )

Рис. 8.2. Система цепных и базисных показателей динамики  
Источник: [3, с. 232, рис. 8.1]

В зависимости от принятой схемы расчета различают цепной ( $\Delta y^u$ ) и базисный ( $\Delta y^b$ ) абсолютные приросты:

$$\Delta y^u = y_i - y_{i-1}, \tag{8.1}$$

$$\Delta y^b = y_i - y_0, \tag{8.2}$$

где  $y_i$  –  $i$ -тый уровень ряда динамики;  
 $y_{i-1}$  – уровень, предшествующий  $i$ -тому уровню;  
 $y_0$  – начальный уровень ряда.

Абсолютный прирост может быть положительной или отрицательной величиной. Знак «минус» при оценке абсолютного прироста указывает на абсолютное снижение анализируемого уровня по сравнению с уровнем, принятым за базу. Абсолютный прирост имеет ту же единицу измерения, что и уровни ряда динамики. При этом следует указывать период, за который оценивается абсолютный прирост.

Цепные абсолютные приросты (абсолютные прирост с переменной базой) называют *скоростью роста*. Между базисными и цепными абсолютными приростами существует определенная связь. Сумма цепных абсолютных приростов за определенный период дает общий, то есть базисный абсолютный прирост за этот же период:

$$\sum_{i=1}^{n-1} \Delta y_i^u = y_n - y_0, \tag{8.3}$$



где  $n$  – число уровней ряда динамики;

$y_0, y_n$  – начальный и последний уровни ряда динамики;

$\Delta y_i^y$  – цепной абсолютный прирост в  $i$ -том периоде (в  $i$ -тый момент времени).

Если цепные абсолютные приросты приблизительно равны между собой, делают вывод о *равномерном развитии* явления и тогда основная тенденция ряда динамики может быть выражена с помощью уравнения прямой.

Если цепные абсолютные приросты изменяются от периода к периоду, то говорят о *неравномерном развитии*. Определяя разности между цепными абсолютными приростами по цепной схеме, получают абсолютные показатели ускорения (иначе их называют *вторые разности*):

$$\Delta_i = \Delta y_i^y - \Delta y_{i-1}^y, \quad (8.4)$$

где  $\Delta_i$  – абсолютное ускорение для  $i$ -того периода (момента времени);

$\Delta y_i^y, \Delta y_{i-1}^y$  – цепные абсолютные приросты уровней для  $i$ -того и  $(i - 1)$  периодов (моментов времени).

Показатель абсолютного ускорения применяется только в цепном варианте.

Положительная величина ускорения говорит об ускоренном росте (развитии), отрицательная величина ускорения свидетельствует о замедлении скорости роста. При равенстве вторых разностей основная тенденция развития может быть описана с помощью уравнения параболы второй степени.

### **Коэффициент роста (темп роста)**

Если производить относительное сравнение уровней ряда динамики, т.е. находить их отношение, то получают показатель, который называют *коэффициентом роста* или *темпом роста*. Название показателя зависит от того, выражается ли он в виде коэффициента или в процентах.

*Коэффициент роста* показывает, во сколько раз данный уровень ряда больше или меньше уровня, принятого за базу сравнения. В зависимости от принятой схемы расчета различают цепные ( $K_p^y$ ) и базисные ( $K_p^b$ ) коэффициенты роста:

$$K_p^y = \frac{y_i}{y_{i-1}}, \quad (8.5)$$

$$K_p^b = \frac{y_i}{y_0}. \quad (8.6)$$

Принятые в данном случае условные обозначения соответствуют принятым в формулах (8.1) и (8.2).

Темп роста ( $T_p$ ) выражают в процентах и получают умножением коэффициента роста на 100 %:

$$T_p^u = \frac{y_i}{y_{i-1}} \cdot 100 \% = K_p^u \cdot 100 \% , \quad (8.7)$$

$$T_p^{\sigma} = \frac{y_i}{y_0} \cdot 100 \% = K_p^{\sigma} \cdot 100 \% . \quad (8.8)$$

Темп роста показывает, сколько процентов составляет сравниваемый уровень по отношению к уровню, принятому за базу.

Величина коэффициента роста, большая единицы, указывает на увеличение уровня анализируемого периода по сравнению с базисным; равная единице – на то, что уровень анализируемого периода не изменился по сравнению с базисным; меньшая единицы – снизился. При этом следует иметь в виду, что коэффициент роста (темп роста) всегда величина положительная.

При наличии цепных коэффициентов роста базисный коэффициент роста находится по формуле

$$K_{P_{n/0}}^{\sigma} = K_{1/0} \cdot K_{2/1} \cdot \dots \cdot K_{n-1/n-2} \cdot K_{n/n-1} . \quad (8.9)$$

Если уровни ряда динамики принимают положительные и отрицательные значения, например, финансовым результатом от реализации продукции может быть прибыль (+), а может быть убыток (–), тогда темп роста (равно как и темп прироста) применять нельзя. В этом случае показатели теряют смысл и не имеют экономической интерпретации.

### Темп прироста

Темп прироста характеризует абсолютный прирост в относительных величинах и показывает, на сколько процентов уровень данного периода больше (меньше) уровня, принятого за базу:

$$T_{np}^{\sigma} = \frac{y_i - y_1}{y_0} \cdot 100 \% , \quad (8.10)$$

$$T_{np}^u = \frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100 \% \quad (8.11)$$

$$\text{или } T_{np} = T_p - 100 \% . \quad (8.12)$$

При темпах роста меньше 100 % (снижение уровней ряда) имеем отрицательные темпы прироста.

Сумма цепных темпов прироста *не равна* базисному темпу прироста. Дело в том, что значения цепных темпов прироста, рассчитанные каждый к своей базе, различаются не только числом процентов, но и величиной абсолютного изменения, составляющей каждый процент. Поэтому складывать цепные темпы прироста нельзя. Не имеет смысла и перемножение темпов прироста.

### **Абсолютное значение 1 % прироста**

При анализе относительных показателей динамики (темпов роста и темпов прироста) не следует рассматривать их изолированно от абсолютных показателей (уровней ряда и абсолютных приростов). Сравнение абсолютного прироста и темпа прироста за одни и те же периоды показывает, что замедление темпов прироста не всегда сопровождается уменьшением абсолютных приростов. Поэтому, чтобы правильно оценить значение полученного темпа прироста, его рассматривают в сопоставлении с показателем абсолютного прироста. В результате получают показатель, который называют *абсолютным значением одного процента прироста* ( $A_i$ ):

$$A_i = \frac{y_i - y_{i-1}}{T_{np}^u (\%)} = \frac{y_i - y_{i-1}}{\frac{y_i - y_{i-1}}{y_{i-1}} \cdot 100 \%} = 0,01 y_{i-1}. \quad (8.13)$$

### **Коэффициент опережения**

В статистической практике применяют еще некоторые аналитические показатели рядов динамики, например, *коэффициент опережения*.

*Коэффициент опережения*  $K_{on}$  определяется как отношение темпов роста за одинаковые периоды по двум динамическим рядам:

$$K_{on} = \frac{T_p^A}{T_p^B}, \quad (14)$$

где  $T_p^A$ ,  $T_p^B$  – темпы роста уровней ряда  $A$  и  $B$  соответственно за один и тот же период.

С помощью коэффициентов опережения могут сравниваться динамические ряды одинакового содержания, но относящиеся к разным объектам (странам, регионам, отраслям, предприятиям, видам продукции и т.д.), или ряды разного содержания, характеризующие один и тот же объект по

взаимосвязанным показателям. Сопоставление темпов роста двух взаимосвязанных показателей позволяет аналитику получить следующие результаты:

– во-первых, коэффициент опережения может оказаться темпом роста третьего показателя, получаемого от деления двух первых. Например, при делении темпа роста объема промышленной продукции на темп роста среднесписочной численности промышленно-производственного персонала (ППП) получают коэффициент опережения, являющийся темпом роста производительности труда ППП. Или другой пример, при сопоставлении темпов роста номинальной средней заработной платы и темпов роста потребительских цен (индекса потребительских цен) получают коэффициент опережения, представляющий собой темп роста реальной заработной платы.

– во-вторых, расчет коэффициентов опережения позволяет делать выводы о соблюдении тех или иных закономерностей, присущих экономическим процессам. Например, коэффициент опережения темпов роста производительности труда над темпами роста средней заработной платы позволит оценить эффективность организации оплаты труда работников.

Коэффициенты опережения могут исчисляться не только по темпам роста, но и на основе темпов прироста. При этом если сравниваемые динамические ряды находятся в причинно-следственной связи, коэффициенты опережения могут быть интерпретированы как коэффициенты эластичности, характеризующие чувствительность следствия к изменению причин.

### **Различные типы динамики**

Изменение уровней ряда динамики происходит с разной интенсивностью, и характер динамики может принимать разные черты. Однако можно указать основные типы динамики [12, с. 331 – 333], которые встречаются чаще других:

- цепные абсолютные приросты падают (растут) (I тип);
- цепные абсолютные приросты стабильны (II тип);
- цепные темпы прироста стабильны (III тип);
- цепные темпы прироста увеличиваются (уменьшаются) (IV тип).

Возьмем для анализа четыре условных ряда динамики с одинаковым исходным уровнем, который примем равным 200-м каким-либо единицам в 2005 г. (допустим, это 200 тыс. шт. за год какого-либо изделия). Затем для каждого типа динамики исчислим цепные абсолютные приросты и цепные темпы прироста. Результаты вычислений представим в табл. 8.6.

Таблица 8.6

## Типы динамики

Годы	Типы динамики											
	I – падающие цепные абсолютные приросты			II – стабильные цепные абсолютные приросты			III – стабильные цепные темпы прироста			IV – растущие цепные темпы прироста		
	уровень показателя, тыс. шт.	цепной абсолютный прирост, тыс. шт.	цепной темп прироста, %	уровень показателя, тыс.шт.	цепной абсолютный прирост, тыс.шт.	цепной темп прироста, %	уровень показателя, тыс.шт.	цепной абсолютный прирост, тыс.шт.	цепной темп прироста, %	уровень показателя, тыс.шт.	цепной абсолютный прирост, тыс.шт.	цепной темп прироста, %
1	2	3	4	5	6	7	8	9	10	11	12	13
2005	200	–	–	200	–	–	200	–	–	200	–	–
2006	220	20	10,0	220	20	10,0	220	20	10	220	20	10
2007	238	18	8,1	240	20	9,1	242	22	10	246	26	12
2008	254	16	6,7	260	20	8,3	266	24	10	280	34	14
2009	268	14	5,5	280	20	7,7	293	27	10	325	45	16
2010	280	12	4,4	300	20	7,1	322	29	10	384	59	18

Источник: составлено автором на основе [12, с. 331]

Во всех приведенных вариантах уровень непрерывно возрастает, но характер роста, его интенсивность различны.

*I тип динамики* характеризуется ростом уровня при непрерывно понижающемся цепном абсолютном приросте.

*II тип* показывает одинаковые цепные абсолютные приросты на протяжении всего периода. Это означает, что каждый год продукция увеличивается на одну и ту же абсолютную величину. При наличии такой динамики цепной темп роста и темп прироста будут падать. Прирост в 20 тыс. шт. по сравнению с 200 тыс. шт. будет составлять 10 %, но эти же 20 тыс. шт., отнесенные к 280 тыс. шт. (предпоследний член динамического ряда, графа 5), составят уже всего 7,1 %. Таким образом, возрастание уровня при наличии стабильных абсолютных приростов означает некоторое уменьшение цепных темпов прироста («затухающие темпы»). Графически наличие динамики этого типа выражается в виде идущей вверх прямой линии.

*III тип* дает одинаковые цепные темпы роста и прироста. При этом цепные абсолютные приросты неуклонно возрастают.

*IV тип* выражает наибольшую интенсивность роста уровня. В этом случае цепные темпы роста не только не затухают, а систематически возрастают. В нашем примере цепные темпы прироста увеличивались по законам арифметической прогрессии с разностью, равной двум. Разумеется, характер возрастания темпов может быть и совсем иным.

Графически все рассмотренные типы динамики представлены на рис. 8.3.

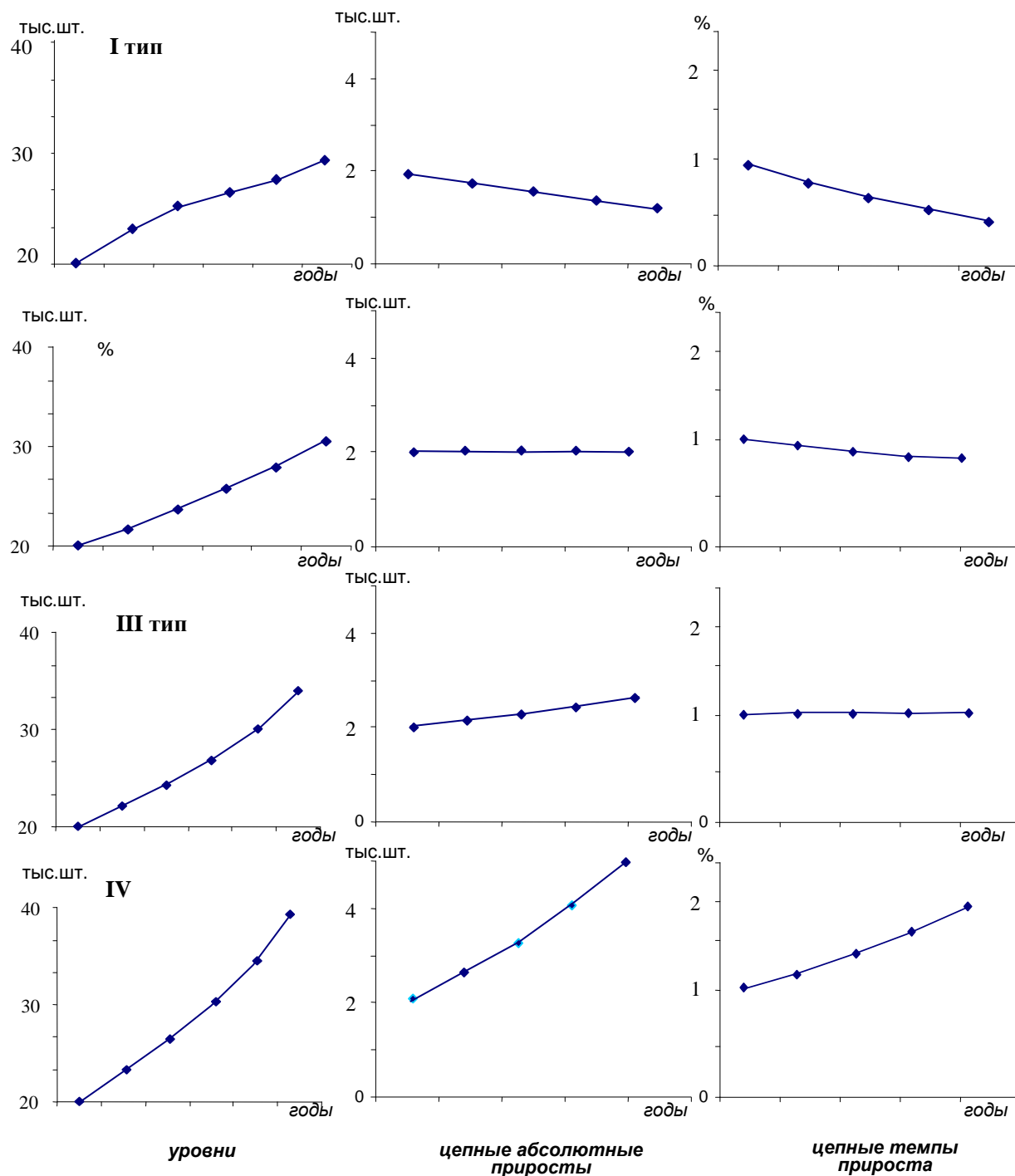


Рис. 8.3. Типы динамики  
Источник: [12, с. 332, рис. 61]

Приведенный анализ типов динамики наглядно показывает, что совместный анализ абсолютных и относительных показателей динамики дает более полную картину развития явлений во времени.

### Пример 8.3

Рассмотрим на конкретном примере результаты расчета показателей интенсивности изменения уровней ряда динамики.

Таблица 8.7

**Динамика объема производства продукции на предприятии**

Год	Объем производства продукции $y_i$ , тыс. шт.	Абсолютный прирост, тыс. шт.		Абсолютный показатель ускорения, тыс. шт.	Темп роста, %		Темп прироста, %		Абсолютное значение 1 % прироста, тыс. шт.
		цепной	базисный		цепной	базисный	цепной	базисный	
2004	100	–	–	–	–	100,0	–	–	–
2005	112	12	12	–	112,0	112,0	12,0	12,0	1,00
2006	128	16	28	4	114,3	128,0	14,3	28,0	1,12
2007	148	20	48	4	115,6	148,0	15,6	48,0	1,28
2008	172	24	72	4	116,2	172,0	16,2	72,0	1,48
2009	200	28	100	4	116,3	200,0	16,3	100,0	1,72
2010	232	32	132	4	116,0	232,0	16,0	132,0	2,0

Анализ представленных в табл. 8.7 данных позволяет сделать вывод о том, что в анализируемом периоде объем производства продукции имел положительную динамику. При этом отмечался равноускоренный рост показателя, о чем свидетельствует равенство цепных абсолютных ускорений (вторых разностей).

### 8.3. Средние показатели ряда динамики

Средние показатели ряда динамики позволяют:

- во-первых, обобщить характеристики динамики за длительный период;
- во-вторых, сравнивать развитие явлений за неодинаковые по длительности периоды.

Для обобщающей характеристики динамики исследуемого явления определяют две группы средних величин:

- *средние уровни ряда* (их называют динамическими, или хронологическими);
- *средние интенсивности развития*.

### **Расчет хронологических средних**

*Средней хронологической (динамической) называется средняя, исчисленная из значений, изменяющихся во времени.* Такие средние обобщают хронологическую вариацию. В хронологической средней отражается совокупное проявление тех условий, в которых развивалось изучаемое явление в анализируемом периоде.

Хронологические средние, как и обычные средние, должны отражать типичный для данного периода уровень явления. Согласно теории средних величин, их вычисление должно вестись по однородным совокупностям. Для развивающихся во времени явлений это означает, что динамические средние должны относиться к периоду с одинаковыми условиями развития.

Обычные средние относятся к одному времени, а хронологические средние обобщают изучаемое явление за разные периоды (моменты времени).

Метод расчета среднего уровня ряда динамики зависит от вида динамического ряда: моментный или интервальный, полный или неполный.

### **Средний уровень интервального ряда**

*Средний уровень интервального ряда с равными интервалами рассчитывается по формуле*

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}, \quad (8.15)$$

где  $n$  – число уровней ряда (периодов, за которые приводятся данные).

Например, средний уровень ряда динамики, представленного в таблице 8.7, следует определять по вышеприведенной формуле, т.к. ряд является интервальным и полным:

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{100 + 112 + 128 + 148 + 172 + 200 + 232}{7} = 156 \text{ (тыс. шт.)}.$$



Если интервалы *не равны*, то средний уровень интервального ряда рассчитывается по формуле

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n t_i}, \quad (8.16)$$

где  $t_i$  – длительность  $i$ -того интервала между уровнями (в месяцах, годах и др.).

### Средний уровень моментного ряда

Средний уровень моментного ряда определяется иначе, т.к. уровни ряда здесь относятся не к определенным периодам, а к моментам времени, не связанным с длительностью периода.

В *моментном ряду с равноотстоящими датами* средний уровень рассчитывают по формуле

$$\bar{y} = \frac{1/2 y_1 + y_2 + \dots + y_{n-1} + 1/2 y_n}{n - 1}. \quad (8.17)$$

### Пример 8.4

Покажем пример расчета среднего уровня моментного ряда динамики с равноотстоящими уровнями.

*Например*, если известны товарные остатки магазина на 1-е число каждого месяца (в млн руб.) за первый квартал:

1.01.2011	1.02.2011	1.03.2011	1.04.2011
180	150	170	210,

то среднемесячный товарный остаток в первом квартале в соответствии с формулой (8.17) составит:

$$\bar{y} = \frac{180/2 + 150 + 170 + 210/2}{4 - 1} = 171,7 \text{ (млн руб.)}.$$

Этот же результат можно получить иначе. При вычислении среднего уровня моментного ряда условно предполагается непрерывное, равномерное изменение уровня в промежутках между двумя датами. Основываясь на этом предположении, определим средние товарные остатки магазина за каждый месяц как полусумму остатков на начало и конец месяца. Средние товарные остатки составят:

в январе:  $\frac{180+150}{2} = 165$  (млн руб.);

в феврале:  $\frac{150+170}{2} = 160$  (млн руб.);

в марте:  $\frac{170+210}{2} = 190$  (млн руб.).

По существу, выполнив эти расчеты, мы получим интервальный ряд динамики, представленный средними величинами, и тогда расчет среднего уровня ряда динамики может быть выполнен по формуле (8.15):

$$\bar{y} = \frac{165+160+190}{3} = 171,7 \text{ (млн руб.)}$$

В моментном ряду с неравноотстоящими датами средний уровень определяется по формулам:

$$\bar{y} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i}, \quad (8.18)$$

где  $y_i$  – уровни ряда динамики, сохранявшиеся без изменения в течение промежутка времени  $t_i$ ,

или

$$\bar{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \dots + (y_{n-1} + y_n)t_{n-1}}{2(t_1 + t_2 + \dots + t_{n-1})} \bar{y}, \quad (8.19)$$

где  $t_1$  – время (в днях, месяцах) между моментами регистрации  $y_1$  и  $y_2$ ;

$t_2$  – время между моментом регистрации  $y_2$  и  $y_3$  и т.д.

Формула (8.19) предполагает, что в промежутках между датами уровни принимали разные значения и мы из двух известных уровней ( $y_i$  и  $y_{i+1}$ ) определяем средние, из которых затем уже рассчитываем общую среднюю.

### Пример 8.5

Рассмотрим пример расчета среднего уровня в моментном ряду.

Известны следующие данные о наличии движения средств по вкладу в первом квартале 2011 г. (в тыс. руб.):

Наличие на 1.01.2011	Дата и размер изменения вклада			
500	05.01.2011 +150	17.01.2011 -200	21.02.2011 +500	28.03.2011 +100

Необходимо определить средние остатки по вкладу в первом квартале.

Так как в данном случае мы имеем моментный ряд с неравноотстоящими датами, между которыми уровни ряда оставались неизменными (изменение вклада происходило только в указанные даты), расчет среднего уровня производим по формуле (8.18):

$$\bar{y} = \frac{\sum_{i=1}^n y_i t_i}{\sum_{i=1}^n t_i}.$$

Для лучшего понимания методики расчета промежуточные данные представим в табл. 8.8.

Таблица 8.8

**Промежуточные расчеты для определения среднего остатка по вкладу**

Периоды	Число дней в периоде $t_i$	Размер вклада $y_i$ , тыс. руб.	$y_i \cdot t_i$
01.01 – 05.01	4	500	2 000
05.01 – 17.01	12	650	7 800
17.01 – 21.02	35	450	15 750
21.02 – 28.03	35	950	33 250
28.03 – 01.04	4	1 050	4 200
Итого	90	–	63 000

Средний остаток по вкладу составит:

$$\bar{y} = \frac{63\,000}{90} = 700 \text{ (тыс. руб.)}.$$

**Пример 8.6**

Известна списочная численность персонала организации на определенные даты 2010 г. (человек):

1.01.2010	1.03.2010	1.06.2010	1.09.2010	1.01.2011
1 640	1 680	1 580	1 600	1 720.

Данный ряд, как и предыдущий, является моментным рядом с неравноотстоящими датами, но его отличие состоит в том, что значения показателя между указанными датами могли изменяться, т.е. не были фиксированными. Поэтому в данном случае для определения среднего уровня следует воспользоваться формулой (8.17):

$$\bar{y} = \frac{(1\ 640 + 1\ 680) \cdot 2 + (1\ 680 + 1\ 580) \cdot 3 + (1\ 580 + 1\ 600) \cdot 3 + (1\ 600 + 1\ 720) \cdot 4}{(2 + 3 + 3 + 4) \cdot 2} =$$

$$= \frac{39\ 240}{24} = 1\ 635 \text{ (чел.)}$$

### Средние показатели интенсивности развития

К этой группе показателей относят:

- *средний абсолютный прирост;*
- *средний коэффициент (темп) роста;*
- *средний темп прироста.*

**Средний абсолютный прирост, или средняя скорость роста,** рассчитывается как средняя арифметическая из показателей скорости роста за отдельные периоды:

$$\bar{\Delta y} = \frac{\Delta_1 + \Delta_2 + \dots + \Delta_{n-1}}{n-1} = \frac{\sum_{i=1}^{n-1} \Delta y_i}{n-1}, \quad (8.20)$$

где  $(n-1)$  – число абсолютных приростов за анализируемый период;

$n$  – число уровней ряда;

$\Delta y_i$  – цепные абсолютные приросты.

Так как 
$$\sum_{i=1}^{n-1} \Delta y_i = y_n - y_1, \quad (8.21)$$

то 
$$\bar{\Delta y} = \frac{y_n - y_1}{n-1}. \quad (8.22)$$

Этот показатель определяет, насколько в среднем за единицу времени (месяц, квартал, год и др.) должен увеличиваться уровень ряда (в абсолютном выражении), чтобы, двигаясь от начального уровня за данное число периодов (например, месяцев, кварталов, лет), достигнуть конечного уровня. Определяющим свойством среднего абсолютного прироста является общий абсолютный прирост за весь период, ограничивающий ряд динамики.

### Пример 8.7

Рассмотрим порядок расчета среднего абсолютного прироста по данным табл. 8.7:

$$\bar{\Delta y} = \frac{12 + 16 + 20 + 24 + 28 + 32}{7-1} = 22 \text{ (тыс. шт.)}$$

или

$$\bar{\Delta y} = \frac{232}{7-1} = 22 \text{ (тыс. шт.)}.$$

**Средний темп роста** ( $\bar{T}_p$ ) рассчитывается по формуле

$$\bar{T}_p = \bar{K}_p \cdot 100 \%, \quad (8.23)$$

где  $\bar{K}_p$  – средний коэффициент роста.

