

**Н.А. Богородская, Е.М.Киселева**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ  
к практическим занятиям по  
СТАТИСТИКЕ**

**раздел**

**ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ**

**Уфа 2003**

## Содержание

|               |                                                                              |    |
|---------------|------------------------------------------------------------------------------|----|
| <i>Тема 1</i> | Сводка и группировка.....                                                    | 3  |
| <i>Тема 2</i> | Обобщающие статистические показатели.....                                    | 8  |
| <i>Тема 3</i> | Средние величины.....                                                        | 13 |
| <i>Тема 4</i> | Показатели вариации.....                                                     | 20 |
| <i>Тема 5</i> | Выборочный метод в экономико-статистических исследованиях ...                | 27 |
| <i>Тема 6</i> | Статистическое изучение связи.....                                           | 32 |
| <i>Тема 7</i> | Ряды динамики и их статистический анализ.....                                | 36 |
| <i>Тема 8</i> | Индексы и их использование в экономико-статистических<br>исследованиях ..... | 43 |
|               | Список рекомендуемой литературы.....                                         | 50 |

## Тема 1 СВОДКА И ГРУППИРОВКА.

### 1. Понятия статистической сводки и группировки. Виды группировок

Собранный в результате статистического наблюдения статистический материал подвергается логическому и арифметическому контролю (проверке смысловой согласованности сведений первичного документа и проверке счетной согласованности). Затем приступают к статистической сводке.

*Статистическая сводка* – систематизация единичных фактов, позволяющая перейти к обобщающим показателям, относящимся ко всей изучаемой совокупности и ее частям, и осуществлять анализ и прогнозирование изучаемых явлений и процессов.

Сводка определяет общий размер изучаемого явления по заданным показателям, представляя общие итоги по изучаемой совокупности в целом без какой-либо предварительной систематизации собранного материала.

Статистическая сводка в широком ее понимании предполагает систематизацию и группировку данных, характеристику образованных групп системой показателей, подсчет соответствующих итогов и представление результатов сводки в виде таблиц, графиков.

*Группировка* – это процесс образования однородных групп на основе расчленения статистической совокупности на части или объединение изучаемых единиц в частные совокупности по существенным для них признакам.

Признаки, по которым производится распределение единиц наблюдаемой совокупности на группы, называются *группировочными признаками*, или *основанием группировки*.

С помощью метода группировок решаются задачи: выделение социально-экономических типов явлений; изучение структуры явления и структурных сдвигов, происходящих в нем; выявление связи и зависимости между явлениями. Для решения этих задач применяют соответственно типологические, структурные и аналитические группировки. Данная классификация видов статистических группировок по выполняемым ими задачам имеет несколько условный характер, поскольку на практике они применяются в комплексе.

*Типологическая* группировка - это расчленение разнородной совокупности на отдельные качественно однородные группы и выявление на этой основе экономических типов явления. При использовании метода типологических группировок важное значение имеет правильный выбор группировочного признака. При атрибутивном признаке с незначительным разнообразием его значений число групп определяется свойствами изучаемого явления (например, группировка предприятий по формам собственности). Выделение типов на основе количественного признака состоит в определении групп с учетом значений изучаемых признаков.

*Структурная* группировка предназначена для изучения состава однородной совокупности по какому-либо варьирующему признаку. Другими словами, выделенные с помощью типологической группировки типы явления могут изучаться с точки зрения их структуры и состава. Однако нередко структурные группировки применяются и без предварительного расчленения совокупности на части.

Для изучения связи между отдельными признаками явления используются *аналитические группировки*.

Образование групп по двум и более признакам называется *комбинированной группировкой*.

### 2. Построение статистических группировок

1. Выбор группировочного признака – признака, по которому производится разбиение совокупности на отдельные группы. В качестве признака необходимо использовать существенные обоснованные признаки.

По форме выражения группировочные признаки бывают *атрибутивными* (не имеющими количественного выражения, например, профессия) и *количественными* (например, число филиалов, величина дохода). При этом количественные признаки могут быть *дискретными* (прерывными, значения которых выражаются только целыми числами, например, число филиалов) и *непрерывными* (принимающими как целые, так и дробные значения, например, величина дохода).

По характеру колеблемости группировочные признаки бывают *альтернативными*, которыми одни единицы обладают, а другие – нет (например, товары – качественные или некачественные), и имеющими множество количественных значений (например, число филиалов, величина дохода).