В свою очередь средний коэффициент роста определяется по формуле средней геометрической из цепных коэффициентов роста:

$$\bar{K}_p = \sqrt[n-1]{K_{p_1} \cdot K_{p_2} \cdot \dots \cdot K_{p_{n-1}}} = \sqrt[n-1]{\prod_{i=1}^{n-1} K_{p_i}}, \quad (8.24)$$

где  $K_{p_i}$  – цепные коэффициенты роста;

$n$  – число уровней ряда динамики.

Расчет среднего коэффициента (темпа) роста может быть выполнен и по другой формуле:

$$\bar{K}_p = \sqrt[n-1]{\frac{y_n}{y_0}}. \quad (8.25)$$

Для ряда динамики, представленного в табл. 8.7, средний темп роста определяется следующим образом:

$$\bar{T}_p = \bar{K}_p \cdot 100 \%,$$

$$\bar{K}_p = \sqrt[6]{1,12 \cdot 1,143 \cdot 1,156 \cdot 1,162 \cdot 1,163 \cdot 1,16} = 1,151$$

или

$$\bar{K}_p = \sqrt[6]{\frac{232}{100}} = 1,151.$$

Тогда  $\bar{T}_p = 1,151 \cdot 100 \% = 115,1 \%$ .

**Средний темп прироста** определяют по формуле

$$\bar{T}_{np} = \bar{T}_p - 100 \%. \quad (8.26)$$

Для динамического ряда, представленного в табл. 8.7, среднегодовой темп прироста составляет 15,1 %.

## 8.4. Характеристика и методы выявления основной тенденции развития явлений

Анализ динамики предполагает *выявление закономерностей развития явлений во времени*. Но закономерности не проявляются четко в каждом конкретном уровне, они проявляются лишь как тенденция в достаточно длительном периоде. Это происходит потому, что на основную закономерность развития накладываются другие составляющие:

- случайные;
- сезонные (периодические);
- циклические.

В общем случае динамика явлений складывается из четырех составляющих (компонент):

– *основной тенденции*, характеризующей основную закономерность развития исследуемого явления. Тенденцию представляют в виде тренда. *Тренд – это некоторая функция времени, которая выражает основную тенденцию ряда динамики;*

– *периодической компоненты*, отражающей влияние фактора сезонности на развития изучаемого явления;

– *циклической компоненты*, характеризующей циклические колебания, свойственные любому воспроизводственному процессу;

– *случайной компоненты*, проявляющейся как результат влияния множества случайных факторов.

Наиболее типичные ситуации, характеризующие состав компонент, формирующих уровни ряда динамики, представлены на рис. 8.4.

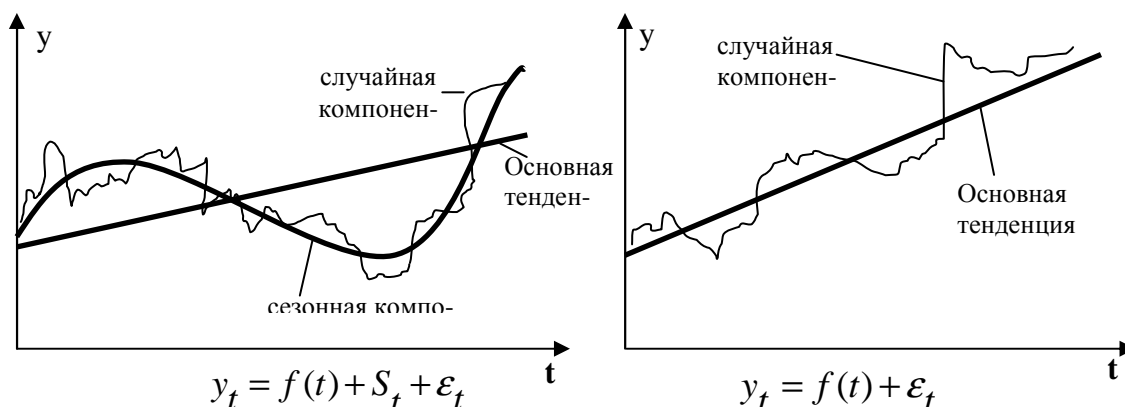


Рис. 8.4. Характеристика составляющих динамического ряда:  
 $f(t)$  – составляющая, характеризующая основную тенденцию развития;  
 $\varepsilon_t$  – случайная компонента;  $S_t$  – сезонная составляющая

Источник: [17, с. 277, рис. 12.1]

Рассмотрим используемые в статистике методы выявления основной тенденции развития: *метод укрупнения интервалов*; *метод скользящей средней* и *аналитическое выравнивание*.

### **Метод укрупнения интервалов ряда динамики**

При использовании этого метода исходный динамический ряд преобразуется и заменяется другим, показатели которого относятся к большим периодам. Например, ряд, содержащий данные о месячном выпуске продукции преобразуется в ряд квартальных данных. Вновь образованный ряд может содержать либо абсолютные величины за укрупненные периоды, либо средние величины.

При суммировании уровней или при выведении средних по укрупненным интервалам отклонения в уровнях, обусловленные случайными причинами, взаимопогашаются и действие основных факторов проявляется более отчетливо.

### **Метод скользящей средней**

Этот метод, как и предыдущий, является лишь эмпирическим приемом предварительного анализа тенденции.

*Скользящая средняя* – подвижная динамическая средняя, которая подсчитывается по динамическому ряду при последовательном передвижении на один интервал.

Продолжительность того периода, который принимается для расчета скользящей средней, называется *периодом скользящей средней*.

Если в динамическом ряду имеются периодические колебания (например, сезонные), период скользящей средней должен совпадать с периодом колебания или быть кратным ему. Например, имея дело с квартальными данными о заготовках сельхозпродуктов, период скользящей средней нужно брать равным четырем кварталам (году), т.к. колебания в таком ряду повторяются ежегодно, можно взять 8, 12 и т.д. кварталов.

Если в ряду периодические колебания отсутствуют, период скользящей средней подбирают, начиная с наименьшего (т.е. с двух уровней), укрупняя его до тех пор, пока в скользящей средней не будет более или менее ясно выступать тенденция развития явления.

Для определения скользящей средней формируются укрупненные интервалы, состоящие из одинакового числа уровней. Каждый последующий интервал получают, постепенно сдвигаясь от начального уровня динамического ряда на один уровень. Полученная средняя относится к середине укрупненного интервала. Для сглаживания скользящей средней удоб-

нее укрупненный интервал составлять из нечетного числа уровней. Нахождение скользящей средней по четному числу уровней создает неудобство, проявляющееся в том, что средняя может быть отнесена только к середине между двумя периодами. В таких случаях проводят последующее центрирование данных.

### **Аналитическое выравнивание ряда динамики**

Для того чтобы количественно описать основную тенденцию развития фактические значения показателей динамического ряда заменяют уровнями, вычисленными по уравнениям того или иного вида. Такой метод выявления основной тенденции развития называют аналитическим выравниванием.

При этом методе основная закономерность ряда динамики определяется как функция

$$\hat{y}_t = f(t), \quad (8.28)$$

где  $\hat{y}_t$  – уровни динамического ряда, вычисленные по соответствующему аналитическому уравнению на момент времени  $t$ .

Уровень ряда изменяется от периода к периоду не потому, что прошло какое-то время, а потому, что в течение этого времени действовали различные факторы, с разной интенсивностью.

С помощью этого метода могут выравниваться уровни ряда как содержащие, так и не содержащие сезонную компоненту. Однако если последняя имеется, то уровни выравниваемых интервальных рядов должны быть не менее *годовых*, т.к. в годовых и больших уровнях *сезонные компоненты нивелируются*. Уровни моментных рядов с сезонной компонентой должны относиться к одинаковым моментам года, в этом случае сезонные колебания также не оказывают влияния на динамику. Уровни рядов, не содержащих периодических колебаний, могут относиться к любым периодам.

В зависимости от исходных данных для описания основной тенденции ряда могут быть выбраны разные типы функций. Если для динамического ряда характерны более или менее постоянные абсолютные цепные приросты, выравнивание производят по прямой; если постоянны темпы роста – по показательной кривой (экспоненте); если постоянно ускорение (вторые абсолютные разности) – по параболе второго порядка.

Рассмотрим методику аналитического выравнивания динамического ряда на примере прямой.

Применительно к динамическим рядам уравнение прямой имеет вид:

$$\hat{y}_t = a + bt, \quad (8.29)$$



где  $\hat{y}_t$  – теоретические уровни, рассчитанные по уравнению;

$t$  – порядковый номер момента времени или периода. Величина нам известная;

$a, b$  – параметры прямой.

Параметры  $a$  и  $b$  в соответствии с методом наименьших квадратов находим из решения системы уравнений

$$\begin{cases} na + b \sum t = \sum y \\ a \sum t + b \sum t^2 = \sum yt, \end{cases} \quad (8.30)$$

где  $n$  – число уровней динамического ряда;

$t$  – время, к которому относится данный уровень ряда динамики. Для введения фактора времени в уравнение используется определенная схема нумерации.

Если  $n$  четное, то, например, для ряда с шестью уровнями ( $n = 6$ )  $t$  будет обозначаться следующим образом: (-5); (-3); (-1); (+1); (+3); (+5). При этом (-5) относится к первому уровню ряда динамики, (-3) – ко второму, (-1) – к третьему и т.д. Если нечетное, то, например, для ряда с семью уровнями  $t$  следует обозначить по схеме: (-3); (-2); (-1); (0); (+1); (+2); (+3) (при  $n = 7$ ).

В обоих случаях  $\sum t = 0$ . С учетом того, что  $\sum t = 0$ , упрощается определение параметров уравнения:

$$a = \frac{\sum y}{n}, \quad (8.31)$$

$$b = \frac{\sum yt}{\sum t^2}. \quad (8.32)$$

При равномерном изменении скорости (стабильном ускорении) выравнивание производится по параболе второго порядка:

$$\hat{y}_t = a + bt + ct^2, \quad (8.33)$$

$$\begin{cases} an + b \sum t + c \sum t^2 = \sum y \\ a \sum t + b \sum t^2 + c \sum t^3 = \sum yt \\ a \sum t^2 + b \sum t^3 + c \sum t^4 = \sum yt^2. \end{cases} \quad (8.34)$$

Найденные теоретические уровни динамического ряда отражают детерминированную составляющую ряда, состоящую из одной основной тенденции или  $f(t) + S$ .

Разность между  $y$  и  $\hat{y}_t$  характеризует случайную составляющую. Она будет тем меньше, чем точнее выбранная функция воспроизводит динамику явления.

Ошибка уравнения:

$$S_{\hat{y}_t} = \sqrt{\frac{\sum (y - \hat{y}_t)^2}{n - m}}, \quad (8.35)$$

где  $n$  – число уровней ряда динамики;

$m$  – число параметров уравнения.

Чем точнее уравнение воспроизводит (моделирует) ряд динамики, тем больше его прикладное значение. В частности, оно может быть использовано для *интерполяции* и *экстраполяции*.

*Интерполяция* – приблизительный расчет недостающих уровней внутри однородного периода, когда известны уровни, лежащие по обе стороны неизвестного.

*Экстраполяция* – приблизительный расчет уровней ряда динамики за пределами анализируемого периода. Возможна экстраполяция в прошлое и будущее.

Экстраполяция на будущее является одним из статистических методов прогнозирования. Основным условием прогнозирования указанным методом является сохранение в будущем условий, определявших тенденцию развития явления в прошлом.

Оценка точности прогнозов производится с помощью доверительных интервалов прогноза:

$$\hat{y}_t - t_{\alpha; n-m} S_{\hat{y}_t} K \leq y_{\text{прог.}} \leq \hat{y}_t + t_{\alpha; n-m} S_{\hat{y}_t} K, \quad (8.36)$$

где  $t_{\alpha; n-m}$  – табличное значение  $t$ -критерия Стьюдента с  $(n - m)$  степенями свободы и уровнем значимости  $\alpha$ ;

$K$  – коэффициент, уточняющий интервальный прогноз в зависимости от  $n$  и периода упреждения ( $L$ ).

Для уравнения прямой этот коэффициент рассчитывается по формуле

$$K = \sqrt{\frac{n+1}{n} + \frac{3(n+2L-1)^2}{n(n^2-1)}}. \quad (8.37)$$

## 8.5. Понятие и оценка сезонной неравномерности развития

*Колебания динамики, связанные со сменой времен года, называют сезонными.* Таким образом, сезонность рассматривают как *внутригодовую*

*динамику*. Сезонность развития явлений может быть обусловлена разными причинами:

- сезонным характером спроса на товары и услуги;
- неравномерностью производственной деятельности в отраслях, связанных с переработкой сельскохозяйственного сырья;
- сезонным характером производства.

Для целей планирования и управления необходимо выявлять и оценивать сезонную неравномерность. Измеряют сезонную неравномерность путем расчета специальных показателей – *индексов сезонности*.

Используют следующие методы расчета индексов сезонности.

Если данные о помесечных уровнях показателя приводятся за несколько лет и отсутствует сколько-нибудь значительная динамика годовых уровней (цепные темпы прироста по абсолютной величине не превышают 5 %) используют формулу

$$i_{\text{сез.}} = \frac{\bar{y}_i}{\bar{y}}, \quad (8.38)$$

где

$$\bar{y}_i = \frac{\sum_{j=1}^n y_{ij}}{n}, \quad (8.39)$$

где  $y_{ij}$  – уровень ряда динамики в  $i$ -том месяце,  $j$ -том году;

$\bar{y}$  – средний месячный уровень ряда за весь анализируемый период:

$$\bar{y} = \frac{\sum_{j=1}^n \sum_{i=1}^{12} y_{ij}}{12n}, \quad (8.40)$$

где  $\bar{y}_i$  – средний уровень ряда динамики в  $i$ -том месяце за  $n$  лет.

Если данные о помесечных уровнях показателя приводятся за несколько лет и обнаруживается тенденция к значительному росту или падению годовых уровней ( $|T_{np}| > 5\%$ ), используют формулу

$$i_{\text{сез.}} = \frac{\sum_{j=1}^n \frac{y_{ij}}{\bar{y}_j}}{n}, \quad (8.41)$$

где  $\bar{y}_j$  – среднемесячный уровень в  $j$ -том году:

$$\bar{y}_j = \frac{\sum_{i=1}^{12} y_{ij}}{12}. \quad (8.42)$$

Для обобщения сезонной неравномерности внутри года, рассчитывают среднее квадратическое отклонение индексов сезонности:

$$\sigma_{i_{\text{сез.}}} = \sqrt{\frac{\sum_{i=1}^{12} (i_{\text{сез.}} - 1)^2}{p}}, \quad (8.43)$$

$$\sigma_{i_{\text{сез.}}} = \sqrt{\frac{\sum_{i=1}^4 (i_{\text{сез.}} - 1)^2}{p}}, \quad (8.44)$$

где  $p$  – число интервалов внутри года (12 месяцев или 4 квартала).

Сравнение  $\sigma_{i_{\text{сез.}}}$  за разные годы показывает сдвиги в сезонности.

Снижение величины  $\sigma_{i_{\text{сез.}}}$  свидетельствует об уменьшении сезонности.

## 8.6. Методика оценки компонент динамики

Оценка компонент динамического ряда проводится поэтапно.

На первом этапе определяют индексы сезонности. В соответствии с методикой оценки индексов сезонности, представленной в предыдущем пункте данной темы, расчет этих показателей сезонной динамики выполняется по следующим формулам:

$$i_{\text{сез}} = \frac{\bar{y}_i}{\bar{y}}, \quad (8.45)$$

$$i_{\text{сез.}} = \frac{\sum_{j=1}^n \frac{y_{ij}}{\bar{y}_j}}{n}. \quad (8.46)$$

Напомним, что выбор формул для расчета индексов сезонности производится в зависимости от интенсивности годовой динамики.

На втором этапе определяют значения уровней ряда динамики, которые не содержат сезонную составляющую, их обозначим через  $d$ . Исключение сезонной неравномерности производят путем следующего расчета:

$$d = \frac{y_t}{i_{\text{сез}}} = \bar{y}_t \cdot c_t \cdot \varepsilon_t. \quad (8.47)$$

Третий этап предполагает определение уровней динамического ряда, представляющих основную тенденцию развития. Расчеты на этом этапе проводятся на основе данных, для которых исключена сезонная неравно-

мерность, т.е. на основе данных, полученных на предыдущем этапе. Использование метода аналитического выравнивания позволяет определить  $\hat{y}_t$ , представляющих основную закономерность развития.

На завершающем четвертом этапе анализа определяют циклическую составляющую. Для этого сначала выделяют циклическую и случайную составляющие:

$$\frac{d}{\hat{y}_t} = \frac{y_t \cdot c_t \cdot \varepsilon_t}{\hat{y}_t} = c_t \cdot \varepsilon_t. \quad (8.48)$$

Затем определяют циклическую составляющую на основе метода скользящей средней, где объектом осреднения являются значения  $(c_t \cdot \varepsilon_t)$ . Полученные усредненные значения на основе скользящей средней и дают циклическую составляющую.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ**

### **Вопросы для самоподготовки и контроля знаний**

1. Что такое ряд динамики?
2. Какова роль анализа динамики в экономических исследованиях?
3. Каковы основные элементы ряда динамики?
4. Какие существуют виды рядов динамики?
5. Какое значение в анализе динамических рядов имеет обеспечение сопоставимости их уровней?
6. Перечислите основные условия сопоставимости уровней ряда динамики.
7. В чем сущность такого приема приведения уровней динамического ряда в сопоставимый вид как смыкание рядов динамики?
8. Какие показатели включают в группу аналитических показателей ряда динамики?
9. В чем сущность цепной и базисной схем расчета показателей динамики?
10. Какова роль средних показателей ряда динамики?
11. Различается ли порядок расчета среднего уровня ряда динамики в моментном и интервальном рядах?
12. Назовите состав компонент, формирующих уровни ряда динамики.
13. Что такое тренд?

14. Какие существуют методы выявления основной тенденции развития?
15. Как рассчитать скользящую среднюю и для каких целей она может быть использована?
16. Каковы преимущества выявления основной тенденции развития на основе аналитического выравнивания?
17. Как оценивается точность уравнения тренда?
18. Что такое экстраполяция?
19. Какие условия должны соблюдаться при прогнозировании на основе экстраполяции?
20. Что такое сезонность в рядах динамики и каково значение ее оценки?
21. Как рассчитываются индексы сезонности?
22. В чем состоит сущность методики оценки циклической составляющей в ряду динамики?

### Тест 8.1

1. Уровни ряда динамики могут быть представлены ...
  - а) абсолютными величинами;
  - б) средними величинами;
  - в) относительными величинами;
  - г) атрибутивными признаками.
2. Ряд динамики характеризует ...
  - а) структуру совокупности по какому-либо признаку;
  - б) изменение значений признака в пространстве;
  - в) изменение значений признаков во времени.
3. Уровни интервального динамического ряда выражают ...
  - а) размер явлений за определенный период;
  - б) размер явлений на определенную дату.
4. Моментным рядом динамики является ряд, представленный ...
  - а) показателями производства велосипедов за каждый квартал отчетного года;
  - б) показателями валютных запасов Республики на конец каждого квартала отчетного года;
  - в) показателями объема капитальных вложений в экономику страны по годам.
5. Интервальным рядом динамики является ряд, представленный ...
  - а) показателями парка автобусов в автотранспортном предприятии на конец каждого квартал отчетного года;

б) показателями производства картофеля в районе за каждый год в период 2007 – 2010 гг.;

в) ежегодный выпуск специалистов университета в период 2007 – 2010 г.г.

6. При цепной системе расчета показателей интенсивности изменения уровней ряда динамики анализируемый уровень ряда сравнивается ...

а) с предыдущим уровнем ряда;

б) с начальным уровнем ряда.

7. Согласны ли Вы с утверждением, что базисные показатели динамики характеризуют интенсивность изменения уровней от периода к периоду в пределах изучаемого интервала времени?

а) да;

б) нет.

8. Если все уровни ряда динамики сравниваются с одним и тем же уровнем, показатели динамики называются ...

а) цепными;

б) базисными.

9. Абсолютный прирост – это ...

а) величина только положительная;

б) величина только отрицательная;

в) может быть положительной и отрицательной.

10. Коэффициент роста – это ...

а) разность двух уровней ряда динамики;

б) отношение двух уровней ряда динамики;

в) отношение абсолютного прироста к уровню, принятому за базу сравнения.

## Тест 8.2

1. При каких значениях темпа роста делают вывод:

а) об увеличении уровней динамического ряда;

б) об уменьшении уровней динамического ряда;

в) о неизменности уровней динамического ряда?

2. Чему равен темп роста, если темп прироста составляет ... ?

а) 4,5 %;

б) (-9,8 %);

в) 0 %.

3. Абсолютный прирост исчисляется как ...

а) как отношение двух уровней ряда динамики;

б) как разность двух уровней ряда динамики;

в) как произведение двух уровней ряда динамики.

4. В 2010 г. по сравнению с 2009 г. численность работающих в организации увеличилась на 5 %. Абсолютное значение 1 % прироста составило 10 человек. Исходя из этого численность работающих в организации в 2010 г. составила ...

- а) 105 человек;
- б) 1 050 человек;
- в) 1 000 человек.

5. Для выявления основной тенденции развития используют ...

- а) аналитическую группировку;
- б) метод укрупнения интервалов;
- в) метод скользящей средней;
- г) метод аналитического выравнивания.

6. Точность прогноза при увеличении периода упреждения ...

- а) увеличивается;
- б) уменьшается.

7. Сезонность – это динамика явлений, которая выявляется в пределах ...

- а) года;
- б) квартала;
- в) полугодия.

8. Сумма месячных индексов сезонности за год составляет ...

- а) 12;
- б) 1;
- в) 0.

9. Скорость роста оценивают с помощью ...

- а) базисных коэффициентов роста;
- б) цепных абсолютных приростов;
- в) базисных абсолютных приростов;
- г) вторых разностей.

10. Показатель абсолютного ускорения применяют ...

- а) только в цепном варианте расчета;
- б) только в базисном варианте расчета;
- в) как в цепном, так и в базисном варианте.

### **Задания и задачи по теме**

#### **Задача 8.1**

Используя взаимосвязь показателей динамики, определите уровни ряда динамики и недостающие в таблице базисные показатели динамики по следующим данным о производстве часов в регионе за 2002-2010 гг.:



Год	Производство часов, млн шт.	Базисные показатели динамики		
		абсолютный прирост, млн шт.	темпы роста, %	темпы прироста, %
2002	55,1	–	100,0	–
2003		2,8		
2004			110,3	
2005				14,9
2006				17,1
2007			121,1	
2008		13,5		
2009				25,4
2010		14,0		

### Задача 8.2

Используя взаимосвязь показателей динамики, определите уровни ряда динамики и недостающие в таблице цепные показатели динамики по следующим данным о производстве продукции предприятиями объединения (в сопоставимых ценах):

Год	Производство продукции, млн руб.	По сравнению с предыдущим годом			
		абсолютный прирост, млн руб.	темпы роста, %	темпы прироста, %	абсолютное значение 1 % прироста, млн руб.
2005	92,5				
2006		4,8			
2007			104,0		
2008				5,8	
2009					
2010		7,0			1,15

### Задача 8.3

Имеются следующие данные об объеме пассажирооборота по автобусным предприятиям города:

Год	Пассажирооборот, млрд пасс.-км	Цепные показатели динамики			
		абсолютный прирост, млрд пасс.-км	коэффициент роста	темпы прироста, %	абсолютное значение 1 % прироста, млрд пасс.-км
2005	127,0	–	–	–	–
2006			1,102		
2007				7,1	
2008	164,60				
2009					
2010				9,9	1,75

#### Задача 8.4

Имеются следующие данные по предприятию о производстве промышленной продукции за 2005 – 2010 гг. (в сопоставимых ценах), млрд руб.:

2005	2006	2007	2008	2009	2010
67,7	73,2	75,6	77,9	82,0	84,4

Определите:

- 1) цепные и базисные абсолютные приросты;
- 2) цепные и базисные темпы роста и темпы прироста;
- 3) для каждого года абсолютное значение 1 % прироста.

Результаты расчетов представьте в табличной форме. Сделайте выводы.

#### Задача 8.5

По данным предыдущей задачи определите:

- 1) средний уровень ряда динамики;
- 2) среднегодовой абсолютный прирост;
- 3) среднегодовой темп роста и прироста.

#### Задача 8.6

Динамика выпуска продукции в ОАО «Альфа» характеризовалась следующими данными:

	2006	2007	2008	2009	2010
Объем производства продукции в сопоставимых ценах, млрд руб.	21,2	22,4	24,9	28,6	31,6

Определите:

- 1) цепные и базисные абсолютные приросты;
- 2) цепные и базисные темпы роста и темпы прироста;
- 3) для каждого года абсолютное значение 1% прироста;
- 4) среднегодовой темп роста и прироста;
- 5) среднегодовой абсолютный прирост.

Результаты расчетов представьте в табличной форме. Сделайте выводы.

#### Задача 8.7

В таблице представлены данные об активах предприятия на начало каждого квартала за 2007 – 2009 гг.

Определить абсолютное и относительное изменение активов предприятия в 2009 г. по сравнению с 2008 и 2007 гг.

Год	Активы предприятия, млрд руб.			
	1.01	1.04	1.07	1.10
2007	62	65	70	68
2008	68	70	75	78
2009	80	84	88	90
2010	95	–	–	–

### Задача 8.8

Имеются следующие данные о поголовье коров в хозяйствах всех категорий области, тыс. голов:

	2003	2004	2005	2006	2007	2008	2009	2010
На 1 января	37,6	38,1	40,1	42,5	–	–	–	–
На 1 июля	–	–	–	44,7	44,8	45,0	46,0	46,1

Укажите причину несопоставимости уровней ряда динамики. Приведите уровни ряда к сопоставимому виду, применив метод смыкания рядов динамики.

### Задача 8.9

Имеются следующие данные о производстве продукта А в республике за 2004 – 2010 гг.:

Год	2004	2005	2006	2007	2008	2009	2010
Производство продукции, млн т	32,0	33,5	35,0	37,0	38,5	40,0	42,0

Выявите основную тенденцию производства продукта А за анализируемый период методом аналитического выравнивания ряда динамики. Выбор вида уравнения тренда обоснуйте. Определите ошибку уравнения тренда. С вероятностью 0,95 определите интервал, в котором можно ожидать объем производства продукта в 2011 г. Укажите условия, при которых прогноз может быть получен методом экстраполяции.

### Задача 8.10

Ежегодный объем производства продукта в регионе характеризуется следующими данными, тыс. т:

2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
20	22	23	25	28	26	27	27	30	29	32

Для изучения основной тенденции производства продукции произведите:

- 1) сглаживание уровней ряда динамики с помощью трехчленной скользящей средней;
- 2) аналитическое выравнивание.

На графике представьте фактические уровни ряда динамики, а также значения, полученные на основе выравнивания методом скользящей средней и методом аналитического выравнивания.

### Задача 8.11

По станциям технического обслуживания легковых автомобилей города имеются следующие данные:

Месяцы	Число заявок, тыс. шт.		Индексы сезонности годовые		Индексы сезонности за 2 года
	2009	2010	2009	2010	
Январь	10,3	13,6			
Февраль	11,3	14,3			
Март	11,5	14,4			
Апрель	12,0	14,6			
Май	12,6	15,6			
Июнь	16,0	17,1			
Июль	15,9	16,9			
Август	16,2	17,0			
Сентябрь	16,4	16,5			
Октябрь	15,2	16,0			
Ноябрь	15,0	14,9			
Декабрь	12,8	13,8			
Итого:	165,0	184,7			

На основе приведенных данных выполните следующее:

- 1) определите индексы сезонности в 2009 и 2010 гг.;
- 2) определите индексы сезонности за два года вместе;
- 3) рассчитайте средние квадратические отклонения индексов сезонности для каждого года и сделайте выводы об изменении сезонной неравномерности по годам.

### Задание 8.12

По периодическим изданиям, данным статистических сборников, сайта Национального статистического комитета Республики Беларусь найдите примеры рядов динамики различных видов и рассчитайте для них показатели интенсивности изменения уровней ряда динамики, а также средние показатели ряда динамики.

### Задание 8.13

По периодическим изданиям, данным статистических сборников, сайта Национального статистического комитета Республики Беларусь найдите ряд динамики с целью выявления основной тенденции развития известными вам методами.

## Тема 9

# ИНДЕКСНЫЙ МЕТОД В СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ

- 9.1. Сущность индексов и значение индексного метода анализа.
- 9.2. Классификация индексов.
- 9.3. Индексы количественных показателей.
- 9.4. Индексы качественных показателей.
- 9.5. Система индексов переменного состава, постоянного состава и структурных сдвигов.
- 9.6. Индексы пространственно-территориального сопоставления.
- 9.7. Применение индексов в факторном анализе.
- 9.8. Решение типовых задач по теме.

Изучив данную тему, студенты должны:

- ясно понимать сущность индексов и их значение в статистико-экономическом анализе;
- знать методику построения индивидуальных и общих индексов количественных и качественных показателей;
- уметь интерпретировать результаты индексного анализа.

### 9.1. Сущность индексов и значение индексного метода анализа

Термин «индекс» происходит от латинского слова *index*, что в переводе означает указатель, показатель.

В статистике индексы относятся к важнейшим обобщающим показателям. С их помощью характеризуется развитие национальной экономики в целом и отдельных ее отраслей, анализируются результаты производственно-хозяйственной деятельности предприятий и организаций, оценивается роль отдельных факторов в формировании важнейших экономических показателей, выявляются резервы улучшения использования ресурсов производства. Индексы используются также в международных сопоставлениях экономических показателей, определении уровня развития человеческого потенциала, мониторинге деловой активности в экономике и т.д.

*Индекс – это относительный показатель, который оценивает изменение социально-экономических явлений во времени, в пространстве или по сравнению с планом.*

В статистической практике индексы принято обозначать символами *i* и *I* (начальная буква латинского слова *index*). Буквой *i* обозначаются индивидуальные (частные) индексы, буквой *I* – общие индексы.

Поскольку индексы относятся к относительным величинам, их рассчитывают путем деления одной величины на другую. Одна из этих величин называется *сравниваемой* (или *текущей*), а вторая, с которой производится сравнение, – *базисной*. Сравниваемой величине присваивается подстрочный индекс «1», базисной – «0».

Выбор базы сравнения зависит от цели анализа.

*Если изучают динамику явления*, то в качестве базы берут размер явления в каком-либо периоде, предшествующем анализируемому. При этом возможны два способа расчета индексов: *базисный* и *цепной*. *Цепные индексы* получают сопоставлением текущих уровней с предшествующим. *Базисные индексы* получают сопоставлением текущих уровней с уровнем периода, принятого за базу сравнения.

*При территориальных сравнениях*, когда исследуется изменение явления в пространстве (по странам, регионам, предприятиям и т.п.), в качестве базы сравнения принимают соответствующие показатели по интересующим исследователя объектам.

*При использовании индексов* как показателей выполнения плана за базу сравнения принимаются плановые показатели.

Результаты расчета индексов выражаются в виде коэффициентов или в процентах.

В теории статистики для обозначения индексируемых показателей используются определенные символы:

$q_i$  – количество произведенной продукции (проданного товара) данного вида в натуральном выражении;

$p_i$  – цена продукции или товара данного вида;

$z_i$  – себестоимость единицы продукции данного вида;

$t_i$  – трудоемкость единицы продукции данного вида;

$T$  – общие затраты рабочего времени;

$F = zq$  – общие затраты на производство продукции данного вида;

$Q = pq$  – общая стоимость произведенной продукции данного вида или товарооборот.

С помощью индексов решаются следующие аналитические задачи:

– определяются средние изменения во времени сложных совокупностей, состоящих из непосредственно несоизмеримых элементов (например, изменение физического объема продукции промышленности). В этом случае индексы выступают как *синтетические показатели динамики*;

– оценивается динамика средней величины экономического показателя, а также влияние изменения структуры явления на индексную вели-

чину. Эта задача решается с помощью *системы индексов переменного состава, постоянного состава и структурных сдвигов*;

– измеряются средние соотношения сложных явлений в пространстве (по разным объектам: странам, регионам, предприятиям и др.). Эта задача решается с помощью *индексов пространственно-территориальных сопоставлений*;

– оценивается роль отдельных факторов в общем изменении сложного явления (результативного признака) путем элиминирования влияния других факторов. В решении этой задачи индексы выступают как *аналитический инструмент*;

– оценивается средняя степень выполнения плана по совокупности в целом (или ее части). В этом случае используют *индексы выполнения плана*;

– производится пересчет стоимостных экономических показателей из фактических цен в сопоставимые. Для этих целей используют *индексы цен*.

Каждая из этих задач решается с помощью различных индексов, виды которых рассматриваются ниже.

## 9.2. Классификация индексов

Используемые в статистико-экономическом анализе индексы можно классифицировать по ряду признаков, их состав представлен на рис. 9.1.

Охарактеризуем указанные виды индексов.

По *степени охвата элементов совокупности* индексы могут быть *индивидуальными, групповыми и общими*.

*Индивидуальные индексы* оценивают изменение интересующего исследователя показателя по одному определенному элементу совокупности. В общем случае индивидуальный индекс рассчитывается по формуле

$$i_x = \frac{x_1}{x_0}, \quad (9.1)$$

где  $x_0, x_1$  – значения индексируемого показателя по сравниваемым вариантам.

Отметим, что *индексируемым называется показатель, изменение которого оценивается с помощью индекса*. Примером индивидуального индекса может служить индекс цены на конкретное изделие, индекс физического объема продукции конкретного вида и т.п.

*Общий индекс* оценивает изменение индексируемой величины по всем элементам исследуемой совокупности. Общий индекс иначе называют *сводным*.

Расчет общих индексов выполняется по формулам:

$$I_x = \frac{\sum x_1 S_1}{\sum x_0 S_1} \quad (9.2)$$

или

$$I_x = \frac{\sum x_1 S_0}{\sum x_0 S_0}, \quad (9.3)$$

где  $x_0, x_1$  – значения индексируемой величины по данному элементу совокупности по сравниваемым вариантам;

$S_0, S_1$  – веса, с помощью которых соизмеряются, а затем суммируются элементы исследуемой совокупности (по сравниваемым вариантам).

Формула (9.2) представляет расчет общего индекса по методу Пааше и ее называют *формулой Пааше*. Формулу (9.3) называют *формулой Ласпейреса*. Выбор формулы для расчета общего индекса главным образом зависит от содержания (вида) индексируемой величины. Правила выбора формулы общего индекса рассматриваются ниже.

Если индексы охватывают не все элементы сложного явления, а только часть из них, то такие индексы называют *групповыми*, или *субиндексами*, например, индексы физического объема продукции по отдельным отраслям промышленности, индексы цен по группам продовольственных и непродовольственных товаров. Групповые индексы отражают закономерности развития отдельных частей изучаемых явлений. В таких индексах проявляется их связь с методом группировок.

По *базе сравнения* индексы могут быть территориальными, выполнения плана и динамики. Их характеристика представлена в п. 9.1 данной темы.

По *виду индексируемых показателей* индексы подразделяются на *индексы количественных* (объемных) и *качественных показателей*.

К индексам количественных показателей относятся индексы физического объема продукции, валового внутреннего продукта, реализации. Большим разнообразием отличаются индексы качественных показателей. К ним относят индексы цен, себестоимости, трудоемкости, производительности труда, фондоотдачи и т.п.

*Количественные (объемные) показатели* характеризуют либо размер совокупности, либо размер признака в совокупности. Например, численность работников предприятия и фонд их заработной платы.

*Качественные признаки* характеризуют интенсивность использования ресурсов, размер признака в расчете на единицу совокупности. Например, производительность труда, фондоотдача, материалоемкость, трудоемкость, цена и др.



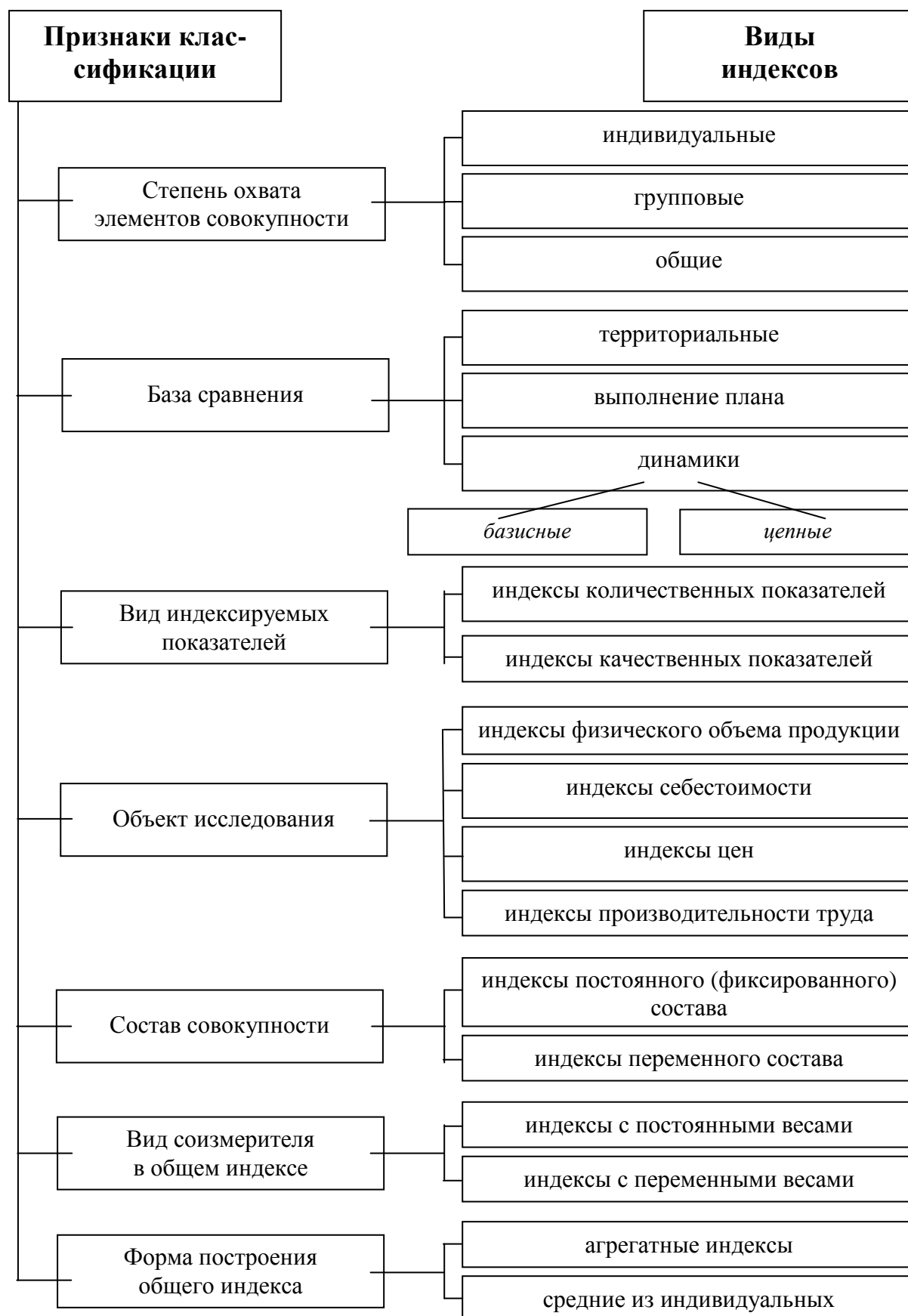


Рис. 9.1. Виды индексов  
 Источник:[4, с. 408, рис. 6.1]

В практике расчета общих индексов динамики придерживаются двух основных правил:

– при построении общих индексов количественных показателей в качестве весов используют качественные показатели, относящиеся к базисному периоду;

– при построении общих индексов качественных показателей в качестве весов используют количественные показатели, относящиеся к отчетному периоду.

Другими словами, при построении общих индексов количественных показателей используют формулу Ласпейреса, а при построении общих индексов качественных показателей – формулу Пааше.

Из указанных правил принятия весов при построении общих индексов динамики количественных и качественных показателей практика делает исключения, ориентируясь на конкретную ситуацию и исследовательскую задачу. Примером такого исключения является расчет индекса потребительских цен, который осуществляется на основе формулы Лайпейреса.

По *объекту исследования* индексы бывают: производительности труда, себестоимости, физического объема продукции, стоимости продукции и др.

По *составу совокупности* выделяют два вида индексов: индексы постоянного (фиксированного) состава и индексы переменного состава. Деление индексов на эти два вида используют для анализа динамики средних показателей. Подробно эти виды индексов рассматриваются в п. 9.5 данной темы.

По *виду соизмерителя* (весов) общие индексы динамики могут быть с постоянными и переменными весами. При построении общих индексов динамики за несколько периодов можно использовать веса либо меняющиеся от периода к периоду, либо использовать один и тоже базисный вес. В первом случае будем иметь общие индексы с переменными весами, во втором – с постоянными весами.

*В зависимости от формы построения* общие индексы могут быть агрегатными и средними из индивидуальных.

Средние из индивидуальных индексов, в свою очередь, делятся на средние арифметические взвешенные индексы и средние гармонические взвешенные.

*Агрегатные индексы* являются основной формой общего индекса. Свое название они получили потому, что в числителе и в знаменателе с помощью весов соизмеряются, а затем суммируются наборы или агрегаты

индексируемых показателей. Числитель и знаменатель агрегатного индекса представляют собой сумму произведений индексируемого показателя по каждому элементу совокупности на его вес. Произведений будет столько, сколько элементов входит в изучаемую совокупность. Первым множителем каждого произведения является индексируемый показатель, характеризующий определенный элемент совокупности: в числителе – отчетного периода, в знаменателе – базисного. Вторым множителем – вес в числителе и знаменателе одинаков, т.е. принимается на одном и том же уровне, что позволяет элиминировать его влияние на величину общего индекса.

Средние из индивидуальных индексов целесообразно использовать, когда информация по индексируемому показателю дана не в виде абсолютных величин, а в виде индивидуальных индексов.

*Средний арифметический индекс* тождественен агрегатному индексу, если весами индивидуальных индексов будут слагаемые знаменателя агрегатного индекса. Средние арифметические индексы чаще всего на практике применяются для расчета общих индексов количественных показателей.

*Средний гармонический индекс* тождественен агрегатному, если индивидуальные индексы взвешены с помощью слагаемых числителя агрегатного индекса. Средние гармонические индексы, как правило, используются для построения общих индексов качественных показателей.

Порядок расчета конкретных общих индексов в агрегатной форме и форме средних из индивидуальных индексов рассматривается в п.п. 9.3 и 9.4 данной темы.

### 9.3. Индексы количественных показателей

Типичным индексом количественных показателей является **индекс физического объема**. Он может быть рассчитан для различных совокупностей в виде:

- индекса физического объема промышленной продукции;
- индекса физического объема реализации;
- индекса физического объема товарооборота;
- индекса физического объема ВВП и др.

*Индивидуальный индекс физического объема  $i_q$*  показывает, во сколько раз возрос (уменьшился) выпуск продукции конкретного вида в отчетном периоде по сравнению с базисным, или сколько процентов составляет рост (снижение) количества продукции (ее физического объема). Динамика физи-

ческого объема продукции конкретного вида оценивается на основе расчета следующего индивидуального индекса:

$$i_q = \frac{q_1}{q_0}, \quad (9.4)$$

где  $q_0, q_1$  – количество продукции данного вида в базисном и отчетном периодах соответственно.

В знаменателе формулы (9.4) может быть приведено не только количество продукции, фактически произведенной за какой-либо предыдущий период, но и плановое значение ( $q_{пл}$ ). Тогда формула (9.4) примет следующий вид:

$$i_q^{ВП} = \frac{q_1}{q_{пл}}. \quad (9.5)$$

Индивидуальный индекс физического объема продукции, полученный на основе формулы (9.5), является индексом выполнения планового задания.

Особенность построения общих индексов физического объема связана с несоизмеримостью элементов сложного явления. Если есть данные о выпуске продукции каждого вида только в натуральном выражении, то нельзя оценить динамику объема выпуска всей продукции в целом, т.к. разные виды продукции неравноценны по количеству затраченного на их производство общественного труда и имеют разные потребительные стоимости. Поэтому нельзя суммировать объемы производства по разным видам продукции. Для этого необходимо привести их в соизмеримый вид с помощью *соизмерителей-весов*. При индексации физического объема продукции в качестве *соизмерителей* можно принимать разные признаки: цену, себестоимость единицы продукции, затраты труда по норме или фактическую трудоемкость. Выбор соизмерителя в каждом конкретном случае зависит от цели исследования и исходной информации.

Более широко при построении общих индексов физического объема используют цены как более универсальный соизмеритель.

Сопоставляя объемы производства продукции в ценах соответствующих периодов, получаем агрегатный индекс стоимости продукции:

$$I_{qp} = \frac{\sum q_1 p_1}{\sum q_0 p_0}. \quad (9.6)$$

Но этот индекс учитывает изменение не только физического объема продукции, но и изменение цен.

Чтобы оценить изменение только физического объема, необходимо элиминировать (устранить) влияние изменения цен. Для этого количество продукции каждого вида в сравниваемых периодах умножают на одинаковые цены:

$$I_q = \frac{q_1^A p^A + q_1^B p^B + \dots + q_1^N p^N}{q_0^A p^A + q_0^B p^B + \dots + q_0^N p^N}. \quad (9.7)$$

Такой индекс называют агрегатным индексом физического объема.

По правилу построения общих индексов количественных показателей в качестве веса следует принимать цену базисного периода, т.е. использовать формулу Ласпейреса:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}. \quad (9.8)$$

Целесообразность выбора этой формулы связана с тем, что если использовать цены отчетного периода ( $p_1$ ), то они применялись бы только один раз и для каждого следующего отчетного года нужно брать новую цену. Для выражения объема производства отчетного периода в ценах базисного периода следует использовать индексы цен, определяемые Национальным статистическим комитетом Республики Беларусь.

От общего индекса  $I_q$  в агрегатной форме переходят к среднему из индивидуальных:

$$I_q = \frac{\sum i_q q_0 p_0}{\sum q_0 p_0} = \sum i_q d q_0 p_0. \quad (9.9)$$

#### 9.4. Построение индексов качественных показателей

К индексам качественных показателей относят:

- индексы себестоимости;
- индексы цен;
- индексы производительности труда;
- индексы трудоемкости продукции и др.

При построении общих индексов качественных показателей в роли весов индексируемой величины выступают количественные показатели, например, количество продукции, численность работников и др.

Напомним, что *при построении общих индексов качественных показателей веса устанавливаются на уровне отчетного периода*. Однако из этого

правила есть исключения. Например, индексы потребительских цен строят с использованием весов базисного периода.

В данной теме рассмотрим порядок построения индивидуальных и общих индексов цен. Индексы других качественных показателей будут представлены в соответствующих темах второго раздела данного курса.

Индексам цен в рыночной системе отводится важное место. Их используют для решения двух основных задач:

- измерения уровня инфляции;
- пересчета стоимостных показателей из фактических цен в сопоставимые.

В практике экономического анализа используют несколько видов индексов цен:

– *индекс потребительских цен*. Он оценивает динамику цен на рынке потребительских товаров и характеризует динамику общего уровня цен на товары и услуги, приобретаемые населением для непроизводственного потребления;

– *индекс цен производителей промышленной продукции производственно-технического назначения*. Этот индекс измеряет изменение цен на указанную продукцию на момент ее первой продажи;

– *индексы-дефляторы*. С помощью индексов-дефляторов производят пересчет стоимостных показателей системы национальных счетов (СНС) из фактических цен в сопоставимые. *Дефлятор* (от англ. *deflate* – **выражать в постоянных ценах**) – коэффициент перевода значения стоимостного показателя отчетного периода в стоимостные показатели базисного периода. Например, индекс-дефлятор валового внутреннего продукта (ВВП) представляет собой индекс цен, применяемый для корректировки номинального объема ВВП с учетом инфляции и получения на этой основе реального объема ВВП.

В практике статистического анализа используют *индивидуальные и общие индексы цен*.

*Индивидуальные индексы цен* характеризуют изменение цены по конкретному виду продукции ( $i_p$ ):

$$i_p = \frac{p_1}{p_0}, \quad (9.10)$$

где  $p_0, p_1$  – цена продукции в базисном и отчетном периодах соответственно.

*Общий индекс цен* оценивает изменение цены по нескольким видам продукции. В общем случае общий индекс цен строится на основе формулы Пааше:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}. \quad (9.11)$$

Это агрегатная форма общего индекса цен. Общий индекс цен в форме среднего гармонического из индивидуальных индексов имеет вид:

$$I_p = \frac{\sum q_1 p_1}{\sum \frac{q_1 p_1}{i_p}}. \quad (9.12)$$

Следует иметь в виду, что для средних индексов в качестве весов могут приниматься не только абсолютные показатели стоимости продукции (например,  $q_1 p_1$ ), но и относительные величины в виде долей или удельного веса отдельных групп продукции (товаров) в общем объеме производства, потребления, товарооборота.

Так, например, сводный индекс потребительских цен (ИПЦ), характеризующий изменение во времени общего уровня цен на потребительскую корзину, представляющую собой минимальный набор продуктов питания, непродовольственных товаров и услуг, необходимых для сохранения здоровья человека и обеспечения его жизнедеятельности, рассчитывается как средняя арифметическая из индексов средних цен на отдельные группы товаров и услуг. При этом в качестве весов выступает доля расходов населения на данные товары (услуги) в общей сумме потребительских расходов населения в базисном году:

$$d_{0_i} = \frac{q_{0_i} p_{0_i}}{\sum q_{0_i} p_{0_i}}. \quad (9.13)$$

С учетом вышесказанного, формула индекса потребительских цен ( $I_{пц}$ ) может быть записана следующим образом:

$$I_{пц} = \sum i_p d_{0_i}. \quad (9.14)$$

Если веса представлены абсолютными показателями, формула индекса потребительских цен имеет вид:

$$I_{пц} = \frac{\sum i_p q_0 p_0}{\sum q_0 p_0}. \quad (9.15)$$

Таким образом, индекс потребительских цен определяется на основе формулы Ласпейреса.

Общий индекс цен может быть определен и как средняя геометрическая из агрегатных индексов цен Ласпейреса и Пааше, на основе формулы

$$I_p = \sqrt{\frac{\sum q_1 p_1}{\sum q_1 p_0} \cdot \frac{\sum q_0 p_1}{\sum q_0 p_0}}. \quad (9.16)$$

Это так называемый идеальный индекс Фишера, рекомендуемый его автором в тех случаях, когда трудно отдать предпочтение весам  $q_0$  или  $q_1$ .

### 9.5. Система индексов переменного состава, постоянного состава и структурных сдвигов

Для качественных показателей может быть построена *система общих индексов переменного состава, постоянного состава и структурных сдвигов*. Расчет этой системы общих индексов выполняется при проведении анализа динамики качественных показателей по совокупности хозяйственных единиц (по совокупности цехов, совокупности предприятий).

Рассмотрим порядок расчета общих индексов переменного состава, общих индексов постоянного состава и влияния структурных сдвигов на примере показателя себестоимости продукции.

Индекс себестоимости переменного состава рассчитывается по формуле

$$I_{\bar{z}} = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_0}{\sum q_0} = \bar{z}_1 : \bar{z}_0, \quad (9.17)$$

где  $\bar{z}_1, \bar{z}_0$  – средняя себестоимость единицы продукции в отчетном и базисном периодах.

Индекс переменного состава иначе называют индексом средних величин, т.к. он оценивает изменение среднего значения индексируемой величины.

Индекс переменного состава зависит от двух факторов:

- от изменения индивидуальных уровней индексируемой величины. Влияние этого фактора измеряют с помощью *индекса постоянного состава*;

- от изменения структуры изучаемой совокупности. Влияние этого фактора оценивают с помощью *индекса влияния структурных сдвигов*.

Расчетная формула индекса себестоимости постоянного состава:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1} = \frac{\sum z_1 q_1}{\sum z_0 q_1}. \quad (9.18)$$



Индекс постоянного состава показывает, как в среднем изменился уровень индексируемой величины.

Индекс влияния структурных сдвигов рассчитывается по формуле

$$I_{cmp} = \frac{\sum z_0 q_1}{\sum q_1} \cdot \frac{\sum z_0 q_0}{\sum q_0}. \quad (9.19)$$

Индекс влияния структурных сдвигов оценивает изменение среднего уровня индексируемой величины за счет структурных сдвигов.

Рассмотренные индексы связаны между собой следующим образом:

$$I_z = I_z \cdot I_{cmp}. \quad (9.20)$$

## 9.6. Индексы пространственно-территориального сопоставления

В статистической практике часто возникает потребность в сопоставлении уровней экономического явления в пространстве: по странам, экономическим районам, областям, т.е. в исчислении территориальных индексов. При построении территориальных индексов приходится решать вопрос, какие веса использовать при их исчислении. Например, если стоит задача сравнить цены двух регионов (*A* и *B*), то можно построить два индекса:

$$I_{A/B} = \frac{\sum p_A q_A}{\sum p_B q_A}, \quad (9.21)$$

$$I_{B/A} = \frac{\sum p_B q_B}{\sum p_A q_B}, \quad (9.22)$$

где  $I_{A/B}$  – индекс, в котором в качестве базы сравнения применяются данные по региону *B*;

$I_{B/A}$  – индекс, используемый в качестве базы сравнения данных по региону *A*.

Эти формулы могут дать совершенно различное представление о соотношении уровней явления. Например, при расчете по формуле (9.21) значение признака будет ниже в регионе *A*, а по формуле (9.22) – в регионе *B*.

Рассмотрим данные, представленные в таблице 9.1. Рассчитав индексы по двум вышеприведенным формулам, получаем:

$$I_{A/B} = \frac{\sum p_A q_A}{\sum p_B q_A} = \frac{236400}{243000} = 0,973 \text{ или } 97,3 \%,$$

$$I_{B/A} = \frac{\sum p_B q_B}{\sum p_A q_B} = \frac{186000}{180800} = 1,029 \text{ или } 102,9 \%.$$

Индексы показывают, что при сравнении региона А с регионом Б цены ниже в регионе А на 2,7 %, а при сравнении региона Б с регионом А оказывается, что цены несколько выше в регионе Б. Таким образом, расчет индексов не позволяет определить, в каком регионе выше цены. Причина заключается в резком различии структуры продаж в отдельных регионах.

В теории и практике статистики предлагаются различные методы построения территориальных индексов, в т.ч. *метод стандартных весов*. Этот метод заключается в том, что значения индексируемой величины взвешиваются не по весам какого-то одного региона, а по весам области, экономического района, республики, в которых находятся сравниваемые регионы.

В нашем примере в качестве весов можно использовать количество продукции, проданной в регионах А и Б, т.е.:

$$I_p = \frac{\sum p_A (q_A + q_B)}{\sum p_B (q_A + q_B)}. \quad (9.23)$$

Определим значение индекса по данным табл. 9.1:

$$I_p = \frac{417200}{429000} = 0,9725 \text{ или } 97,25 \%.$$

Итак, цены региона А ниже, чем цены в региона Б, в среднем на 2,75 %.

## 9.7. Применение индексов в факторном анализе

Важнейшее значение индексов в статистическом анализе заключается в том, что они позволяют измерить влияние отдельных факторов на изменение того или иного результативного показателя. Такой анализ, позволяющий оценить изменение результативного показателя (в абсолютной или относительной форме) под влиянием каждого из факторов, называется факторным анализом.

Рассмотрим порядок проведения факторного анализа на основе индексного метода.

Если оценивают относительное изменение результативного показателя (У) под влиянием каждого из факторов (а, b, с), то используют следующую формулу:

$$I_Y = \frac{y_1}{y_0} = \frac{a_1 \cdot b_1 \cdot c_1}{a_0 \cdot b_0 \cdot c_0} = I_a \cdot I_b \cdot I_c, \quad (9.24)$$

где  $I_a, I_b, I_c$  – частные индексы, которые оценивают меру влияния соответствующего факторного признака на результативный (в относительной форме).

Таблица 9.1

**Цены и количество проданной продукции по двум регионам**  
(цифры условные)

Вид продукции	Регион <i>A</i>		Регион <i>B</i>		Товарооборот		$p_B q_A$	$p_A q_B$	Количество продукции, проданной в регионах <i>A</i> и <i>B</i> $p_A + q_B$	$p_A(q_A + q_B)$	$p_B(q_A + q_B)$
	Цена тыс.руб., $p_A$	Кол. проданного товара, $q_A$	Цена тыс.руб., $p_B$	Кол. проданного товара, $q_B$	Регион <i>A</i> $p_A q_A$	Регион <i>B</i> $p_B q_B$					
<i>A</i>	1	2	3	4	5=1·2	6=3·4	7=3·2	8=1·4	9=2+4	10=1·9	11=3·9
Продукция №1, ед.	95	1 700	100	1 200	161 500	120 000	170 000	114 000	2 900	275 500	290 000
Продукция №2, ед.	83	800	80	600	66 400	48 000	64 000	49 800	1 400	116 200	112 000
Продукция №3, ед.	17	500	18	1 000	8 500	18 000	9 000	17 000	1 500	25 500	27 000
Итого	–	3 000	–	2 800	236 400	186 000	243 000	180 800	–	417 200	429 000

Более сложной является задача оценки абсолютного изменения результативного показателя под влиянием каждого из показателей-факторов. Поясним это на основе простейшей модели:

$$Y = a \cdot b. \quad (9.25)$$

Рассматривая разность результативных показателей в двух сравниваемых периодах и выполнив необходимые элементарные подстановки, раскрывая скобки и приводя подобные члены, в итоге получаем:

$$\Delta y = Y_1 - Y_0 = a_1 b_1 - a_0 b_0 = (a_0 + \Delta a)(b_0 + \Delta b) - a_0 b_0 = \Delta a b_0 + \Delta b a_0 + \Delta a \Delta b. \quad (9.26)$$

Из приведенной формулы видно, что при анализе двухфакторной модели абсолютный прирост представлен тремя слагаемыми. Если пользоваться трехфакторной моделью мультипликативного вида  $Y = abc$ , то число слагаемых составит уже 7. В этом нетрудно убедиться, проделав аналогичные приведенным выше, элементарные преобразования. Трудности интерпретации результатов анализа в такой ситуации резко возрастают с увеличением числа факторов, а также в связи с тем, что знак произведения  $\Delta a \Delta b$  зависит не от абсолютных (по модулю) величин приращений, а от их знаков. Так, если факторы  $a$  и  $b$  в отчетном периоде по сравнению с базисным уменьшились по величине (отрицательные абсолютные приросты), произведение приростов окажется положительным, а если, допустим, фактор  $a$  уменьшается очень мало ( $\Delta a < 0$ ), а фактор  $b$  увеличивается на сколь угодно большую величину ( $\Delta b > 0$ ), произведение приростов всегда будет отрицательным.

Трудности такого рода и привели к тому, что на практике обычно слагаемое, представляющее собой остаточный член ( $\Delta a \Delta b$ ), присоединяют к какому-либо из двух первых слагаемых, руководствуясь при этом *экономическим смыслом показателей, содержанием поставленной задачи и эмпирическим правилом расположения факторов-сомножителей в исходной модели*. Целесообразным признается всегда ставить на первое место качественный фактор (характеризующий размер признака, приходящийся в среднем на одну единицу совокупности), а на второе – фактор количественный (характеризующий объем совокупности). Так, присоединяя остаточный член к первому слагаемому, получим окончательно следующую формулу, по которой определяют влияние абсолютного изменения каждого из факторов на абсолютное изменение результативного показателя:

$$a_1 b_1 - a_0 b_0 = \Delta a b_0 + \Delta a \Delta b + \Delta b a_0 = \Delta a (b_0 + \Delta b) + \Delta b a_0 = \Delta a b_1 + \Delta b a_0. \quad (9.27)$$

Исходя из этого, сформируем основные правила проведения факторного анализа на основе индексного метода.

### ***Правило 1***

В исходной модели на первое место ставят качественный факторный признак и порядок расположения факторов должен быть таким, чтобы последовательное попарное их перемножение давало показатель, имеющий экономический смысл.

### ***Правило 2***

Оценка влияния каждого данного факторного признака на результативный производится поочередно, при элиминировании влияния оставшихся факторов. Элиминирование влияния этих факторов производится путем фиксации их значений по сравниваемым периодам на одном и том же уровне.

### ***Правило 3***

Выбор периода, на уровне которого фиксируются значения факторов, влияние которых элиминируется, производят следующим образом:

- если оценивается влияние изменения качественного признака, то количественные признаки принимаются на уровне отчетного периода («1»);
- если оценивается влияние изменения количественного признака, качественные признаки принимаются на уровне базисного периода («0»).

### ***Правило 4***

Если исходная модель имеет более двух факторов, то значения факторов, влияние которых элиминируется, относят к тому или иному уровню (к «0» или «1») в зависимости от содержания показателя, полученного при их перемножении.

Если перемножение факторов, влияние которых элиминируется, не дает показателя, имеющего экономический смысл, то выбор уровня, на котором они фиксируются, делается для каждого фактора отдельно, исходя из их содержания.

Описанный метод называют *индексным методом определения влияния факторов*.

Покажем применение этого метода проведения факторного анализа на основе следующей трехфакторной мультипликативной модели:

$$\Phi = Z \cdot W \cdot \bar{T}, \quad (9.28)$$

где  $\Phi$  – фонд заработной платы за период;

$Z$  – зарплатоемкость единицы продукции ( $Z = \frac{\Phi}{Q}$ );

$W$  – производительность труда работников ( $W = \frac{Q}{T}$ );

$\bar{T}$  – среднесписочная численность работников;

$Q$  – объем продукции в стоимостном выражении.

Относительное изменение фонда заработной платы за счет изменения каждого из факторов определяется следующим расчетом:

$$I_F = \frac{\Phi_1}{\Phi_0} = \frac{Z_1 W_1 \bar{T}_1}{Z_0 W_0 \bar{T}_0} = I_Z \cdot I_W \cdot I_{\bar{T}}, \quad (9.29)$$

где  $I_\Phi$  – общее относительное изменение фонда заработной платы;

$I_Z$  – относительное изменение фонда заработной платы, обусловленное изменением зарплатоемкости продукции;

$I_W$  – относительное изменение фонда заработной платы, обусловленное изменением производительности труда;

$I_{\bar{T}}$  – относительное изменение фонда заработной платы, обусловленное изменением среднесписочной численности работников.

Общее абсолютное изменение фонда заработной платы определяется расчетом:

$$\Delta\Phi = \Phi_1 - \Phi_0. \quad (9.30)$$

Абсолютное изменение фонда заработной платы, обусловленное влиянием изменения зарплатоемкости продукции ( $\Delta\Phi_Z$ ), определяется по формуле

$$\Delta\Phi_Z = Z_1 W_1 \bar{T}_1 - Z_0 W_1 \bar{T}_1 = (Z_1 - Z_0) W_1 \bar{T}_1. \quad (9.31)$$

В этом случае элиминируется влияние факторов  $W$  и  $\bar{T}$ . Их перемножение ( $W\bar{T} = Q$ ) дает показатель (объем продукции), который по своему содержанию является количественным, поэтому оба эти фактора (и  $W$ , и  $\bar{T}$ ) относим к отчетному периоду.

Абсолютное изменение фонда заработной платы, обусловленное влиянием изменения производительности труда ( $\Delta\Phi_W$ ), оцениваем по формуле

$$\Delta\Phi_W = Z_0 W_1 \bar{T}_1 - Z_0 W_0 \bar{T}_1 = Z_0 (W_1 - W_0) \bar{T}_1. \quad (9.32)$$

В этом случае элиминируется влияние факторов  $Z$  и  $\bar{T}$ . Так как их перемножение  $Z \cdot \bar{T} = \frac{\Phi}{Q} \cdot \bar{T}$  не дает показателя, имеющего экономический

смысл, каждый из этих факторов относим к соответствующему уровню, исходя из их содержания. Так как  $Z$  – качественный факторный признак,

его принимаем на уровне базисного периода («0»),  $\bar{T}$  – количественный, его значение принимаем на уровне отчетного периода («1»).

Абсолютное изменение фонда заработной платы, обусловленное влиянием изменения среднесписочной численности работников ( $\Delta\Phi_T$ ), определяется по формуле

$$\Delta\Phi_T = Z_0W_0\bar{T}_1 - Z_0W_0\bar{T}_0 = Z_0W_0(\bar{T}_1 - \bar{T}_0). \quad (9.33)$$

В этом случае элиминируется влияние изменения зарплатоемкости продукции и производительности труда. Перемножение этих показателей  $Z \cdot W = \frac{\Phi}{Q} \cdot \frac{Q}{T} = \bar{Z}$  дает среднюю заработную плату одного работника, т.е. получаем показатель, имеющий экономический смысл. Так как по содержанию этот показатель является качественным,  $Z$  и  $W$  относим к базисному периоду («0»).

**Проверка:** общее абсолютное изменение фонда заработной платы должно быть равно сумме его изменений за счет каждого из рассмотренных факторов:

$$\Delta\Phi = \sum_{i=1}^n \Delta\Phi_i = \Delta\Phi_Z + \Delta\Phi_W + \Delta\Phi_T. \quad (9.34)$$

Расчеты, оценивающие влияние факторных признаков на результативный показатель, служат основой для формулировки выводов о силе и направленности влияния каждого из факторов. Это позволит определить направления улучшения анализируемого результативного показателя.

## 9.8. Решение типовых задач по теме

### Пример 9.1

По данным приведенной ниже таблицы определите:

- индивидуальные индексы себестоимости продукции и физического объема продукции;
- общие индексы физического объема, себестоимости продукции и затрат на производство:

Вид продукции	Себестоимость 1 т, тыс. руб.		Произведено, тыс. т		Индивидуальные индексы	
	базисный период	отчетный период	базисный период	отчетный период	себестоимости $i_z$	физического объема продукции $i_q$
А	13,2	12,1	1,27	1,31	...	...
Б	23,9	24,6	4,16	2,96	...	...

Решение:

1. Определяем индивидуальные индексы себестоимости единицы продукции:

$$i_z^A = \frac{z_1^A}{z_0^A} = \frac{12,1}{13,2} = 0,92 \text{ или } 92 \%,$$

$$i_z^B = \frac{z_1^B}{z_0^B} = \frac{24,6}{23,9} = 1,03 \text{ или } 103 \%.$$

2. Определяем индивидуальные индексы физического объема продукции:

$$i_q^A = \frac{q_1^A}{q_0^A} = \frac{1,31}{1,27} = 1,03 \text{ (103 \%)},$$

$$i_q^B = \frac{q_1^B}{q_0^B} = \frac{2,96}{4,16} = 0,71 \text{ (71 \%)}.$$

3. Определяем общий индекс физического объема продукции (т.е. по обоим видам продукции):

$$I_q = \frac{\sum q_1 z_0}{\sum q_0 z_0} = \frac{1,31 \cdot 13,2 + 2,96 \cdot 23,9}{1,27 \cdot 13,2 + 4,16 \cdot 23,9} = 0,758 \text{ или } 75,8 \%.$$

В отчетном периоде по сравнению с базисным физический объем продукции снизился на 24,2 %.

4. Определяем общий индекс себестоимости продукции (т.е. по обоим видам продукции вместе):

$$I_z = \frac{\sum z_1 q_1}{\sum z_0 q_1} = \frac{12,1 \cdot 1,31 + 24,6 \cdot 2,96}{13,2 \cdot 1,31 + 23,9 \cdot 2,96} = 1,007 \text{ или } 100,7 \%.$$

В отчетном периоде по сравнению с базисным себестоимость продукции в среднем возросла на 0,7 %.

5. Определяем общий индекс затрат на производство.

*Первый вариант.* Расчет выполняется на основе общего индекса затрат на производство:

$$I_{zq} = \frac{\sum z_1 q_1}{\sum z_0 q_0} = \frac{12,1 \cdot 1,31 + 24,6 \cdot 2,96}{13,2 \cdot 1,27 + 23,9 \cdot 4,16} = 0,763 \text{ (76,3 \%)}.$$

*Второй вариант.* Расчет выполняется на основе формулы, отражающей связь между индексами:

$$I_{zq} = I_z \cdot I_q = 1,007 \cdot 0,758 = 0,763 \text{ (76,3 \%)}.$$

Затраты на производство продукции в отчетном периоде по сравнению с базисным сократились на 23,7 %. Основной фактор снижения затрат – уменьшение физического объема продукции.



### Пример 9.2

По данным таблицы рассчитайте общий индекс физического объема реализации:

Вид товара	Объем реализации в базисном периоде, тыс. руб.	Темп прироста физического объема в текущем периоде по сравнению с базисным, %	Расчетные данные	
			$i_q$	$i_q p_0 q_0$
Мандарины	46 000	-6,4	0,936	43056
Грейпфруты	27 000	-8,2	0,918	24056
Апельсины	51 000	+1,3	1,013	51663
Итого	124 000	–	–	119505

Решение:

Так как нет данных об абсолютных показателях объема продукции и известны индивидуальные индексы физического объема реализации, рассчитаем общий индекс физического объема реализации на основе среднего арифметического индекса:

$$I_q = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0} = \frac{119505}{124000} = 0,964 \text{ (96,4 \%)}.$$

Физический объем реализации в отчетном периоде по сравнению с базисным уменьшился на 3,6 %.

### Пример 9.3

По данным таблицы рассчитайте:

а) общие индексы: цены, физического объема продаж и стоимости продукции;

б) величину изменения расходов покупателей в октябре по сравнению с сентябрем вследствие изменения цен

Вид продукции	сентябрь		октябрь	
	цена за единицу, тыс. руб.	продано, ед.	цена за единицу, тыс. руб.	продано, ед.
А	18	2630	19	2410
Б	15	880	15	920
В	22	1450	24	1230

Решение:

1. Определяем общий индекс цен:

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{19 \cdot 2410 + 15 \cdot 920 + 24 \cdot 1230}{18 \cdot 2410 + 15 \cdot 920 + 22 \cdot 1230} = 1,058 \text{ (105,8 \%)}.$$

Рост цен в октябре по сравнению с сентябрем составил 105,8 %.

2. Определяем общий индекс физического объема реализации:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{2410 \cdot 18 + 920 \cdot 15 + 1230 \cdot 22}{2630 \cdot 18 + 880 \cdot 15 + 1450 \cdot 22} = 0,911 \text{ (91,1 \%)}.$$

В октябре по сравнению с сентябрем физический объем реализации сократился на 8,9 %.

3. Общий индекс стоимости продукции:

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{19 \cdot 2410 + 15 \cdot 920 + 24 \cdot 1230}{18 \cdot 2630 + 15 \cdot 880 + 22 \cdot 1450} = 0,964 \text{ (96,4 \%)}.$$

В октябре по сравнению с сентябрем стоимость продукции уменьшилась на 3,6 %. Основным фактором уменьшения стоимости продукции явился фактор уменьшения физического объема реализации.

4. Определяем сумму изменения расходов покупателей в октябре по сравнению с сентябрем вследствие изменения цен:

$$\Delta F = \sum q_1 p_1 - \sum q_1 p_0 = (19 \cdot 2410 + 15 \cdot 920 + 24 \cdot 1230) - (18 \cdot 2410 + 15 \cdot 920 + 22 \cdot 1230) = 89110 - 84240 = 4870 \text{ (тыс. руб.)}.$$

### Пример 9.4

Имеются следующие данные о выпуске продукции «А» по двум заводам.

Завод	Базисный период		Отчетный период	
	произведено продукции, тыс. шт.	себестоимость единицы, тыс. руб.	произведено продукции, тыс. шт.	себестоимость единицы, тыс. руб.
№ 1	60	24	80	20
№ 2	60	20	120	18

Определите: общие индексы себестоимости переменного состава, постоянного состава и структурных сдвигов.

Решение:

1. Определяем общий индекс себестоимости переменного состава:

$$I_{\bar{z}} = \frac{\sum z_1 q_1}{\sum q_1} \cdot \frac{\sum z_0 q_0}{\sum q_0} = \frac{20 \cdot 80 + 18 \cdot 120}{80 + 120} \cdot \frac{24 \cdot 60 + 20 \cdot 60}{60 + 60} = 18,8 : 22 = 0,855 \text{ (85,5 \%)}.$$

Средняя себестоимость в отчетном периоде по сравнению с базисным снизилась на 14,5 %.

2. Определяем индекс себестоимости постоянного состава:

$$I_z = \frac{\sum z_1 q_1}{\sum q_1} \cdot \frac{\sum z_0 q_1}{\sum q_1} = 18,8 \cdot \frac{24 \cdot 80 + 20 \cdot 120}{80 + 120} = 18,8 : 21,6 = 0,87 \text{ (87 \%)}.$$

Индивидуальная себестоимость единицы продукции в отчетном периоде по сравнению с базисным в среднем снизилась на 13 %.

3. Определяем индекс влияния структурных сдвигов:

$$I_{стр} = \frac{\sum z_0 q_1}{\sum q_1} \cdot \frac{\sum z_0 q_0}{\sum q_0} = 21,6 : 22 = 0,982 \text{ (98,2 \%)}.$$

Средняя себестоимость за счет структурных сдвигов в объеме производства снизилась в отчетном периоде по сравнению с базисным на 1,8 %.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ ПО ТЕМЕ**

### **Вопросы для самоподготовки и контроля знаний**

1. Что такое индекс как статистический показатель?
2. Какие аналитические задачи решаются на основе применения индексов?
3. Какие виды индексов выделяют в зависимости от степени охвата элементов совокупности?
4. Какие индексы называются индивидуальными?
5. Что характеризуют общие индексы?
6. Что такое индексируемая величина?
7. Какие индексы называют цепными и какие базисными?
8. Какая связь существует между цепными и базисными индексами?
9. Какие существуют формы построения общих индексов?
10. Укажите правила построения агрегатных индексов количественных и качественных показателей.
11. В чем состоит различие формул Пааше и Ласпейреса?
12. Какие виды средних индексов используются в статистической практике?
13. Какой вид средних величин используется при построении идеального индекса Фишера?
14. Какая формула, Пааше или Ласпейреса, лежит в основе расчета индекса потребительских цен?
15. В чем суть индексов переменного и фиксированного состава?
16. Что характеризует индекс структурных сдвигов?
17. Как связаны между собой общие индексы переменного состава, постоянного состава и структурных сдвигов?
18. Назовите виды индексов качественных показателей.

19. Что характеризует разность числителя и знаменателя общего индекса физического объема, общего индекса себестоимости продукции?
20. Какова роль индексов цен в экономике?
21. Что такое индекс-дефлятор?
22. Какие проблемы возникают при построении территориальных индексов и как они решаются?

### Тест 9.1

1. В зависимости от степени охвата единиц совокупности различают ...

- а) агрегатные индексы;
- б) индивидуальные индексы;
- в) общие индексы;
- г) средние из индивидуальных.

2. К индексам качественных показателей относят ...

- а) индекс себестоимости продукции;
- б) индекс физического объема продукции;
- в) индекс цены.

3. Общий индекс стоимости продукции исчисляется по формуле ...

- а)  $\frac{\sum p_1 q_1}{\sum p_0 q_1}$ ;
- б)  $\frac{\sum p_1 q_1}{\sum p_0 q_0}$ ;
- в)  $\frac{\sum q_1 p_1}{\sum q_0 p_1}$ .

4. Приведенная система общих индексов стоимости продукции

$$\frac{\sum p_1 q_1}{\sum p_0 q_0}; \quad \frac{\sum p_2 q_2}{\sum p_1 q_1}; \quad \frac{\sum p_n q_n}{\sum p_{n-1} q_{n-1}}$$

является ...

- а) цепной;
- б) базисной.

5. Данная формула  $\frac{\sum z_1 q_1}{\sum q_1} : \frac{\sum z_0 q_1}{\sum q_1}$  представляет ...

- а) индекс переменного состава;
- б) индекс постоянного состава;
- в) индекс структурных сдвигов.

6. Индекс цен Ласпейреса определяется по формуле ...

а)  $\frac{\sum p_1 q_1}{\sum p_0 q_1}$ ;

б)  $\frac{\sum p_1 q_0}{\sum p_0 q_0}$ ;

в)  $\sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \cdot \frac{\sum p_1 q_1}{\sum p_0 q_1}}$ .

7. Индекс количества продукции, произведенной в единицу времени, рассчитывается по формуле ...

а)  $\frac{q_1}{T_1} : \frac{q_0}{T_0}$ ;

б)  $\frac{t_0}{t_1}$ ;

в)  $\frac{q_1 P}{T_1} : \frac{q_0 P}{T_0}$ ;

г)  $\frac{t_1 q_1}{t_0 q_0}$ .

8. Система базисных индексов физического объема продукции с постоянными весами имеет следующий вид: ...

а)  $\frac{\sum q_1 P_0}{\sum q_0 P_0}$ ;  $\frac{\sum q_2 P_0}{\sum q_1 P_0}$ ; ...  $\frac{\sum q_n P_0}{\sum q_{n-1} P_0}$ ;

б)  $\frac{\sum q_1 P_0}{\sum q_0 P_0}$ ;  $\frac{\sum q_2 P_0}{\sum q_0 P_0}$ ; ...  $\frac{\sum q_n P_0}{\sum q_0 P_0}$ ;

в)  $\frac{\sum q_1 P_0}{\sum q_0 P_0}$ ;  $\frac{\sum q_2 P_2}{\sum q_0 P_0}$ ; ...  $\frac{\sum q_n P_n}{\sum q_0 P_0}$ .

9. Если себестоимость снизилась на 10 %, а количество продукции увеличилось на 5 %, то индекс затрат на производство будет равен ...

а) 105 %;

б) 94,5 %;

в) 95,0 %.

10. Если индекс переменного состава равен 125 %, а индекс структурных сдвигов 110 %, то индекс фиксированного состава равен ...

а) 100 %;

б) 119 %;

в) 115 %.

## Задания и задачи по теме

### Задача 9.1

Имеются следующие данные за два периода о ценах и объемах реализации трех видов товаров по одному из торговых предприятий:

Вид товара	Базисный период		Отчетный период	
	Цена, тыс. руб. $p_0$	Продано товаров, шт. $q_0$	Цена, тыс. руб. $p_1$	Продано товаров, шт. $q_1$
А	45	2500	63	1700
Б	27	830	35	2300
В	12	610	14	1000

Определите:

- 1) индивидуальные индексы цен;
- 2) индивидуальные индексы физического объема реализации товаров;
- 3) общий индекс цен;
- 4) общий индекс физического объема реализации товаров;
- 5) общий индекс товарооборота (стоимости товаров).

### Задача 9.2

Имеются следующие данные об изменении физического объема розничного товарооборота в регионе:

Группа товаров	Индекс физического объема товарооборота по группам товаров в 2010 г. по сравнению с 2009 г., %	Структура товарооборота в 2009 г., %
Продовольственные	102,3	42
Непродовольственные	104,1	58

Определите общий индекс физического объема розничного товарооборота в регионе в 2010 г. по сравнению с 2009 г.

### Задача 9.3

Динамика физического объема производства промышленной продукции на предприятии характеризуется следующими данными:

Год	2005	2006	2007	2008	2009	2010
Цепной индекс физического объема производства, %	94,7	101,6	102,1	102,8	103,2	104,4

Определите:

- 1) индекс физического объема промышленного производства в 2010 г. по сравнению с 2005 г.;
- 2) среднегодовой темп изменения физического объема производства за указанный период.

#### Задача 9.4

Имеются данные об общем индексе цен на продукцию промышленного предприятия за период 2007 – 2010 гг.:

Год	2007	2008	2009	2010
Цепной индекс цен, %	107,8	108,2	106,5	108,6

Определите изменение цен на продукцию предприятия за весь период и их среднегодовое изменение.

#### Задача 9.5

Имеются следующие данные по предприятию за два года:

Вид продукции	2009		2010	
	Себестоимость единицы продукции, тыс. руб.	Количество произведенной продукции, тыс. шт.	Себестоимость единицы продукции, тыс. руб.	Количество произведенной продукции, тыс. шт.
<i>A</i>	3,0	22	3,5	25
<i>B</i>	7,0	13	8,3	10
<i>B</i>	9,0	42	8,6	48

Определите:

- 1) общий индекс себестоимости продукции;
- 2) общий индекс изменения затрат на производство;
- 3) абсолютное изменение затрат на производство общее и в том числе, обусловленное изменением себестоимости единицы продукции отдельных видов и изменением физического объема произведенной продукции.

#### Задача 9.6

Имеются следующие данные по промышленному предприятию:

Изделие	Общие затраты на производство в 2010 г., млн руб.	Темп прироста себестоимости изделия в 2010 г. по сравнению с 2009 г., %
<i>A</i>	691,2	+8,0
<i>B</i>	1267,2	+5,6
<i>B</i>	2214,0	-2,5

Определите:

- 1) общий индекс себестоимости продукции в 2010 г. по сравнению с 2009 г.,
- 2) размер экономии или дополнительных затрат, вызванных изменением себестоимости продукции.

### Задача 9.7

Деятельность торгового предприятия за два месяца 2010 г. характеризуется следующими данными:

Товар	Товарооборот, тыс. руб.	
	март	апрель
Кофе растворимый	340	305
Кофе молотый	1650	1710
Чай	970	1054
Какао	800	840

Определите общий индекс физического объема реализации с учетом того, что в апреле предприятие повысило все цены в среднем на 8 %.

### Задача 9.8

Имеются следующие данные о реализации товара А на рынках города:

Рынок	Январь		Февраль	
	Цена за кг, тыс. руб.	Продано, тыс. кг	Цена за кг, тыс. руб.	Продано, тыс. кг
1	8,5	2,45	8,9	2,52
2	9,0	1,87	9,0	1,92
3	8,9	2,36	9,2	2,36

Определите:

- 1) общий индекс цен переменного состава;
- 2) общий индекс цен фиксированного состава;
- 3) общий индекс цен структурных сдвигов.

По результатам расчетов сделайте выводы, объяснив значение каждого индекса и связи между ними.

### Задача 9.9

Строительно-производственная деятельность двух строительных предприятий города характеризуется следующими данными:



Домостроительное предприятие	Построено жилья, тыс. м <sup>2</sup>		Себестоимость 1 м <sup>2</sup> , млн руб.	
	в базисном периоде	в отчетном периоде	в базисном периоде	в отчетном периоде
Предприятие 1	53	68	4,34	4,71
Предприятие 2	102	84	3,52	3,96

Определите:

- 1) общий индекс себестоимости переменного состава;
- 2) общий индекс себестоимости постоянного состава;
- 3) общий индекс себестоимости структурных сдвигов.

По результатам расчетов сделайте выводы.

### Задача 9.10

На основе данных, представленных в нижеследующей таблице, выполните следующие расчеты:

- 1) определите общий индекс физического объема продукции;
- 2) индивидуальный индекс производительности труда по каждому виду продукции;
- 3) общий индекс производительности труда по предприятию в целом;
- 4) общее абсолютное изменение затрат рабочего времени на производство продукции и в т.ч. за счет изменения: а) производительности труда; б) физического объема продукции.

Вид продукции	Выпуск продукции, тыс. ед.		Затраты времени на единицу продукции, чел.-час.	
	базисный период	отчетный период	базисный период	отчетный период
А	50	56	4,2	4,0
Б	200	150	3,0	3,0
В	100	80	2,0	1,9

### Задание 9.11

По данным периодической печати приведите примеры экономических индексов. Укажите, какие из них являются индексами качественных и количественных показателей.

### Задание 9.12

Сформируйте месячную потребительскую корзину для вашей семьи, указав наименование товаров и услуг, объем потребления по каждому виду товаров и услуг, цену товаров на начало и конец месяца. Рассчитайте индивидуальные индексы цен товаров и услуг, а также общий индекс цен. Сделайте выводы.

## Тема 10

# СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СВЯЗИ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

- 10.1. Виды и характеристика взаимосвязей в статистике.
- 10.2. Статистические методы выявления корреляционной связи между двумя признаками.
  - 10.2.1. Элементарные методы анализа.
  - 10.2.2. Дисперсионный анализ.
  - 10.2.3. Аналитические возможности и условия применения корреляционно-регрессионного анализа (КРА).
- 10.3. Измерение степени тесноты корреляционной связи между двумя признаками.
  - 10.3.1. Коэффициент корреляции знаков.
  - 10.3.2. Линейный коэффициент корреляции.
  - 10.3.3. Корреляционное отношение.
- 10.4. Расчет и интерпретация однофакторных уравнений регрессии.
- 10.5. Понятие о множественной корреляции и факторном анализе.
- 10.6. Практические примеры анализа корреляционной связи.

Изучив данную тему, студенты должны:

- понимать сущность корреляционной связи;
- знать статистические методы выявления корреляционной связи;
- уметь оценивать степень тесноты корреляционной связи между признаками;
- уметь проводить регрессионный анализ и интерпретировать его результаты.

### 10.1. Виды и характеристика взаимосвязей в статистике

Исследуя социально-экономические явления, важно не только определить уровень их развития, но и объяснить, под влиянием каких причин сформировался такой уровень. Изучение социально-экономических явлений и процессов показывает, что изменчивость каждого анализируемого признака зависит от вариации других, связанных с ним признаков.

Раздел статистики, который выявляет и измеряет связи между признаками, называется *теорией корреляции*.

Основу изучения связей между признаками составляет *анализ причинно-следственных отношений* между ними. Причинно-следственные отношения проявляются в том, что одни признаки выступают в качестве причин (факторов), обуславливающих изменение других, связанных с ними признаков.

Признаки, обуславливающие изменение других, связанных с ними признаков, называются *факторными*, или просто *факторами*.

Признаки, изменяющиеся под действием факторных признаков, называют *результативными*.

Различают два типа связей между различными явлениями и их признаками:

- *функциональные*, или жестко детерминированные;
- *статистические*, или стохастически детерминированные.

Строго определить различие этих типов связи можно, если дать им математическое описание. Для простоты понимания будем говорить о связи двух явлений или двух признаков, один из которых факторный, другой – результативный.

***Функциональной называют такую связь, при которой каждому значению факторного признака (x) обязательно соответствует одно<sup>1</sup> строго определенное значение результативного признака (y).***

При функциональной связи изменение результативного признака полностью обусловлено действием факторного признака (x), поэтому функциональную связь иначе называют *полной*. Математически это выражается следующим образом:

$$y = f(x). \quad (10.1)$$

Характерной особенностью функциональной связи является то, что она проявляется одинаково для каждой единицы изучаемой совокупности, поэтому ее можно распространить как на каждую единицу, так и на всю совокупность в целом. В экономике примером функциональной связи может служить зависимость между объемом производства продукции с одной стороны, и уровнем производительности труда и среднесписочной численностью работников, с другой стороны.

*Статистическая (стохастически детерминированная) связь – это вид причинной связи, проявляющейся не в каждой единице совокупности, а в среднем для совокупности в целом.*

Корреляционная связь представляет частный случай статистической связи. *Корреляционной называют такую связь, при которой разным значениям факторного признака соответствуют различные средние значения результативного.<sup>2</sup>*

---

<sup>1</sup> В некоторых случаях при функциональной связи одному значению факторного признака может соответствовать несколько строго определенных значений результативного, например, если  $y = \sqrt{x}$ .

<sup>2</sup> Если же с изменением значений факторного признака x среднее значение результативного признака не изменяется закономерным образом, но с определенной закономерностью изменяется другая статистическая характеристика (показатели вариации, асимметрии, эксцесса), то связь, являясь статистической, не будет корреляционной.

Как следует из сказанного выше, корреляционная связь характеризуется тем, что с изменением значения факторного признака ( $x$ ) закономерным образом изменяется среднее значение результативного признака ( $y$ ). При этом каждому конкретному значению ( $x$ ) может соответствовать распределение значений  $y$  (с различными вероятностями). Эта особенность корреляционной связи обусловлена тем, что результативный признак формируется не только под влиянием выделенных для анализа факторных признаков, но и под влиянием других, неучтенных нами, причин. При этом состав этих причин, направление и сила их воздействия на отдельные единицы совокупности могут различаться. Поэтому корреляционную связь называют *неполной*. Математически это можно представить следующим образом:

$$y = f(x) + \varepsilon, \quad (10.2)$$

где  $f(x)$  – уровень результативного признака, обусловленный влиянием факторного признака;

$\varepsilon$  – уровень результативного признака, обусловленный влиянием неучтенных факторов. В состав таких факторов входят второстепенные, случайные, неизвестные, не имеющие количественного выражения.

Резюмируя характеристику сущности корреляционной связи, отметим характерные ее черты:

- каждому значению факторного признака может соответствовать не одно значение результативного признака, а распределение его значений;
- с изменением факторного признака изменяется среднее значение результативного признака;
- корреляционная связь является неполной;
- она проявляется не в каждом единичном случае, а только в массе явлений.

Корреляционная связь между признаками может возникнуть различными путями. Рассмотрим три наиболее типичных ситуации объяснения связей между явлениями.

**Первая ситуация** характеризуется наличием четко определяемой причинной обусловленностью вариации результативного признака вариацией факторного признака. Например, балл оценки плодородия почв рассматривается как независимая переменная ( $x$ ), а урожайность выступает как зависимая переменная ( $y$ ).

**Вторая ситуация** ориентирует на правильную интерпретацию изучения корреляционной связи между двумя следствиями одной общей причины. Такую корреляцию нельзя рассматривать как связь причины и след-

ствия. Классическим примером такого случая является отмечаемая прямая корреляция между числом пожарных команд в городах и размером убытков от пожаров. На самом деле оба признака являются следствием общей причины – размера города. Вполне объяснимо, что в крупных городах больше пожарных частей, но больше и пожаров, и убытков от них, чем в мелких городах.

**Третья ситуация** связана с изучением корреляции между признаками, каждый из которых может быть и причиной и следствием. Такова, например, корреляция между уровнем производительности труда рабочих и уровнем их оплаты. В такой системе признаков допустимы две постановки задачи (в зависимости от целей исследования): каждый признак может выступать и в роли независимой переменной ( $x$ ) и в качестве зависимой переменной ( $y$ ).

Охарактеризуем корреляционную связь на основе классификации по ряду признаков.

*По направлению* корреляционные связи бывают *прямыми*, когда зависимая переменная растет с увеличением факторного признака, и *обратными*, при которых рост факторного признака сопровождается уменьшением результативного.

*По форме и аналитическому выражению* корреляционные связи бывают *линейными* и *нелинейными*. Если корреляционная связь между явлениями может быть приближенно выражена уравнением прямой, то ее называют *линейной связью*; если же она выражается уравнением какой-либо кривой линии (параболы, гиперболы, степенной и др.), то *нелинейной* или *криволинейной*.

*В зависимости от количества взаимодействующих признаков* различают *парную* и *множественную корреляцию*. Парная корреляция изучает связь между двумя признаками. Множественная корреляция изучает зависимость результативного признака от двух и более факторных признаков.

*По степени тесноты* различают *слабые* и *сильные корреляционные связи*. Эта важная характеристика корреляционных связей измеряется с помощью специальных статистических показателей (линейного коэффициента корреляции, корреляционного отношения, ранговых коэффициентов корреляции, коэффициента ассоциации, коэффициента контингенции и др.). В статистике используют специальные шкалы, которые в зависимости от значений показателей тесноты связи, дают качественную оценку силы связи. Одна из таких шкал приводится в табл. 10.1.

**Критерии оценки тесноты связи**

Величина коэффициента корреляции	Характер связи
до $ 0,3 $	практически отсутствует
$ 0,3  -  0,5 $	слабая
$ 0,5  -  0,7 $	умеренная
$ 0,7  -  0,99 $	сильная

## 10.2. Статистические методы выявления корреляционной связи между двумя признаками

Статистический анализ связей между признаками начинается с ответа на вопрос, существует ли между признаками корреляционная зависимость. Для этого статистика использует следующие методы:

- *элементарные приемы анализа;*
- *дисперсионный анализ;*
- *корреляционно-регрессионный анализ.*

### 10.2.1. Элементарные методы анализа

В статистике не всегда требуется количественное измерение корреляционной связи. Часто достаточно выявить лишь ее наличие (отсутствие), определить направление и предположить форму аналитического выражения этой связи. Для ответа на эти вопросы можно ограничиться применением элементарных методов:

- сопоставление двух параллельных рядов значений факторного и результативного признаков;
- графическое представление данных с помощью поля корреляции и эмпирической линии регрессии;
- построение аналитической группировки.

*Метод сопоставления двух параллельных рядов* предполагает сопоставление ряда значений факторного признака и соответствующих ему значений результативного признака. Значения факторного признака располагают в возрастающем порядке и затем прослеживают направление изменения величины результативного.

В случаях, когда возрастание величины факторного признака влечет за собой возрастание и величины результативного признака, предполагают наличие прямой корреляционной связи. Если же с увеличением факторного признака величина результативного имеет тенденцию к уменьшению, – обратную корреляционную связь между признаками.

При большом объеме изучаемой совокупности восприятие таких параллельных рядов затруднительно и для установления наличия корреляци-

онной связи целесообразнее воспользоваться корреляционными таблицами или аналитической группировкой.

*Графические методы* выявления корреляционной связи основаны на использовании *поля корреляции и эмпирической линии регрессии*.

*Поле корреляции* представляет собой точечный график в системе координат, который строится на основе индивидуальных значений признака-фактора и соответствующих им значений результативного признака.

Применение поля корреляции позволяет:

*во-первых*, сделать предположение о направлении связи на основе анализа расположения точек в системе координат. Если точки расположены слева-направо-вверх – связь прямая, т.к. с ростом причины увеличивается следствие. Если точки расположены слева-направо-вниз, можно предполагать наличие обратной связи;

*во-вторых*, сделать предположения о степени тесноты корреляционной связи. Чем плотнее располагаются точки вокруг определенной линии, тем связь теснее.

На основе поля корреляции в большинстве случаев сложно делать предположения о форме связи. Чем больше рассеивание точек на графике, тем сложнее судить о форме связи. На рис. 10.1 представлены варианты интерпретации корреляционной связи на основе поля корреляции.

*Эмпирическая линия регрессии* строится в системе координат на основе сгруппированных данных: каждому значению факторного признака ( $x$ ) соответствует среднее значение ( $y$ ). На основе эмпирической линии регрессии можно установить направление связи (прямая, обратная) и предполагать форму связи. Если эмпирическая линия связи по своему виду приближается к прямой линии, то можно предполагать наличие прямолинейной корреляционной связи между признаками. Если же имеется тенденция неравномерного изменения значений результативного признака, и эмпирическая линия регрессии будет приближаться к какой-либо кривой, то это может быть связано с наличием криволинейной корреляционной связи.

Эмпирическая линия регрессии не дает представление о тесноте связи.

Использование *аналитической группировки* в анализе корреляционной связи предполагает, что вся совокупность разбивается на группы в зависимости от величины признака-фактора, и по каждой группе вычисляются средние значения результативного признака. Проследив изменение средних значений результативного признака по группам, можно сделать вывод о наличии и направлении корреляционной связи. Более подробно методика проведения аналитической группировки рассматривалась в теме 3.

Благодаря группировке ряды уплотняются, и связь проявляется более отчетливо.

### Характеристика тесноты корреляционной связи

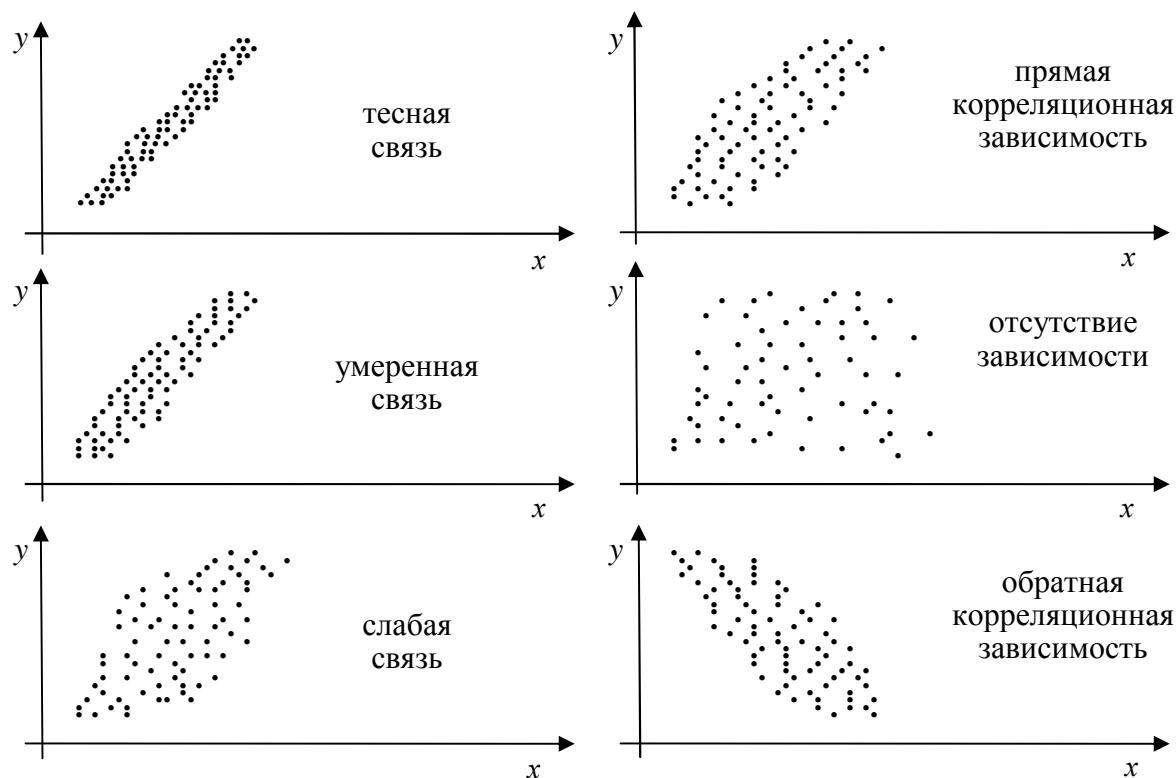


Рис. 10.1. Анализ корреляционной связи на основе поля корреляции

### 10.2.2. Дисперсионный анализ

Применение дисперсионного анализа в изучении корреляционных связей позволяет:

- установить, в какой мере вариация результативного признака обусловлена влиянием факторного признака;
- измерить степень тесноты связи между факторным и результативным признаками;
- проверить с помощью  $F$ -критерия существенность влияния факторного признака на результативный;
- отобрать факторы, наиболее существенно влияющие на результативный признак.

### 10.2.3. Аналитические возможности и условия применения корреляционно-регрессионного анализа (КРА)

Корреляционно-регрессионный анализ включает два направления:

- **корреляционный анализ**, в задачи которого входит оценка тесноты связи между анализируемыми признаками и выявление факторов, оказывающих наиболее сильное влияние на результативный признак;



– *регрессионный анализ*, т.е. установление формы корреляционной связи, ее аналитическое выражение в виде уравнения регрессии и проверка адекватности (надежности) последнего.

Следует отметить, что ряд авторов рассматривают корреляционный анализ широко, включая в него, наряду с собственно корреляционным анализом, и регрессионный анализ.

КРА достаточно сложный метод статистического анализа, поэтому его проведение включает несколько этапов:

- постановка задачи, анализ причинно-следственных отношений между признаками и определение факторных и результативного признаков;
- сбор статистических данных и проверка их пригодности для проведения КРА;
- предварительное изучение связи между признаками с помощью элементарных методов;
- измерение тесноты связи между признаками;
- аналитическое выражение корреляционной связи с помощью уравнения регрессии;
- оценка надежности уравнения регрессии, его экономическая интерпретация и определение сферы и условий практического применения.

Чтобы результаты корреляционного анализа достаточно точно характеризовали взаимосвязь между признаками, необходимо соблюдать следующие требования к формированию исходных статистических данных для анализа.

*Во-первых, корреляционный анализ должен проводиться для достаточно больших по объему совокупностей (выборок).* Обычно считают, что число наблюдений должно быть не менее чем в 5 – 6 раз, а лучше в 10 раз больше числа факторов, включенных в уравнение регрессии. Формирование достаточных по объему совокупностей для анализа обеспечит проявление закона больших чисел, и таким образом закономерности связи между признаками проявятся более отчетливо.

*Во-вторых, должна быть обеспечена однородность совокупности по величине факторного и результативного признаков.* Нарушение этого требования сделает результаты КРА ненадежными. Напомним, что однородность совокупности может быть оценена с помощью коэффициента вариации.

*В-третьих, распределение единиц совокупности по результативному и факторным признакам должно подчиняться нормальному закону распределения вероятностей.* Это условие связано с применением метода наименьших квадратов (МНК) при расчете параметров уравнения регрес-

сии: только при нормальном распределении метод наименьших квадратов дает оценку параметров, отвечающую принципам максимального правдоподобия. На практике это требование чаще всего выполняется приближенно.

*В-четвертых, число включаемых в исследование факторов должно быть ограничено.* Отобрать следует независимые друг от друга факторные признаки, наиболее существенно влияющие на результативный признак. Факторные признаки, функционально связанные с результативным, должны быть исключены.

*Пятое требование заключается в том, что при построении уравнений корреляционной связи факторные признаки должны иметь количественное выражение, иначе они не могут быть включены в модель.*

Построение корреляционно-регрессионных моделей, какими бы сложными они не были, само по себе не вскрывает полностью всех причинно-следственных связей. Основой их адекватности во всех случаях является содержательный анализ, основанный на учете специфики и особенностей сущности исследуемых социально-экономических явлений и процессов.

### **10.3. Измерение степени тесноты корреляционной связи между двумя признаками**

*Показатели степени тесноты связи характеризуют зависимость вариации результативного признака от вариации факторного признака.*

Показатели тесноты корреляционной связи позволяют решать следующие исследовательские задачи:

- установить целесообразность изучения данной связи между признаками и возможность практического применения результатов ее количественного выражения;
- сопоставляя показатели тесноты связи для различных ситуаций, можно судить о степени и природе различий в ее проявлении для конкретных условий;
- сопоставляя показатели тесноты связи результативного признака с различными факторами, можно отобрать те, которые в данных конкретных условиях являются наиболее значимыми и в решающей степени влияют на изменение величины результативного признака.

Рассмотрим порядок расчета и применения в корреляционном анализе основных показателей тесноты связи.

### 10.3.1. Коэффициент корреляции знаков

Коэффициент корреляции знаков (иначе его называют коэффициентом Фехнера) основан на оценке степени согласованности направлений отклонений индивидуальных значений факторного и результативного признаков от их средних величин:

$$K_{\phi} = \frac{n_c - n_{nc}}{n_c + n_{nc}}, \quad (10.3)$$

где  $n_c$  – число совпадений знаков отклонений индивидуальных значений факторного и результативного признаков от их средних значений;

$n_{nc}$  – число несовпадений знаков отклонений индивидуальных значений факторного и результативного признаков от их средних значений.

*Коэффициент корреляции знаков ( $K_{\phi}$ )* изменяется в пределах от -1 до +1. Знак при численном значении коэффициента Фехнера указывает направление связи: если  $K_{\phi} > 0$ , предполагают наличие прямой связи, при  $K_{\phi} < 0$  – обратной связи. При  $K_{\phi} = 0$  делают вывод об отсутствии связи. Чем ближе значение  $K_{\phi}$  по абсолютной величине к 1, тем связь теснее. Однако равенство коэффициента Фехнера единице ни в коей мере нельзя воспринимать как свидетельство функциональной зависимости между  $x$  и  $y$ .

Достоинством данного коэффициента является простота его расчета. Однако он лишь приблизительно оценивает тесноту корреляционной связи, т.к. не учитывает величину отклонений значений признаков от их средних величин. Коэффициент Фехнера практически характеризует не столько тесноту связи, сколько выявляет ее наличие и направление. Поэтому на основании этого коэффициента нельзя делать выводы о существенности корреляционной связи. Пример расчета коэффициента Фехнера приведен в разделе 10.6 данной темы.

### 10.3.2. Линейный коэффициент корреляции

*Линейный коэффициент корреляции ( $r$ ) характеризует тесноту и направление корреляционной связи между двумя количественными признаками в случае, если эта связь по форме является линейной.* Если заданы значения факторного ( $x$ ) и результативного ( $y$ ) признаков, то линейный коэффициент корреляции вычисляется по формуле

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}, \quad (10.4)$$

где  $r$  – линейный коэффициент корреляции.

Можно использовать и другие формулы, но результат должен быть одинаковым для всех вариантов расчета.

Линейный коэффициент корреляции изменяется в интервале от -1 до +1:  $-1 \leq r \leq 1$ . Чем ближе  $|r|$  к единице, тем связь между признаками теснее. Знак указывает направление связи: при  $r > 0$  связь прямая, при  $r < 0$  – обратная. Если  $|r| = 1$ , это означает, что связь между  $x$  и  $y$  является функциональной. Если же  $r = 0$ , то делают вывод об отсутствии *линейной* корреляционной связи между  $x$  и  $y$ . Однако это не исключает наличия *нелинейной* корреляционной связи, что может быть проверено с помощью *корреляционного отношения*.

Интерпретация выходных значений линейного коэффициента корреляции представлена в табл. 10.2.

Таблица 10.2

**Оценка линейного коэффициента корреляции**

Значение линейного коэффициента корреляции	Характер связи	Интерпретация связи
$r = 0$	Отсутствует линейная корреляционная связь	–
$0 < r < 1$	Прямая	С увеличением $x$ увеличивается $y$
$-1 < r < 0$	Обратная	С увеличением $x$ уменьшается $y$ , и наоборот
$r = 1$	Функциональная	Каждому значению факторного признака соответствует одно строго определенное значение результативного признака

Величина  $r^2$  представляет собой *коэффициент детерминации*, который показывает, какая доля вариации результативного признака обусловлена влиянием факторного признака.

При расчете линейного коэффициента корреляции необходимо оценить его существенность, или статистическую значимость. Необходимость такой проверки связана с тем, что коэффициент корреляции в основном оценивается на основе выборочных данных, т.е. является случайной величиной, вследствие чего вывод о наличии корреляционной связи в генеральной совокупности требует дополнительного подтверждения. С уменьшением числа наблюдений надежность выборочного коэффициента корреляции снижается.

Для проверки существенности (не случайности) линейного коэффициента корреляции используют соответствующие статистические критерии.

Для небольшого объема выборки ( $n < 30$ ) значимость линейного коэффициента корреляции проверяется на основе  $t$ -критерия Стьюдента [14, с. 228]:

$$t_p = \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}}. \quad (10.5)$$

где  $t_p$  – расчетное значение  $t$ -критерия, которое сопоставляется с табличным значением ( $t_{\alpha,k}$ ).

Если расчетное значение  $t$ -критерия больше табличного, т.е.  $t_p > t_{\alpha,k}$ , делают вывод о том, линейный коэффициент корреляции статистически значим, а связь является реальной. Если  $t_p < t_{\alpha,k}$ , то считается, что связь между  $x$  и  $y$  отсутствует и значение  $r$ , отличное от нуля, получено случайно.

В приведенных неравенствах табличное значение  $t_{\alpha,k}$  при уровне значимости  $\alpha$  и числе степеней свободы  $k = n - 2$  определяется по таблицам распределения Стьюдента (прил. 3).

При большом объеме выборки ( $n > 50$ ) [14, с. 227]  $r$  считается статистически значимым (неслучайным), если выполняется условие:

$$t_p > t_{\alpha}. \quad (10.6)$$

В этом случае расчетное значение критерия ( $t_p$ ) определяется по формуле

$$t_p = \frac{|r|}{\sigma_r}, \quad (10.7)$$

где  $\sigma_r$  – средняя квадратическая ошибка коэффициента корреляции, расчет которой выполняется по формуле

$$\sigma_r = \frac{1-r^2}{\sqrt{n}}. \quad (10.8)$$

Величина  $t_{\alpha}$  определяется по таблице интеграла Лапласа, где  $\alpha$  – уровень значимости, устанавливающий доверительную вероятность, с которой мы делаем вывод о существенности коэффициента корреляции. Уровень значимости устанавливается исследователем самостоятельно. Обычно  $\alpha$  принимают равным 0,05 или 0,01.

В литературе [14, с.227] приводится несколько упрощенный вариант проверки существенности линейного коэффициента корреляции при  $n > 50$ : если линейный коэффициент корреляции  $r$  превышает свою сред-

ную квадратическую ошибку более чем в три раза, т.е. если  $\frac{|r|}{\sigma_r} > 3$ , то он считается значимым, а связь реальной.

Линейный коэффициент корреляции достаточно точно оценивает степень тесноты связи лишь в случае линейной зависимости между признаками. При нелинейной корреляционной связи  $r$  недооценивает степень тесноты связи и даже может быть равен 0. Поэтому в таких случаях, для оценки степени тесноты корреляционной связи между двумя признаками, используют другой показатель: корреляционное отношение (индекс корреляции). Пример расчета линейного коэффициента корреляции приведен в п. 10.6 данной темы.

### 10.3.3. Корреляционное отношение

*Корреляционное отношение используют для измерения тесноты связи в случае наличия как линейной, так и нелинейной зависимости между количественными признаками.*

Различают эмпирическое и теоретическое корреляционное отношение. Последний показатель иногда называют *индексом корреляции*.

Эмпирическое корреляционное отношение ( $\eta_9$ ) рассчитывается на основе результатов аналитической группировки по формулам (10.9), (10.10):

$$\eta_9 = \sqrt{\frac{\delta^2}{\sigma^2}} \quad (10.9)$$

или

$$\eta_9 = \sqrt{1 - \frac{\overline{\sigma_j^2}}{\sigma^2}}, \quad (10.10)$$

где  $\delta^2$  – межгрупповая (факторная) дисперсия результативного признака;  
 $\sigma^2$  – общая дисперсия результативного признака;  
 $\overline{\sigma_j^2}$  – средняя из внутригрупповых дисперсий результативного признака.

Теоретическое корреляционное отношение ( $\eta$ ) рассчитывают по теоретическим значениям результативного признака ( $\hat{y}_x$ ), полученным по уравнению регрессии. Для расчета ( $\eta$ ) используют формулы:

$$\eta = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_{x_i} - \bar{y})^2}}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (10.11)$$

$$\eta = \sqrt{1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_{x_i})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (10.12)$$

или

$$\eta = \sqrt{\frac{\sigma_{\hat{y}_x}^2}{\sigma^2}}, \quad (10.13)$$

где  $\hat{y}_x$  – теоретические значения результативного признака, полученные по уравнению регрессии;

$\sigma_{\hat{y}_x}^2$  – факторная дисперсия результативного признака, рассчитанная для теоретических значений результативного признака:

$$\sigma_{\hat{y}_x}^2 = \frac{\sum_{i=1}^n (\hat{y}_{x_i} - \bar{y})^2}{n}. \quad (10.14)$$

Корреляционное отношение изменяется в пределах от 0 до 1. Чем ближе  $\eta$  к единице, тем связь теснее. Знак корреляционного отношения в интерпретации связи между признаками не используется.

Оценка существенности корреляционного отношения проводится с помощью  $F$ -критерия (критерия Фишера). Расчетное значение  $F$ -критерия вычисляют по формулам:

$$F_P = \frac{\eta_{\mathcal{G}}^2}{1 - \eta_{\mathcal{G}}^2} \cdot \frac{n - m}{m - 1}, \quad (10.15)$$

$$F_P = \frac{\eta^2}{1 - \eta^2} \cdot \frac{n - m}{m - 1}. \quad (10.16)$$

В формуле (10.15)  $m$  – число групп, в формуле (10.16)  $m$  – число параметров уравнения регрессии. Например, для уравнения прямой  $m = 2$ .

Если  $F_P > F_{\alpha; k_1; k_2}$  – корреляционное отношение признается существенным и, следовательно, можно говорить о статистической значимости связи между  $x$  и  $y$ . Напомним, что  $k_1 = m - 1$ ;  $k_2 = n - m$ ;  $F_{\alpha; k_1; k_2}$  – табличное значение  $F$ -критерия находят по таблицам  $F$ -распределения (прил. 1).

В случае множественной корреляции (два и более факторных признака) степень тесноты связи оценивается с помощью:

- линейного множественного коэффициента корреляции;
- множественного корреляционного отношения.

## 10.4. Расчет и интерпретация однофакторных уравнений регрессии

Задачей регрессионного анализа является моделирование корреляционной связи. Это означает, что нужно:

- 1) выбрать вид уравнения связи;
- 2) определить его параметры;
- 3) оценить достоверность найденного уравнения связи.

Выбор вида уравнения связи осуществляют на основе:

- содержательного анализа природы изучаемой зависимости;
- анализа эмпирической линии регрессии;
- опыта предыдущих исследований;
- выбора наиболее адекватного вида уравнения связи по минимальной величине средней квадратической ошибки ( $S_{\hat{y}_x}$ ) или средней ошибки аппроксимации ( $\bar{\epsilon}$ );

– параллельного сопоставления рядов значений факторного и результативного признаков, что позволяет исследовать характер изменения разностей между значениями  $x$  и  $y$ . Если равномерному изменению факторного признака соответствует равномерное изменение результативного, предполагают наличие линейной зависимости; если равномерному изменению факторного признака соответствует равноускоренное изменение результативного признака, имеем параболу второй степени; если результативный признак изменяется одинаковыми темпами – связь экспоненциальная (показательная  $y = ab^x$ ).

*Уравнение, описывающее корреляционную связь, называют уравнением корреляционной связи, или уравнением регрессии. Линия, построенная по точкам уравнения регрессии, называется теоретической линией регрессии.*

Уравнение регрессии по форме похоже на уравнение функциональной зависимости, но по сути отличается тем, что оно справедливо лишь в среднем для совокупности, а не для каждого явления, и зависит от объема совокупности (чем она больше, тем параметры уравнения типичнее, точнее). Чтобы отметить, что зависимость между двумя признаками выражается в среднем, значения результативного признака, найденные по уравнению регрессии, обозначают  $\hat{y}_x$ .

Рассмотрим уравнения связи для парной корреляции при разных типах функций.

*А. Линейная зависимость.* Линейная зависимость применяется тогда, когда эта форма связи выявляется в результате содержательного анализа характера связей изучаемых явлений или когда форма связи не установлена.



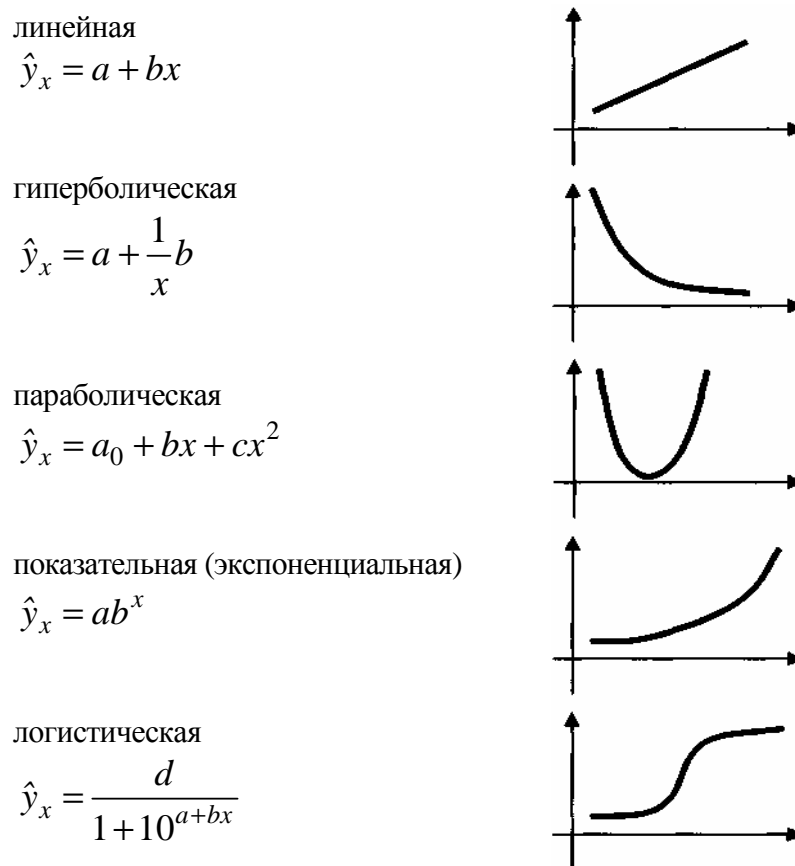


Рис. 10.2. Типы функций, используемых для характеристики связей экономических показателей

При линейной зависимости результативный признак изменяется под влиянием факторного равномерно. Уравнение регрессии в этом случае имеет вид:

$$\hat{y}_x = a + bx, \quad (10.17)$$

где  $\hat{y}_x$  – выровненные (теоретические) значения результативного признака;

$a, b$  – параметры уравнения прямой;

$x$  – факторный признак.

Нахождение параметров уравнения прямой проводится на основе выравнивания по способу наименьших квадратов, что приводит к системе нормальных уравнений:

$$\begin{cases} an + b \sum x = \sum y \\ a \sum x + b \sum x^2 = \sum xy. \end{cases} \quad (10.18)$$

Решение этой системы в общем виде дает следующие значения параметров  $a$  и  $b$ :

$$a = \frac{\sum y \sum x^2 - \sum xy \sum x}{n \sum x^2 - \sum x \sum x}, \quad (10.19)$$

$$b = \frac{n \sum xy - \sum y \cdot \sum x}{n \sum x^2 - \sum x \sum x}, \quad (10.20)$$

где  $b$  – коэффициент регрессии. Если связь прямая, то  $b$  имеет знак «плюс», если связь обратная, то знак «минус».

Коэффициент регрессии показывает, насколько в среднем изменяется результативный признак при изменении факторного признака на единицу. Геометрически  $b$  – это тангенс угла наклона теоретической линии регрессии к оси абсцисс.

Свободный член уравнения регрессии  $a$  не всегда может иметь экономический смысл. Геометрически это значение  $\hat{y}_x$  при  $x = 0$ .

Коэффициент регрессии  $b$  может быть определен на основе коэффициента корреляции:

$$b = r \frac{\sigma_y}{\sigma_x}. \quad (10.21)$$

Коэффициент регрессии применяют для определения коэффициента эластичности:

$$\mathcal{E}_x = b \frac{\bar{x}}{y}. \quad (10.22)$$

Коэффициент эластичности показывает, на сколько процентов в среднем изменится величина результативного признака при изменении факторного признака на один процент.

Определение параметров уравнения регрессии еще не позволяет сделать вывод о том, насколько полученное уравнение точно воспроизводит изучаемую связь между признаками.

Уравнение регрессии дает характеристику связи лишь в среднем. Но мы знаем, что расчет средней величины должен быть дополнен анализом показателей вариации, которые определяют, насколько средняя типична. Поэтому для оценки надежности уравнения регрессии нужно знать меру вариации (рассеивания) эмпирических точек относительно теоретической линии регрессии. Иногда рассеивание (разброс) эмпирических точек настолько велико, что найденное уравнение связи не имеет практической ценности.

Достоверность уравнения регрессии оценивают с помощью средней квадратической ошибки уравнения

$$S_{\hat{y}_x} = \sqrt{\frac{\sum (y - \hat{y}_x)^2}{n - m}}, \quad (10.23)$$

где  $n$  – число единиц совокупности;

$m$  – число параметров уравнения регрессии (для линейной связи  $m = 2$ );

$S_{\hat{y}_x}$  – средняя квадратическая ошибка уравнения регрессии.

Величина  $S_{\hat{y}_x}$  служит показателем точности, а, следовательно, и практической ценности уравнения регрессии.

Средняя квадратическая ошибка уравнения регрессии дает нам возможность для каждого конкретного значения факторного признака  $x_0$  определить доверительные границы результативного признака в генеральной совокупности:

$$\hat{y}_{x_0} - t_\alpha \frac{S_{\hat{y}_x}}{\sqrt{n}} \sqrt{1 + \frac{(x_0 - \bar{x})^2}{\sigma_x^2}} \leq y \leq \hat{y}_{x_0} + t_\alpha \frac{S_{\hat{y}_x}}{\sqrt{n}} \sqrt{1 + \frac{(x_0 - \bar{x})^2}{\sigma_x^2}}, \quad (10.24)$$

где  $t_\alpha$  – табличное значение  $t$ -критерия определяется по таблице распределения Стьюдента для уровня значимости  $\alpha$  и числа степеней свободы ( $n - 2$ ).

Так как параметры уравнения регрессии определены по выборочным данным, необходимо оценить их средние квадратические ошибки и доверительные интервалы.

Для параметра  $a$  средняя квадратическая ошибка рассчитывается по следующим формулам:

$$S_a = S_{\hat{y}_x} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sigma_x^2 \cdot n}} \quad (10.25)$$

или

$$S_a = S_{\hat{y}_x} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2}}, \quad (10.26)$$

Если  $\frac{|a|}{S_a} > t_{\alpha; n-2}$ , то значение параметра  $a$  признается существенным.

При этом  $t_{\alpha; n-2}$  соответствует  $t$ -критерию Стьюдента при заданном уровне значимости  $\alpha$  и  $(n - 2)$  степенях свободы.

Средняя квадратическая ошибка для параметра  $b$  определяется по формулам:

$$S_b = \frac{S_{\hat{y}_x}}{\sqrt{n\sigma_x^2}} \quad (10.27)$$

или

$$S_b = \frac{S_{\hat{y}_x}}{\sqrt{\sum (x - \bar{x})^2}} \quad (10.28)$$

Если  $\frac{|b|}{S_b} > t_{\alpha;n-2}$ , значение параметра  $b$  существенно. Здесь, также,

как и в предыдущем случае, правая часть неравенства представляет табличное значение  $t$ -критерия Стьюдента при уровне значимости  $\alpha$  и числе степеней свободы  $(n - 2)$ .

Зная средние квадратические ошибки параметров уравнения регрессии, можно определить их доверительные интервалы:

– для параметра  $a$ :

$$a - t_{\alpha;n-2}(S_a) \leq A \leq a + t_{\alpha;n-2}(S_a); \quad (10.29)$$

– для параметра  $b$ :

$$b - t_{\alpha;n-2}(S_b) \leq B \leq b + t_{\alpha;n-2}(S_b). \quad (10.30)$$

Если надежность уравнения регрессии подтверждена в результате проверки с помощью статистических критериев, оно может быть использовано для целей планирования и прогнозирования. Например, мы можем найти значения  $y$  за пределами области наших наблюдений.

Но обязательным условием применения уравнения регрессии является сохранение условий формирования уровней признаков, для которых было получено уравнение регрессии.

*Б. Параболическая зависимость.* Применяется при ускоренном возрастании или убывании результативного признака в сочетании с равномерным изменением факторного признака:  $\hat{y}_x = a + bx + cx^2$ .

Параметры уравнения параболы вычисляются путем решения системы трех нормальных уравнений.

$$\begin{cases} na + b\sum x + c\sum x^2 = \sum y \\ a\sum x + b\sum x^2 + c\sum x^3 = \sum yx \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum yx^2. \end{cases} \quad (10.31)$$

Средняя квадратическая ошибка уравнения регрессии определяется по формуле (10.23) при  $m = 3$ .

Коэффициент эластичности для этой формы связи определяется по формуле

$$\varepsilon = (b + c) \frac{\bar{x}}{y}. \quad (10.32)$$

*В. Уравнение гиперболы.* Уравнение регрессии в форме гиперболы имеет следующий вид:

$$\hat{y}_x = a + b \frac{1}{x}. \quad (10.33)$$

Гиперболические зависимости характерны для связей, в которых результативный признак не может варьировать неограниченно, его вариация имеет односторонний предел. Например, при освоении нового оборудования его производительность возрастает, но рост замедлится по мере приближения к конструктивно-технологическому пределу производительности агрегата.

Нормальные уравнения метода наименьших квадратов для гиперболы таковы:

$$\begin{cases} na + b \sum \frac{1}{x} = \sum y \\ a \sum \frac{1}{x} + b \sum \frac{1}{x^2} = \sum \frac{y}{x}. \end{cases} \quad (10.34)$$

Ошибка уравнения регрессии определяется по (10.23) при  $m = 2$ .

Для того чтобы подобрать тип функции, наиболее точно воспроизводящий изучаемую корреляционную связь, можно построить разные модели и выбрать те из них, которые имеют минимальные  $S_{\hat{y}_x}$  или  $\bar{\varepsilon}$ .

Для удобства восприятия средняя квадратическая ошибка уравнения регрессии может быть представлена относительной форме:

$$A = \frac{S_{\hat{y}_x}}{\bar{y}} \cdot 100 \%. \quad (10.35)$$

Если величина  $A$  не превышает 10 – 15 %, делают вывод, что уравнение регрессии достаточно точно описывает изучаемую корреляционную связь и его можно использовать для целей анализа, планирования и прогнозирования.

Средняя ошибка аппроксимации оценивается в процентах и рассчитывается по формуле

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \hat{y}_x|}{y} \cdot 100 \%. \quad (10.36)$$

Рекомендуемое значение средней ошибки аппроксимации не более 10 %.

Следует отметить, что наряду с проверкой значимости параметров уравнения регрессии, как это было рассмотрено для линейной зависимости, можно оценить значимость уравнения регрессии в целом. Эта задача решается на основе расчета  $F$ -критерия Фишера и сопоставления его с табличным (критическим) значением.

$F$ -критерий можно рассчитать через линейный коэффициент корреляции ( $r$ ) (для линейной связи) по формуле

$$F_p = \frac{r^2}{1-r^2} \cdot \frac{n-m}{m-1} \quad (10.37)$$

или через теоретическое корреляционное отношение ( $\eta$ ) для любой формы связи:

$$F_p = \frac{\eta^2}{1-\eta^2} \cdot \frac{n-m}{m-1}, \quad (10.38)$$

где  $m$  – число параметров в уравнении регрессии;

$n$  – число единиц совокупности.

Расчетное значение  $F$ -критерия ( $F_p$ ) сопоставляется с табличным, определяемым по прил. 1 для числа степеней свободы  $k_1 = m - 1$  и  $k_2 = n - m$  при заданном уровне значимости  $\alpha$  ( $F_{\alpha; k_1; k_2}$ ).

Если  $F_p > F_{\alpha; k_1; k_2}$ , уравнение регрессии считается значимым.

Пример проведения регрессионного анализа, в т.ч. с использованием средств Microsoft Excel, рассматривается в п. 10.6 данной темы.

## 10.5. Понятие о множественной корреляции и факторном анализе

Выше рассматривалась связь между двумя признаками, т.е. парная корреляция. В действительности величина результативного признака зависит от многих причин. Исследование зависимости результативного признака от двух и более факторных признаков проводят на основе методов множественной корреляции.

Методика множественной корреляции основывается на тех же принципах, что и парная корреляция. Но она имеет свои особенности и использует более сложный математический аппарат.

Рассмотрим основные этапы исследования множественной корреляции.

1. *Отбор факторных признаков.* Основу отбора должен составлять содержательный анализ связи. Для анализа отбирают наиболее существен-

ные факторы, количественно измеряемые и по которым имеется достоверная информация. Из набора возможных факторных признаков исключают те, которые находятся между собой в функциональной или тесной корреляционной связи ( $|r| \geq 0,8$ ). При введении таких факторов в уравнение наблюдают явление коллинеарности (если факторов два) или мультиколлинеарности, если факторов более двух. Явление коллинеарности выявляют путем построения матрицы парных коэффициентов корреляции, в которой показывается теснота связи каждого фактора с результативным признаком и с каждым из других факторов. Состав отобранных факторов должен быть практически целесообразным, минимальным, но достаточным для решения поставленной задачи.

*2. Выбор аналитического выражения, наилучшим образом описывающего связь факторных признаков с результативным.*

Выбрать форму связи достаточно сложно. Решение этой задачи на практике основывается на априорном теоретическом анализе изучаемого явления и подборе известных типов математических моделей. Среди многофакторных регрессионных моделей выделяют линейные и нелинейные. Наиболее простыми для построения являются многофакторные линейные модели (формула (10.39)), которые содержат независимые переменные только в первой степени:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n. \quad (10.39)$$

Если связь между результативным признаком и анализируемыми факторами нелинейна, то выбранная для ее описания нелинейная многофакторная модель (степенная, показательная и т.п.) может быть сведена к линейной путем линеаризации.

Простейшее линейное уравнение множественной корреляционной зависимости результативного признака от двух факторных ( $x_1, x_2$ ) можно записать следующим образом:

$$\hat{y}_{x_1, x_2} = a_0 + a_1x_1 + a_2x_2, \quad (10.40)$$

где  $a_0$  – свободный член;

$a_1, a_2$  – коэффициенты регрессии;

$x_1, x_2$  – значения факторных признаков.

*3. Определение параметров уравнения множественной корреляционной связи.*

Параметры уравнения множественной регрессии, как и парной, рассчитывают методом наименьших квадратов. При этом решается система нормальных уравнений с  $(n + 1)$  неизвестными:

$$\begin{cases} a_0 n + a_1 \sum x_1 + a_2 \sum x_2 = \sum y \\ a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum y x_1 \\ a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum y x_2. \end{cases} \quad (10.41)$$

Эту систему уравнений можно решить разными способами, в т.ч. через коэффициенты парной корреляции и средние квадратические отклонения:

$$a_0 = \bar{y} - a_1 \bar{x}_1 - a_2 \bar{x}_2, \quad (10.42)$$

$$a_1 = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1 x_2}}{1 - r_{x_1 x_2}^2} \cdot \frac{\sigma_y}{\sigma_{x_1}}, \quad (10.43)$$

$$a_2 = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1 x_2}}{1 - r_{x_1 x_2}^2} \cdot \frac{\sigma_y}{\sigma_{x_2}}. \quad (10.44)$$

По параметрам полученного уравнения можно оценить влияние каждого из факторов на изменение резульативного признака.

#### 4. Интерпретация результатов решения множественной регрессии.

Для характеристики влияния факторных признаков на резульативный используют следующие показатели:

- коэффициенты регрессии при каждом из факторов;
- частные коэффициенты эластичности  $\mathcal{E}_{x_j}$ ;
- стандартизированные частные коэффициенты регрессии ( $\beta$ -коэффициенты);
- дельта коэффициенты ( $\Delta$ -коэффициенты).

Рассмотрим эти показатели.

*Коэффициенты уравнения множественной регрессии* показывают абсолютный размер влияния факторов на анализируемый резульативный показатель при фиксировано (среднем) уровне других факторов, входящих в модель.

Для сравнения оценок роли различных факторов в формировании резульативного показателя следует дополнить коэффициенты регрессии относительными величинами (коэффициентами эластичности,  $\beta$ -коэффициентами,  $\Delta$ -коэффициентами).

*Частные коэффициенты эластичности* показывают, на сколько процентов в среднем изменяется  $y$  с изменением признака-фактора на один процент при фиксированных значениях других факторов и рассчитывается как

$$\mathcal{E}_{x_j} = a_j \frac{\bar{x}_j}{\bar{y}}, \quad (10.45)$$

где  $a_j$  – коэффициент регрессии при  $j$ -факторе.



Стандартизированные частные коэффициенты регрессии ( $\beta$ -коэффициенты) показывают, на какую часть среднего квадратического отклонения  $\sigma_y$  изменится зависимая переменная  $y$  с изменением соответствующего фактора  $x_j$  на величину своего среднеквадратического отклонения ( $\sigma_j$ ). Этот коэффициент позволяет сравнивать влияние колеблемости различных факторов на вариацию результативного показателя, на основе чего выявляются факторы, в развитии которых заложены наибольшие резервы изменения этого показателя:

$$\beta_j = a_j \frac{\sigma_j}{\sigma_y}. \quad (10.46)$$

Дельта-коэффициенты определяют долю влияния каждого фактора в суммарном влиянии факторов, включенных в уравнение регрессии:

$$\Delta_j = \frac{r_{jy} \cdot \beta_j}{\sum_j r_{jy} \cdot \beta_j} = \frac{r_{jy} \cdot \beta_j}{R^2}. \quad (10.47)$$

Содержательный анализ уравнений множественной регрессии в целях уточнения приоритетности факторов опирается на сравнение перечисленных коэффициентов. В этих целях, особенно при достаточно большом числе факторов, включаемых в уравнение регрессии, производится ранжирование факторов по величине коэффициентов эластичности, бета- и дельта-коэффициентов.

5. Оценка тесноты множественной корреляционной связи производится с помощью теоретического корреляционного отношения (совокупный индекс корреляции) по следующей формуле (для двухфакторной модели):

$$i_{y, x_1, x_2} = \sqrt{1 - \frac{\sigma_{(y - \hat{y}_{x_1, x_2})}^2}{\sigma_y^2}}, \quad (10.48)$$

где  $\sigma_{(y - \hat{y}_{x_1, x_2})}^2$  – дисперсия эмпирических значений результативного признака относительно его значений, рассчитанных по уравнению регрессии, которая определяется делением суммы квадратов указанных отклонений результативного признака на  $(n - k - 1)$ ;

$\sigma_y^2$  – дисперсия эмпирических значений результативного признака;

$n$  – объем выборки;

$k$  – количество факторных признаков.

Множественное корреляционное отношение варьирует в тех же пределах и имеют тот же смысл, что и в парной корреляции. Множественное корреляционное отношение измеряет тесноту корреляционной связи как при линейной, так нелинейной форме связи. Квадрат множественного корреляционного отношения представляет собой множественный коэффициент детерминации, который показывает, долю влияния учтенных в модели факторов в суммарном влиянии факторов.

6. *Оценка существенности и надежности параметров множественной корреляции.*

Для оценки значимости каждого коэффициента регрессии необходимо рассчитать значение  $t$ -критерия Стьюдента (отношение коэффициента регрессии к его средней ошибке) [14, с. 277]:

$$t_{расч.}^j = \frac{|a_j|}{\sigma_{a_j}}. \quad (10.49)$$

Средняя (стандартная) ошибка коэффициента регрессии может быть найдена по формуле [14, с. 227]:

$$\sigma_{a_j}^2 = \frac{\sigma_y^2}{k}, \quad (10.50)$$

где  $\sigma_y^2$  – дисперсия результативного признака;

$k$  – количество факторных признаков.

Коэффициент регрессии считается статистически значимым, если  $t_{расч.}^j$  превышает табличное (теоретическое)  $t_{табл.}$  значение  $t$ -критерия Стьюдента для заданного уровня значимости  $\alpha$  и  $(n - k - 1)$  степеней свободы:

$$t_{расч.}^j > t_{табл.}^{\alpha; n-k-1}.$$

Вывод об адекватности уравнения регрессии и правильности выбора связи можно проводить с помощью  $F$ -критерия [14, с. 278]:

$$F_{расч.} = \frac{R^2}{1 - R^2} \cdot \frac{n - k - 1}{k}, \quad (10.51)$$

где  $R^2$  – совокупный коэффициент множественной детерминации.

Величина  $F_{табл.}$  находится по таблицам  $F$ -критерия Фишера (прил. 1). Входами в таблицу являются: заданный исследователем уровень значимости  $\alpha$  и значения  $\nu_1 = k$ ,  $\nu_2 = n - k - 1$ . Если  $F_{расч.} > F_{табл.}$ , связь признается существенной (статистически значимой).

В исследовании связей между явлениями используются *метода факторного анализа* [18. с. 589]. Факторный анализ представляет собой совокупность методов, которые на основе реально существующих связей анализируемых признаков или связей самих наблюдаемых объектов позволяют выявлять скрытые (неявные, латентные) обобщающие характеристики организационной структуры и механизма развития изучаемых явлений, процессов.

Методы факторного анализа в исследовательской практике применяются, главным образом, с целью сжатия информации, получения небольшого числа обобщающих признаков, объясняющих вариативность (дисперсию) элементарных признаков (*R*-техника факторного анализа) или вариативность наблюдаемых объектов (*Q*-техника факторного анализа).

Алгоритмы факторного анализа основываются на использовании редуцированной матрицы парных корреляций. Редуцированная матрица – это такая матрица, на главной диагонали которой расположены не единицы (оценки полной корреляции) или не оценки полной дисперсии, а их редуцированные, несколько уменьшенные величины. При этом постулируется, что в результате анализа будет объяснена не вся дисперсия изучаемых признаков (объектов), а ее некоторая, обычно бóльшая, оставшаяся необъясненная часть дисперсии – это характерность, возникшая из-за специфически наблюдаемых объектов или ошибок, допускаемых при регистрации явлений, процессов, т.е. ненадежности вводных данных.

## 10.6. Практические примеры анализа корреляционной связи

### Пример 10.1

По данным табл. 10.4 оцените тесноту связи между суммой прибыли банков и размером их активов с помощью коэффициента корреляции знаков (коэффициента Фехнера).

Таблица 10.4

Размер активов и прибыли по группе банков

Номер банка	Активы банка, млн руб.	Прибыль, млн руб.
1	866	39,6
2	328	17,8
3	207	12,7
4	185	14,9
5	109	4,0
6	104	15,5
7	327	6,4
8	113	10,1
9	91	3,4
10	849	13,4

Решение:

1. Определим причинно-следственные связи между анализируемыми признаками:

- $x$  (факторный признак) – размер активов;
- $y$  (результативный признак) – размер прибыли.

2. Составим вспомогательную таблицу. Факторный признак располагаем в возрастающем порядке.

Таблица 10.5

**Вспомогательная расчетная таблица**

№	$x$ , млн руб.	$y$ , млн руб.	Знак отклонения от средней величины		Совпадение ( $n_c$ ) или несовпадение ( $n_{nc}$ ) знаков
			для $x$	для $y$	
1	91	3,4	–	–	( $n_c$ )
2	104	15,5	–	+	( $n_{nc}$ )
3	109	4,0	–	–	( $n_c$ )
4	113	10,1	–	–	( $n_c$ )
5	185	14,9	–	+	( $n_{nc}$ )
6	207	12,7	–	–	( $n_c$ )
7	327	6,4	+	–	( $n_{nc}$ )
8	328	17,8	+	+	( $n_c$ )
9	849	13,4	+	–	( $n_{nc}$ )
10	866	39,6	+	+	( $n_c$ )
Итого	$\Sigma 3179$	$\Sigma 137,8$	–		–

3. Определяем среднее значение для  $x$  и  $y$ :

$$\bar{x} = \frac{\sum x}{n} = \frac{3179}{10} = 317,9 \text{ (млн руб.)},$$

$$\bar{y} = \frac{\sum y}{n} = \frac{137,8}{10} = 13,78 \text{ (млн руб.)}.$$

4. Рассчитаем коэффициент корреляции знаков:

$$K_{\phi} = \frac{n_c - n_{nc}}{n_c + n_{nc}} = \frac{6 - 4}{10} = +0,2.$$

5. *Вывод о степени тесноты связи:* можно предполагать, что между активами банка и величиной прибыли существует прямая корреляционная связь.

### Пример 10.2

По данным таблицы из задания 10.1 оцените степень тесноты связи между анализируемыми признаками с помощью линейного коэффициента корреляции. Сделайте выводы о статистической значимости связи между размером активов и величиной прибыли.

Решение:

Расчетная формула линейного коэффициента корреляции:

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

Для применения данной формулы необходимо выполнить вспомогательные расчеты. Необходимые исходные данные для расчета линейного коэффициента корреляции представлены в табл. 10.6.

Таблица 10.6

Вспомогательная расчетная таблица

№	$x$	$y$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$xy$
1	866	39,6	300 413,61	666,6724	34 293,6
2	328	17,8	102,01	16,1604	5 838,4
3	207	12,7	12 298,81	1,1664	2 628,9
4	185	14,9	17 662,41	1,2544	2 756,5
5	109	4,0	43 639,21	95,6484	436
6	104	15,5	45 753,21	2,9584	1 612
7	327	6,4	82,81	54,4644	2 092,8
8	113	10,1	41 984,01	13,5424	1 141,3
9	91	3,4	51 483,61	107,7444	309,4
10	849	13,4	282 067,21	0,1444	11 376,6
итого	$\Sigma 3179$	$\Sigma 137,8$	795 486,9	959,756	62 485,5

$$\bar{x} = 317,9; \quad \bar{y} = 13,78.$$

$$\sigma_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{795\,486,9}{10}} = 282,04.$$

$$\sigma_y = \sqrt{\frac{\sum (y - \bar{y})^2}{n}} = \sqrt{\frac{959,756}{10}} = 9,797.$$

$$\overline{xy} = \frac{\sum xy}{n} = \frac{62\,485,5}{10} = 6\,248,55.$$

$$r = \frac{6\,248,5 - 317,9 \cdot 13,78}{282,04 \cdot 9,797} = \frac{6\,248,5 - 4\,380,662}{2\,763,15} = 0,68.$$

Проверка существенности  $r$ :

$$t_p = \frac{|r| \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,67 \sqrt{10-2}}{\sqrt{1-0,67^2}} = \frac{1,895}{0,742} = 2,55.$$

Табличное значение  $t$ -критерия при уровне значимости  $\alpha = 0,05$  и числе степеней свободы  $n - 2 = 10 - 2 = 8$  определяем по таблице  $t$ -распределения Стьюдента:  $t_{0,05;8} = 2,306$  (прил. 3).

Так как  $t_p > t_{\alpha,k}$ , делаем вывод о существенности (статистической значимости) корреляционной связи между размером активов банка и величиной прибыли. Связь между признаками прямая, умеренная, статистически значимая.

### Пример 10.3

По данным таблицы из задания 10.1 оцените степень тесноты связи между активами и прибылью по группе банков с помощью коэффициента корреляции рангов Спирмена. Сделайте выводы о существенности коэффициента корреляции рангов Спирмена.

Решение:

1. Расчет коэффициента корреляции рангов Спирмена выполняется по формуле

$$r = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}.$$

Для расчета  $r$  необходимо провести ранжирование банков по каждому из признаков. Результаты ранжирования представим в табл. 10.7.

Таблица 10.7

**Ранжирование значений факторного и результирующего признаков**

№	Активы банка $x$ , млн. руб.	Ранг по $x$	Прибыль $y$ , млн. руб.	Ранг по $y$	$d_i = (R_{x_i} - R_{y_i})$	$d_i^2$
1	91	1	3,4	1	0	0
2	104	2	15,5	8	-6	36
3	109	3	4,0	2	1	1
4	113	4	10,1	4	0	0
5	185	5	14,9	7	-2	4
6	207	6	12,7	5	1	1
7	327	7	6,4	3	4	16
8	328	8	17,8	9	1	1
9	849	9	13,4	6	3	9
10	866	10	39,6	10	0	0
Итого					0	68

$$r = 1 - \frac{6 \cdot 68}{10(10^2 - 1)} = 0,588.$$

Проверим существенность коэффициента корреляции рангов. По прил. 4 «Значения коэффициента корреляции рангов Спирмена» определяем таблич-

ное значение коэффициента корреляции рангов при  $\alpha = 0,05$  и  $n = 10$ :  $P_{\alpha;n} = P_{0,05;10} = 0,6364$ . Так как  $p < P_{0,05;10}$ , с вероятностью 95 % можно утверждать, что величина  $p = 0,588$  не является статистически значимой и связь между анализируемыми признаками нельзя признать существенной.

Однако, если принять уровень значимости  $\alpha = 0,1$ , получим:  $P_{0,1;10} = 0,5515$  и  $p > P_{0,1;10}$ . Следовательно, с доверительной вероятностью 90 % можно утверждать, что связь между величиной активов и размером прибыли является статистически значимой.

Поэтому общий вывод по результатам анализа может состоять в рекомендации проведения расчетов по большей совокупности банков, а при отсутствии такой возможности, к результатам расчетов следует относиться с достаточной осторожностью.

### Пример 10.4

Компанию, доставляющую товары почтой, интересует, существенно ли различие между регионами по количеству заказов, сделанных на основе разосланных каталогов. Исходные данные для анализа представлены в табл. 10.8.

Таблица 10.8

**Распределение каталогов по регионам и по характеру отклика**  
(кол-во каталогов)

Распределение разосланных каталогов по регионам	Отклик на разосланные каталоги				Итого
	Заказ сделан		Заказ не сделан		
Восток	926	$a$	22 113	$b$	23 039 ( $a + b$ )
Запад	352	$c$	10 617	$d$	10 969 ( $c + d$ )
Итого	1278	$(a + c)$	32 730	$(b + d)$	34 008

На основе расчета коэффициента ассоциации и коэффициента контингенции сделайте вывод о том, необходимо ли компании дифференцировать свою сбытовую политику по регионам.

Решение:

Определим коэффициент ассоциации и коэффициент контингенции:

$$K_a = \frac{ad - bc}{ad + bc} = \frac{926 \cdot 10\,617 - 352 \cdot 22\,113}{926 \cdot 10\,617 + 352 \cdot 22\,113} = +0,116.$$

$$K_k = \frac{ad - bc}{\sqrt{(a+b)(b+d)(d+c)(a+c)}} = \frac{926 \cdot 10\,617 - 352 \cdot 22\,113}{\sqrt{23\,039 \cdot 32\,730 \cdot 10\,969 \cdot 1\,278}} = +0,02.$$

**Вывод:** т.к.  $K_a < 0,5$ , а  $K_k < 0,3$ , различие между регионами по количеству заказов, сделанных на основе разосланных каталогов, несущественно и компании нет необходимости дифференцировать свою сбытовую политику по регионам.

### Пример 10.5

По результатам опроса о заинтересованности людей в повышении своего образовательного уровня получены следующие данные:

Группы респондентов по полу	Заинтересованность в повышении образовательного уровня					Итого
	заинтересованы в очень высокой степени	заинтересованы в значительной степени	заинтересованы отчасти	не заинтересованы	загрудняюсь ответить	
Мужчины	14	12	6	7	1	40
Женщины	12	20	13	15	2	62
Итого	26	32	19	22	2	102

С помощью коэффициентов сопряженности Пирсона и Чупрова исследуйте связь между полом респондентов и мотивацией к повышению образовательного уровня.

Решение:

Исходные формулы имеют вид:

– коэффициент взаимной сопряженности Пирсона:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{1 + \varphi^2}};$$

– коэффициент взаимной сопряженности Чупрова:

$$K_{\Pi} = \sqrt{\frac{\varphi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}},$$

где  $\varphi^2$  – показатель взаимной сопряженности

$$\varphi^2 = \sum \frac{n_{xy}^2}{n_x \cdot n_y} - 1 \text{ и } 1 + \varphi^2 = \sum \frac{n_{xy}^2}{n_x \cdot n_y} = \sum \frac{\sum n_{xy}^2}{n_y} = \sum \frac{\sum n_{xy}^2}{n_x};$$



$n_{xy}^2$  – квадрат частоты каждой клетки;

$n_x$  – итоговая частота соответствующей строки;

$n_y$  – итоговая частота соответствующего столбца;

$K_1$  – число значений первого признака;

$K_2$  – число значений второго признака.

Для лучшего восприятия вышеприведенной формулы приведем вспомогательную таблицу:

Таблица 10.9

Вспомогательная таблица для расчета  $\phi^2$

$x$	$y$	I	II	III	IV	Итого
I		$n_{xy}$	$n_{xy}$	$n_{xy}$	$n_{xy}$	$n_x$
II		$n_{xy}$				$n_x$
III		$n_{xy}$				$n_x$
Итого		$n_y$	$n_y$	$n_y$	$n_y$	$n$

$$1 + \phi^2 = \sum \frac{n_{xy}^2}{n_x n_y} = \frac{14^2}{26} + \frac{12^2}{32} + \frac{6^2}{19} + \frac{7^2}{22} + \frac{1^2}{3} + \frac{12^2}{26} + \frac{20^2}{32} + \frac{13^2}{19} + \frac{15^2}{22} + \frac{2^2}{3} = 1,03;$$

$$\phi^2 = 0,03.$$

Коэффициент сопряженности Пирсона:

$$K_{II} = \sqrt{\frac{\phi^2}{1 + \phi^2}} = \sqrt{\frac{0,03}{1,03}} = 0,17.$$

Коэффициент взаимной сопряженности Чупрова:

$$K_{Ч} = \sqrt{\frac{\phi^2}{\sqrt{(K_1 - 1)(K_2 - 1)}}} = \sqrt{\frac{0,03}{\sqrt{(2 - 1)(5 - 1)}}} = 0,12.$$

**Вывод:** влияние пола респондентов на их мотивацию несущественно, т.к. связь между признаками слабая.

### Пример 10.6

По данным табл. 10.10 обоснуйте выбор вида уравнения корреляционной связи, найдите параметры уравнения регрессии и проверьте его надежность.

Таблица 10.10

**Показатели уровня механизации производства и объема продукции  
по группе предприятий отрасли**

№	Уровень фондовооруженности труда, тыс. руб./чел.	Объем продукции (услуг) за год, млн руб.
1	60	77
2	64	91
3	77	85
4	83	90
5	95	95
6	100	99
7	105	100
8	108	110
9	110	105
10	118	120

Решение:

Определим роли признаков в причинно-следственных отношениях:

– факторный признак ( $x$ ) – уровень фондовооруженности труда, тыс. руб./чел.;

– результативный признак ( $y$ ) – объем продукции (услуг), млн руб.

Используем элементарные методы выявления корреляционной связи. Сопоставление параллельных рядов значений  $x$  и  $y$  позволяет предполагать наличие между признаками прямой корреляционной связи (см. табл. 10.10).

Таблица 10.11

**Расчетные данные для проведения КРА**

№	$x$	$y$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$xy$	$X^2$	$\hat{y}_x$	$(y - \hat{y}_x)^2$
1	60	77	1024	408,04	4 620	3 600	79,065	4,2642
2	64	91	784	38,44	5 824	4 096	81,337	93,3736
3	77	85	225	148,84	6 545	5 929	88,721	13,8458
4	83	90	81	51,84	7 470	6 889	92,129	4,5326
5	95	95	9	4,84	9 025	9 025	98,945	15,5630
6	100	99	64	3,24	9 900	10 000	101,785	7,7562
7	105	100	169	7,84	10 500	11 025	104,625	21,3906
8	108	110	256	163,84	11 880	11 664	106,329	13,4762
9	110	105	324	60,84	11 550	12 100	107,465	6,0762
10	118	120	676	519,84	14 160	13 924	112,009	63,6861
Итого	920	972	3612	1407,6	91 474	88 252	–	243,9645

Графическое представление связи между признаками с помощью поля корреляции позволяет предположить наличие линейной по форме связи.

По результатам использования элементарных методов анализа корреляционной связи выбираем для ее моделирования уравнение прямой:

$$\hat{y}_x = a + bx.$$

Параметры уравнения  $a$  и  $b$  определим на основе МНК, решив систему нормальных уравнений:

$$\begin{cases} na + b \sum x = \sum y \\ a \sum x + a \sum x^2 = \sum yx. \end{cases}$$

$$a = \frac{\sum y \sum x^2 - \sum xy \sum x}{n \sum x^2 - \sum x \sum x}, \quad b = \frac{n \sum xy - \sum y \sum x}{n \sum x^2 - \sum x \sum x}.$$

Вспомогательные расчеты для определения параметров  $a$  и  $b$  представлены в табл. 10.11.

$$\bar{x} = \frac{\sum x}{n} = \frac{920}{10} = 92; \quad \bar{y} = \frac{\sum y}{n} = \frac{972}{10} = 97,2.$$

$$\sigma_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{3612}{10}} = 19,01;$$

$$\sigma_y = \sqrt{\frac{\sum (y - \bar{y})^2}{n}} = \sqrt{\frac{1407,6}{10}} = 11,86.$$

$$V_x = \frac{\sigma_x}{\bar{x}} \cdot 100 = \frac{19,01}{92} \cdot 100 = 20,6 \%;$$

$$V_y = \frac{\sigma_y}{\bar{y}} \cdot 100 = \frac{11,86}{97,2} \cdot 100 = 12,2 \%.$$

Таким образом, требование однородности исследуемой совокупности по  $x$  и  $y$  соблюдается:  $V_x < 33 \%$ ;  $V_y < 33 \%$ .

$$a = \frac{972 \cdot 88\,252 - 91\,474 \cdot 920}{10 \cdot 88\,252 - 920 \cdot 920} = \frac{85\,780\,944 - 84\,156\,080}{882\,520 - 846\,400} =$$

$$= \frac{1\,624\,864}{36\,120} = 44,985;$$

$$b = \frac{10 \cdot 91\,474 - 972 \cdot 920}{36\,120} = \frac{914\,740 - 894\,240}{36\,120} = 0,568.$$

Следовательно, уравнение регрессии имеет вид:

$$\hat{y}_x = 44,985 + 0,568x.$$

Рассчитаем теоретические значения результативного признака на основе полученного уравнения. Результаты расчета сведены в табл. 10.11.

Оцениваем точность уравнения регрессии с помощью средней квадратической ошибки:

$$S_{\hat{y}_x} = \sqrt{\frac{\sum (y - \hat{y}_x)^2}{n - m}} = \sqrt{\frac{243,9645}{10 - 2}} = 5,52 \text{ (млн руб.)}.$$

В относительном выражении величина ошибки составит:

$$A = \frac{S_{\hat{y}_x}}{\bar{y}} \cdot 100 \% = \frac{5,52}{97,2} \cdot 100 \% = 5,68 \%$$

Допустимый уровень средней квадратической ошибки составляет 10 – 15 %, следовательно, можно говорить о том, что полученное уравнение регрессии достаточно точно воспроизводит исследуемую корреляционную связь.

Так как параметры уравнения регрессии определяются по выборочным данным, оценки параметров  $a$  и  $b$  содержат некоторую погрешность. Поэтому необходимо провести оценку статистической значимости параметров уравнения регрессии. Параметры уравнения регрессии признаются статистически значимыми, если выполняются условия:

$$\frac{|a|}{S_a} > t_{\alpha;n-2} \quad \text{и} \quad \frac{|b|}{S_b} > t_{\alpha;n-2},$$

где  $S_a$  – средняя квадратическая ошибка параметра  $a$ :

$$S_a = S_{\hat{y}_x} \cdot \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{n \cdot \sigma_x^2}} = 5,52 \cdot \sqrt{\frac{1}{10} + \frac{92^2}{10 \cdot 361,2}} = 8,628;$$

$S_b$  – средняя квадратическая ошибка параметра  $b$ :

$$S_b = \frac{S_{\hat{y}_x}}{\sqrt{n \cdot \sigma_x^2}} = \frac{5,52}{\sqrt{10 \cdot 361,2}} = 0,0918;$$

$t_{\alpha;n-2}$  – табличное значение  $t$ -критерия, определяем по таблицам  $t$ -распределения Стьюдента:

$$t_{0,05;8} = 2,306.$$

$$\frac{|a|}{S_a} = \frac{44,985}{8,628} = 5,21.$$

Таким образом,  $t_p > t_{\alpha;n-2}$  ( $5,21 > 2,306$ ).

$$\frac{|b|}{S_b} = \frac{0,568}{0,0918} = 6,19 \quad (6,19 > 2,306).$$

Следовательно, значения параметров уравнения регрессии статистически значимы (не случайны).

Вывод о возможности практического применения полученного уравнения регрессии:

Надежность уравнения регрессии подтверждается, во-первых, статистической значимостью параметров  $a$  и  $b$  и, во-вторых, тем, что ошибка уравнения регрессии не превышает допустимого уровня ( $A = 5,68 \%$ ). Та-

ким образом, уравнение регрессии может быть использовано в практической работе для целей планирования, анализа и прогнозирования. При этом в качестве нормативов можно использовать коэффициент регрессии  $b$  и коэффициент эластичности  $\mathcal{E}$ . В нашем случае  $b = 0,568$ . Это означает, что при увеличении уровня механизации производства на единицу (один процентный пункт) объем продукции увеличится на 0,568 млн руб.

$$\mathcal{E} = b \frac{\bar{x}}{\bar{y}} = 0,568 \cdot \frac{92}{97,2} = 0,538.$$

Коэффициент эластичности показывает, что при увеличении уровня механизации производства на 1 %, объем продукции увеличится на 0,538 %.

Следует иметь в виду, что практическое использование уравнения регрессии будет корректным только в условиях, соответствующих условиям формирования уровней признаков, которые лежат в основе определения параметров уравнения корреляционной связи.

В прил. 5 представлены результаты решения примера 10.6 с помощью средств Microsoft Excel.

## **ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ И ПРАКТИЧЕСКИХ ЗАНЯТИЙ ПО ТЕМЕ**

### **Вопросы для самоподготовки и контроля знаний**

1. В чем сущность корреляционной связи?
2. В чем состоят различия между корреляционной и функциональной связью?
3. Почему корреляционную связь называют неполной?
4. По каким признакам проводят классификацию корреляционных связей?
5. Что собой представляет парная корреляция?
6. Чем характеризуется множественная корреляция?
7. Охарактеризуйте элементарные методы выявления корреляционной связи.
8. Назовите основные этапы проведения корреляционно-регрессионного анализа.
9. Какие требования предъявляются к исходным данным для проведения корреляционно-регрессионного анализа?
10. В чем разница между корреляционным анализом и регрессионным анализом?
11. Какие показатели являются мерой тесноты связи между двумя признаками?

12. Почему необходимо оценивать существенность показателей степени тесноты корреляционной связи?
13. Как оценить существенность линейного коэффициента корреляции?
14. Какие показатели используют для измерения степени тесноты корреляционной связи между качественными признаками?
15. Как измеряется теснота корреляционной связи с помощью корреляционного отношения?
16. Что такое уравнение регрессии?
17. Что характеризует коэффициент регрессии?
18. Какая существует связь между линейным коэффициентом корреляции и коэффициентом регрессии?
19. Как проверить надежность уравнения регрессии?
20. Какой размер средней ошибки аппроксимации считается допустимым?
21. Что означает величина коэффициента эластичности, равная 0,58?
22. Укажите сферы возможного практического применения уравнений регрессии.
23. Укажите, в чем заключаются условия корректного использования уравнений регрессии на практике.

### Тест 10.1

1. Укажите, какие утверждения относятся к корреляционной связи:
  - а) значению факторного признака соответствует строго определенное значение результативного;
  - б) связь проявляется в каждом единичном случае;
  - в) факторный признак полностью определяет значение результативного;
  - г) значению факторного признака соответствует распределение значений результативного.
2. Параллельное сопоставление рядов, не сгруппированных значений факторного и результативного признаков, позволяет ...
  - а) измерить степень тесноты корреляционной связи;
  - б) установить наличие корреляционной связи;
  - в) определить направление корреляционной связи.
3. Укажите, какие характеристики относятся к функциональной связи:
  - а) связь является полной;
  - б) жестко детерминированная связь;
  - в) стохастически детерминированная связь.
4. Аналитическое выражение корреляционной связи осуществляется на основе ...

- а) корреляционного анализа;
- б) регрессионного анализа;
- в) аналитической группировки.

5. Какой вид анализа (корреляционный, или регрессионный) применяется в каждой из описанных ниже ситуаций:

- а) установление наличия какой-либо взаимосвязи между расходами на рекламу и объемом продаж;
- б) создание инструмента формирования сметы, позволяющего выражать затраты в зависимости от количества произведенных изделий;
- в) анализ данных с целью определения силы взаимосвязи между моральным состоянием работников и их производительностью.

6. Постройте поле корреляции, которое бы иллюстрировало каждую из указанных ниже ситуаций (для ответа необязательно использовать какие-либо данные, можно просто рисовать точки):

- а) взаимосвязь между  $x$  и  $y$  отсутствует;
- б) линейная взаимосвязь с сильной положительной корреляцией;
- в) линейная взаимосвязь со слабой отрицательной корреляцией.

7. Линейный коэффициент корреляции используют для оценки степени тесноты корреляционной связи между двумя количественными признаками ...

- а) только при линейной связи;
- б) только при нелинейной связи;
- в) независимо от типа связи.

8. Какой вывод о связи между признаками можно сделать, если линейный коэффициент корреляции:

- а) равен единице;
- б) равен нулю;
- в) равен (-0,85).

9. При наличии линейной корреляционной связи между двумя признаками линейный коэффициент корреляции составил 0,8. Укажите правильное значение коэффициента детерминации:

- а) -0,64;
- б) 0,64;
- в) 0,86.

10. Корреляционное отношение используют для оценки степени тесноты корреляционной связи ...

- а) только при криволинейной связи;
- б) только при линейной связи;
- в) при любом типе связей.

11. Укажите, в каких пределах изменяется корреляционное отношение:

- а)  $-1 \leq \eta \leq +1$ ;
- б)  $0 \leq \eta \leq +1$ ;
- в)  $0 \leq \eta \leq +2$ .

12. Для измерения степени тесноты связи между двумя качественными признаками с альтернативной вариацией используют ...

- а) коэффициент корреляции рангов Спирмена;
- б) коэффициент Фехнера;
- в) коэффициент ассоциации;
- г) коэффициент контингенции.

13. Оценка статистической значимости параметров уравнения регресс осуществляется на основе ...

- а) коэффициента корреляции;
- б) коэффициента эластичности;
- в)  $t$ -критерия Стьюдента.

14. Укажите, в каких случаях возможно применение коэффициента корреляции рангов Спирмена:

- а) анализируется связь между двумя количественными признаками;
- б) анализируется связь между двумя качественными признаками, каждый из которых измеряется в порядковой шкале;
- в) анализируется связь между количественным и качественным признаком, если последний измеряется в порядковой шкале;
- г) анализируется связь между двумя качественными признаками, если один из них измеряется в порядковой, а другой в номинальной шкале.

### **Задания и задачи по теме**

#### **Задача 10.1**

По данным таблицы, полагая, что зависимость между  $x$  и  $y$  линейная, выполните следующее:

- 1) определите линейный коэффициент корреляции и проверьте его статистическую значимость;
- 2) определите параметры линейного уравнения регрессии;
- 3) рассчитайте среднюю квадратическую ошибку среднюю ошибку аппроксимации для уравнения регрессии;
- 4) проверьте статистическую значимость параметров уравнения регрессии;
- 5) рассчитайте коэффициент эластичности и коэффициент детерминации;
- б) сделайте выводы о надежности полученного уравнения регрессии.



### Исходные данные для расчета

$x$	1	4	7	11	15	17	22
$y$	3	6	10	14	18	24	30

#### Задача 10.2

По следующим данным постройте линейное уравнение регрессии, вычислите линейный коэффициент корреляции:

$$\overline{xy} = 100, \bar{x} = 10, \bar{y} = 8, \bar{x}^2 = 136, \bar{y}^2 = 100, a = 4,8.$$

#### Задача 10.3

Используя следующие данные, определите параметры линейного уравнения ( $a$  и  $b$ ) регрессии:  $\bar{x} = 20$ ,  $\bar{y} = 10$ ,  $\Theta_x = 0,8$ .

#### Задача 10.4

Для оценки степени тесноты связи между уровнем выработки рабочих и стажем их работы было рассчитано корреляционное отношение, оказавшееся равным 0,9. Определите величину средней из внутригрупповых дисперсий, если известно, что общая дисперсия выработки рабочих составляет 6,6.

#### Задача 10.5

Компания «Garden Groceries» владеет 12 магазинами. Финансовый менеджер рассматривает целесообразность слияния ряда мелких магазинов для увеличения прибыли компании. Ему необходимо установить связь между прибылью и объемом продаж. Данные для расчета приведены в таблице:

Номер магазина	Годовая прибыль, тыс. фунт. ст.	Объем продаж за год, тыс. фунт. ст.
1	2	50
2	4	60
3	11	85
4	17	85
5	18	100
6	28	120
7	34	140
8	36	155
9	48	180
10	55	210
11	71	250
12	85	300

Постройте уравнение регрессии для описания связи между прибылью и объемом продаж и проверьте его надежность. Дайте совет финансовому менеджеру по поводу целесообразности слияния магазинов.

## ЛИТЕРАТУРА

1. Теория статистики: учебник / Р.А. Шмойлова [и др.]; под ред. Р.А. Шмойловой. – 4-е изд., перераб. и доп. – М.: Финансы и статистика, 2003. – 656 с.
2. О государственной статистике: Закон Респ. Беларусь от 28.11.2004 г. №345-З.: с изм. и доп. от 4.01.2008 г., №309-З; от 2.07.2009г. №31-З // Нац. реестр правовых актов Респ. Беларусь. – 2009. – № 161. – 2/1583.
3. Статистика: учеб. пособие / под ред. М.Р. Ефимовой. – М.: ИНФРА-М, 2000. – 336 с.
4. Евдропова, В.Н. Общая теория статистики: учебник./ В.Н. Евдропова, М.В. Малафеева. – 2-е изд., перераб. и доп. – М.: Магистр, 2007. – 606 с.
5. Бююль, Ахим SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей /Ахим Бююль, Петер Цёфель; пер. с нем. – СПб.: ООО «ДиаСофтЮП», 2002. – 608 с.
6. Богданович, А.В. Региональные особенности развития демографической ситуации в Республике Беларусь / А.В. Богданович, Н.В. Гордеева // Белорусский экономический журнал. – 2010. – № 1. – С. 50 – 60.
7. Маркусенко, Л.Н. Оценка занятых в сельском хозяйстве Беларуси / Л.Н. Маркусенко, О.А. Пашкевич, Н.А. Пашкевич // Экономический бюллетень НИЭИ Министерства экономики Республики Беларусь. – 2010. – № 6. – С. 22 – 33.
8. Сивенков, О.В. Оценка и перспективы развития сферы бизнес-услуг в Беларуси / О.В. Сивенков // Экономический бюллетень НИЭИ Министерства экономики Республики Беларусь. – 2007. – № 6. – С. 58 – 65.
9. Панкевич, С.П. Иностраные инвестиции в реальный сектор экономики Республики Беларусь в 2006 г. / С.П. Панкевич // Планово-экономический отдел. – 2007. – № 4. – С. 16 – 23.
10. Популярный экономико-статистический словарь-справочник/ под ред. И.И. Елисейевой. – М.: Финансы и статистика, 1993. – 192 с.
11. Беларусь в цифрах // Сайт Национального статистического комитета Республики Беларусь [Электронный ресурс]. – Режим доступа: <http://belstat.gov.by>. – Дата доступа: 8.06.2011.
12. Урланис, Б.Ц. Общая теория статистики / Б.Ц. Урланис. – М.: Статистика, 1973. – 440 с.
13. Громыко, Г.Л. Теория статистики: практикум / Г.Л. Громыко. – 3-е изд., доп. и перераб. – М.: ИНФРА-М, 2004. – 205 с.
14. Теория статистики: учебник/ под ред. проф. Г.Л. Громыко. – 2-е изд., перераб. и доп. – М.: ИНФРА-М, 2009. – 476 с.
15. Статистика: учебник / под ред. И.И. Елисейевой. – М.: Высш. образование, 2009. – 566 с.
16. Елисейева, И.И. Практикум по общей теории статистики: учеб. пособие / И.И. Елисейева, Н.А. Флуд, М.М. Юзбашев; под ред. И.И. Елисейевой. – М.: Финансы и статистика, 2008. – 512 с.
17. Суслов, И.П. Общая теория статистики: учебник / И.П. Суслов. – М.: Статистика, 1978. – 391 с.
18. Статистика: показатели и методы анализа: справ. пособие /Н.Н. Бондаренко [и др.]; под ред. М.М. Новикова. – Минск: Современная школа, 2005. – 628 с.
19. Положение о Едином государственном регистре юридических лиц и индивидуальных предпринимателей: утв. Постановлением Совета Министров Респ. Беларусь от 23.02.2009, № 229. – [Электронный ресурс]. – Режим доступа: <http://www.pravoby.info/docum09/part02/akt02239.htm>. – Дата доступа: 19.06.2011.

20. Инструкция по заполнению в формах государственных статистических наблюдений статистических показателей по труду: утв. постановлением Министерства статистики и анализа Респ. Беларусь от 29.07.2008, № 92: с изм., утв. постановлением Нац. статистического комитета Респ. Беларусь от 14.07.2010, № 118 // Сайт Национального статистического комитета Республики Беларусь [http://belstat.gov.by/homep/ru/statinstrum/main\\_new.php](http://belstat.gov.by/homep/ru/statinstrum/main_new.php). – Рубрика: Государственные статистические наблюдения.

21. Профессии рабочих и должности служащих: общегосударственный классификатор Респ. Беларусь: ОКРБ 006-2009. – Утв. постановлением Министерства труда и социальной защиты Респ. Беларусь 22.10.2009 г., № 125. – 421 с.

22. Трудовой кодекс Республики Беларусь: с изм. и доп. по состоянию на 31.12.2009 г. [Электронный ресурс]. – Режим доступа: <http://www.tamby.info/kodeks/tk.htm>. – Дата доступа: 17. 04.2011.

23. Основные положения по составу затрат, включаемых в себестоимость продукции (работ, услуг): утв. постановлением Министерства экономики Респ. Беларусь, Министерства финансов Респ. Беларусь и Министерства труда и социальной защиты Респ. Беларусь от 30.10.2008 г. №210/161/151 // Экономическая газета. – 2009. – № 10. – С. 6 – 16.

24. О некоторых вопросах регулирования заработной платы: Декрет № 3 Президента Респ. Беларусь от 15.02.2002 г.: с изм. и доп. по состоянию на 10.07.2009 г. – [Электронный ресурс]. – Режим доступа: <http://www.levonevski.net/pravo/norm2009/num33/d33020.html>. – Дата доступа: 25.06.2011.

25. Об установлении и порядке повышения размера минимальной заработной платы: Закон Респ. Беларусь от 17 июля 2002 года № 124-3: с изм. и доп. по состоянию на 01.11.2007 г. – [Электронный ресурс]. – Режим доступа: <http://pravo.levonevsky.org/bazaby/zakon/zakb0497.htm>. – Дата доступа: 17.04.2011.

26. О формировании и использовании минимального потребительского бюджета: Закон Респ. Беларусь от 09.01.1992 г. № 1383-XXII: с изм. и доп. по состоянию на 01.11.2007 г. – [Электронный ресурс]. – Режим доступа: <http://pravo.levonevsky.org/bazaby/zakon/zakb1412.htm>. – Дата доступа: 13.04.2011.

27. О прожиточном минимуме в Республике Беларусь: Закон Респ. Беларусь от 06.01.1999 г. № 239-3: с изм. и доп. по состоянию на октябрь 2009г. – [Электронный ресурс]. – Режим доступа: <http://pravo.levonevsky.org/bazaby/zakon/zakb1412.htm>. – Дата доступа: 25.06.2011.

28. Забродская, Н.Г. Экономика и статистика предприятия: учеб. пособие / Н.Г. Забродская. – М.: Издательство деловой и учебной литературы, 2005. – 352 с.

29. Временный республиканский классификатор основных средств и нормативных сроков их службы: утв. постановлением Министерства экономики Респ. Беларусь от 21 нояб. 2001 г. № 186 // Нац. реестр правовых актов Респ. Беларусь. – 2001. – № 114. – 8/7489.

30. Основные средства и нематериальные активы: общегосударственный классификатор Респ. Беларусь: ОКРБ-020-2002. – Утв. постановлением Комитета по стандартизации, метрологии и сертификации при Совете Министров Респ. Беларусь от 27 сент. 2002 г. № 48. – [Электронный ресурс]. – Режим доступа: <http://tnpa.by/ViewFileText.php?UrlRid=6826&UrlOnd=ОКРБ 020-2002>. – Дата доступа: 26.06.20011.

31. Инструкция по бухгалтерскому учету основных средств: утв. постановлением Министерства финансов Респ. Беларусь 12.12.2001 г. № 118: в ред. постановлений Минфина от 29.12.2007 № 207, от 30.10.2009 № 132, от 26.02.2010 № 15, от 25.06.2010, № 77). – [Электронный ресурс]. – Режим доступа:

[http://www.belgopatent.org.by/russian/docs/economika/pologenie\\_118.do](http://www.belgopatent.org.by/russian/docs/economika/pologenie_118.do). – Дата доступа: 26.06.2011.

32. Инструкция о порядке переоценки основных средств, не завершенных строительством объектов и неустановленного оборудования: утв. постановлениями Министерства экономики Респ. Беларусь, Министерства финансов Респ. Беларусь, Министерства архитектуры и строительства Респ. Беларусь от 05.11.2010 г., №162/131/37 // Нац. реестр правовых актов Респ. Беларусь. – 2011. – № 17. – 8/23296.

33. О лимите стоимости активов, учитываемых в бухгалтерском учете в составе средств в обороте: Постановление Министерства финансов Респ. Беларусь от 23.03.2004 № 41: в ред. от 06.11.2009.

34. Виды экономической деятельности: общегосударственный классификатор Респ. Беларусь: 005-2006. – Утв. постановлением Госстандарта Респ. Беларусь 28.12.2006 № 65: с изм. по состоянию на 1.10.2010. – Сайт Нац. статист. комитета Респ. Беларусь [Электронный ресурс]. – Рубрика «Классификаторы».

35. Указания по заполнению в формах государственной статистической отчетности статистических показателей о производстве промышленной продукции (работ, услуг): утв. постановлением Нац. статист. комитета Респ. Беларусь 20.12.2010, № 268. – Сайт Нац. статист. комитета Респ. Беларусь [Электронный ресурс]. – Рубрика «Формы государственных статистических наблюдений».

36. Номенклатура промышленной продукции для составления государственной статистической отчетности и обработки статистических данных (информации) о производстве продукции в натуральном выражении: утв. постановлением Нац. статист. комитета Респ. Беларусь от 4.03.2010 г. № 25. – Сайт Нац. статистического комитета Респ. Беларусь [Электронный ресурс]. – Рубрика «Формы государственных статистических наблюдений».

37. Инструкция о порядке формирования и применения цен и тарифов: утв. Постановлением Министерства экономики Респ. Беларусь 10.09.2008 № 183: в ред. постановлений Минэкономики от 13.11.2008 № 216, от 31.03.2009 № 60, от 19.08.2009 № 141, от 20.01.2010 № 9, от 01.12.2010 № 179, от 30.12.2010 № 196). – [Электронный ресурс]. – Режим доступа: [http://www.calc.of.by/postanovlenie\\_183.phtml](http://www.calc.of.by/postanovlenie_183.phtml). – Дата доступа: 30.06.2011.

38. Микроэкономическая статистика: учебник / под ред. С.Д. Ильенковой. – М.: Финансы и статистика, 2004. – 544 с.

39. Расходы организации: инструкция по бухгалтерскому учету: утв. постановлением Минфина от 26.12.2003 № 182: с изм. и доп. – [Электронный ресурс]. Режим доступа: <http://www.levonevski.net/pravo/norm2009/num26/d26271.html>. – Дата доступа: 1.07.2011.

40. Инструкция по анализу и контролю за финансовым состоянием и платежеспособностью субъектов предпринимательской деятельности: утв. Постановлением Министерства финансов Респ. Беларусь, Министерства экономики Респ. Беларусь и Министерства статистики и анализа Респ. Беларусь 14.05.2004 № 81/128/65: в редакции по состоянию на 08.05.2008 // Нац. реестр правовых актов Респ. Беларусь. – 2008. – №8/18833 от 04.06.2008. – С. 80 – 100.

41. Руководство Осло: рекомендации по сбору и анализу данных по инновациям: совместная публикация ОЭСР и Евростата / пер. на русский язык. – 2-е изд., испр. – М., 2010. – 107 с. – [Электронный ресурс]. – Режим доступа: [http://www.pnzgu.ru/dep/o\\_niid/files/ruk.oslo.pdf](http://www.pnzgu.ru/dep/o_niid/files/ruk.oslo.pdf). – Дата доступа: 27.07.2011.

42. Отчет об инновационной деятельности организации: указания по заполнению формы государственной статистической отчетности 1-нт (инновация): утв. постановлением Нац. статистического комитета Респ. Беларусь 20.12.2010 г., № 270.

Приложение 1

Значения *F*-критерия Фишера при уровне значимости 0,05

$df_1$ ( $v_1$ ) $df_2$ ( $v_2$ )	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	30	$\infty$
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	250	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41	19,42	49,43	19,44	19,46	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74	8,71	8,69	8,66	8,62	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,87	5,84	5,80	5,74	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68	4,64	4,60	4,56	4,50	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,96	3,92	3,87	3,81	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57	3,52	3,49	3,44	3,38	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28	3,23	3,20	3,15	3,08	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07	3,02	2,98	2,93	2,86	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91	2,86	2,82	2,77	2,70	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79	2,74	2,70	2,65	2,57	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69	2,64	2,60	2,54	2,46	2,30
13	4,67	3,81	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60	2,55	2,51	2,46	2,38	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53	2,48	2,44	2,39	2,31	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48	2,43	2,39	2,33	2,25	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,37	2,33	2,28	2,20	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38	2,33	2,29	2,23	2,15	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,29	2,25	2,19	2,11	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31	2,26	2,21	2,15	2,07	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28	2,23	2,18	2,12	2,04	1,84

Нормальный закон распределения

Значения функции  $\Phi(t) = P(|T| \leq t_{\text{табл}})$   $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-\frac{t^2}{2}} dt$

Целые и десятые доли t	Сотые доли t									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0,0797	0,0876	0,0955	0,1034	0,1113	0,1192	0,1271	0,1350	0,1428	0,1507
0,2	0,1585	0,1663	0,1741	0,1819	0,1897	0,1974	0,2051	0,2128	0,2205	0,2282
0,3	0,2358	0,2434	0,2510	0,2586	0,2661	0,2737	0,2812	0,2886	0,2960	0,3035
0,4	0,3108	0,3182	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,5	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,6	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,7	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,8	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,9	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6679	0,6729	0,6778
1,0	0,6828	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	0,7287	0,7330	0,7373	0,7415	0,7457	0,7499	0,7540	0,7580	0,7620	0,7660
1,2	0,7699	0,7737	0,7775	0,7813	0,7850	0,7887	0,7923	0,7959	0,7994	0,8029
1,3	0,8064	0,8098	0,8132	0,8165	0,8198	0,8230	0,8262	0,8293	0,8324	0,8355
1,4	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,5	0,8664	0,8690	0,8715	0,8740	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,6	0,8904	0,8926	0,8948	0,8969	0,8990	0,9011	0,9031	0,9051	0,9070	0,9090
1,7	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,8	0,9281	0,9297	0,9312	0,9327	0,9342	0,9357	0,9371	0,9385	0,9399	0,9412
1,9	0,9426	0,9439	0,9451	0,9464	0,9476	0,9488	0,9500	0,9512	0,9523	0,9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	0,9643	0,9651	0,9660	0,9668	0,9676	0,9684	0,9692	0,9700	0,9707	0,9715

Окончание прил. 2

2,2	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,9780
2,3	0,9786	0,9791	0,9797	0,9802	0,9807	0,9812	0,9817	0,9822	0,9827	0,9832
2,4	0,9836	0,9841	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,5	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,6	0,9907	0,9910	0,9912	0,9915	0,9917	0,9920	0,9922	0,9924	0,9926	0,9928
2,7	0,9931	0,9933	0,9935	0,9937	0,9938	0,9940	0,9942	0,9944	0,9946	0,9947
2,8	0,9949	0,9951	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,9960	0,9961
2,9	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9880
3,1	0,9981	0,9981	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,5	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997
3,6	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	0,9998	0,9998
3,7	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
4,0	0,999936	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
4,5	0,999994	—	—	—	—	—	—	—	—	—
5,0	0,99999994	—	—	—	—	—	—	—	—	—

**Значения  $\alpha$ -процентных пределов  $t_{\alpha,k}$   
в зависимости от  $k$  степеней свободы и заданного уровня значимости  $\alpha$   
для распределения Стьюдента**

$k \backslash \alpha$	10,0	5,0	2,5	2,0	1,0	0,5	0,3	0,2	0,1
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,600
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,922
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,869
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,408
8	1,860	2,306	2,752	2,696	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3,930	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	3,428	3,610	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,704
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
$\infty$	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291



**Значения коэффициента корреляции рангов Спирмэна  
для двухсторонних пределов уровня значимости  $\alpha$**

$n \backslash \alpha$	0,20	0,10	0,05	0,02	0,01	0,02
4	0,8000	0,8000				
5	0,7000	0,8000	0,9000	0,9000		
6	0,6000	0,7714	0,8286	0,8857	0,9429	
7	0,5357	0,6786	0,7450	0,8571	0,8929	0,9643
8	0,5000	0,6190	0,7143	0,8095	0,8571	0,9286
9	0,4667	0,5833	0,6833	0,7667	0,8167	0,9000
10	0,4424	0,5515	0,6364	0,7333	0,7818	0,8667
11	0,4182	0,5273	0,6091	0,7000	0,7455	0,8364
12	0,3986	0,4965	0,5804	0,6713	0,7273	0,8182
13	0,3791	0,4780	0,5549	0,6429	0,6978	0,7912
14	0,3626	0,4593	0,5341	0,6220	0,6747	0,7670
15	0,3500	0,4429	0,5179	0,6000	0,6536	0,7464
16	0,3382	0,4265	0,5000	0,5824	0,6324	0,7265
17	0,3260	0,4118	0,4853	0,5637	0,6152	0,7083
18	0,3148	0,3994	0,4716	0,5480	0,5975	0,6904
19	0,3070	0,3895	0,4579	0,5333	0,5825	0,6737
20	0,2977	0,3789	0,4451	0,5203	0,5684	0,6586
21	0,2909	0,3688	0,4351	0,5078	0,5545	0,6455
22	0,2829	0,3597	0,4241	0,4963	0,5426	0,6318
23	0,2767	0,3518	0,4150	0,4852	0,5306	0,6186
24	0,2704	0,3435	0,4061	0,4748	0,5200	0,6070
25	0,2646	0,3362	0,3977	0,4654	0,5100	0,5962
26	0,2588	0,3299	0,3894	0,4564	0,5002	0,5856
27	0,2540	0,3236	0,3822	0,4481	0,4915	0,5757
28	0,2490	0,3175	0,3749	0,4401	0,4828	0,5660
29	0,2443	0,3113	0,3685	0,4320	0,4744	0,5567
30	0,2400	0,3059	0,3620	0,4251	0,4665	0,5479

Результаты регрессионного анализа с использованием средств Microsoft Excel

60	77	ВЫВОД ИТОГОВ						
64	91							
77	85	<i>Регрессионная статистика</i>						
83	90	Множественный коэффициент	0,909160086					
95	95	R-квадрат	0,826572062					
100	99	Нормированный R-квадрат	0,80489357					
105	100	Стандартная ошибка	5,524006302					
108	110	Наблюдения	10					
110	105							
118	120	Дисперсионный анализ						
			<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
		Регрессия	1	1163,482835	1163,482835	38,1286694	0,000266649	
		Остаток	8	244,117165	30,51464563			
		Итого	9	1407,6				
			<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
		Y-пересечение	<i>a</i> =44,98516058	8,634606942	5,209867789	0,000812559	25,07372128	64,89659987
		Переменная X 1	<i>b</i> =0,567552602	0,09191371	6,17484165	0,000266649	0,355599208	0,779505997



*Учебное издание*

ГОРДИЕНКО Ольга Ильинична

## СТАТИСТИКА

Учебно-методический комплекс  
для студентов специальности 1-26 02 02  
«Менеджмент»

В двух частях

Часть 1

### **Общая теория статистики**

Редактор *Т. А. Дарьянова*

Дизайн обложки *Л. И. Вайдашевич*

---

Подписано в печать 19.09.12. Формат 60x84 1/16. Бумага офсетная.  
Ризография. Усл. печ. л. 18,33. Уч.-изд. л. 18,04. Тираж 30 экз. Заказ 1347.

---

Издатель и полиграфическое исполнение –  
учреждение образования «Полоцкий государственный университет».

ЛИ № 02330/0548568 от 26.06.09

ЛП № 02330/0494256 от 27.05.09

Ул. Блохина, 29, 211440, г. Новополоцк.