По роли во взаимосвязи изучаемых явлений признаки подразделяются на *факторные*, воздействующие на другие признаки, и *результативные*, испытывающие на себе влияние других.

2. Выбор количества групп. Если в основание группировки положен атрибутивный признак, то количество групп будет столько, сколько существует градаций (уровней) данного признака. Если основание группировки – количественный признак, то необходимо обратить внимание на число единиц исследуемого объекта и степень колеблемости группировочного признака. В каждом конкретном случае следует исходить не только из степени колеблемости признака, но и из особенностей объекта и цели исследования. Если совокупность состоит из большого числа единиц и распределение единиц по группировочному признаку близко к нормальному, используют формулу Стерджесса:

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

3. Определение интервала группировки. *Интервал* – это значение варьирующего признака, лежащее в определенных границах. Под *величиной* интервала понимают разность между максимальным и минимальным значениями признака в группе. При этом максимальное значение признака в группе называется *верхней* границей интервала, а минимальное – *нижней* границей. В зависимости от степени колеблемости группировочного признака, характера распределения статистической совокупности устанавливаются интервалы равные или неравные. Если вариация признака происходит в сравнительно узких границах и распределение носит равномерный характер, то строят группировку с равными интервалами; величина интервала определяется по формуле:

$$h = \frac{x_{\max} - x_{\min}}{n}$$

где  $x_{\max}$  – максимальное значение признака в изучаемой совокупности

$x_{\min}$  – минимальное значение признака в изучаемой совокупности

$n$  – количество групп

В экономической практике чаще применяются неравные интервалы, прогрессивно возрастающие или убывающие. Такая необходимость возникает, когда колеблемость признака осуществляется неравномерно и в больших пределах.

### 3. Статистические ряды распределения

Результаты сводки и группировки материалов статистического наблюдения оформляются в виде статистических рядов распределения и таблиц. *Ряд распределения* – это упорядоченное распределение единиц совокупности на группы по определенному признаку. Другими словами, это группировка, в которой для характеристики групп применяется численность группы.

*Атрибутивные* ряды распределения – ряды распределения, построенные по качественным признакам.

*Вариационные ряды распределения* – ряды распределения, построенные по количественным признакам. Вариационный ряд состоит из двух элементов: варианты и частота. *Варианта* (обозначается  $x$ ) – отдельное значение варьирующего признака, которое он принимает в ряду распределения. *Частота* (обозначается  $f$ ) – численность отдельных вариантов, т.е. частота повторения каждого варианта. Частота, выраженная в долях единицы или в процентах к итогу, называется *частотой* (обозначается  $w$ ).

По способу построения вариационные ряды бывают дискретными и интервальными.

*Дискретный* вариационный ряд характеризует распределение единиц совокупности по дискретному признаку, принимающему только целые значения. Для его построения следует перечислить все встречающиеся варианты значений признака и подсчитать частоту повторения. При графическом изображении дискретных вариационных рядов используется *полигон распределения*, или *полигон частот*. Для его построения в прямоугольной системе координат по оси абсцисс в одинаковом масштабе откладываются ранжированные значения варьирующего признака, а по оси ординат наносится шкала для выражения величины частот. Полученные на пересечении абсцисс и ординат точки соединяются прямыми линиями, в результате чего получают ломаную линию.

*Интервальный* вариационный ряд строится в случае непрерывной вариации признака у единиц совокупности (величина может принимать в определенных пределах любые значения, отличающиеся друг от друга на сколь угодно малую величину), а также в случае, когда число вариантов дискретного признака достаточно велико. Для графического изображения интервального вариационного ряда применяется *гистограмма*. При построении гистограммы на оси абсцисс откладываются величины интервалов, а частоты изображаются прямоугольниками, построенным на соответствующих интервалах. Высота столбиков должна быть пропорциональна частотам. В результате мы получим график, на котором ряд распределения изображен в виде смежных друг с другом столбиков.

В ряде случаев для изображения вариационных рядов (как дискретным, так и интервальным) используется *кумулятивная кривая* (или *кумулята*). Для ее построения надо рассчитать накопленные частоты или частоты. Накопленные частоты (обозначаются  $S$ ) показывают, сколько единиц совокупности имеют значение признака не больше, чем рассматриваемое, и определяются последовательным суммированием частот интервалов. При построении кумуляты интервального ряда распределения нижней границе первого интервала соответствует частота, равная нулю, а верхней границе – частота данного интервала.

#### 4. Примеры решения задач

**Пример 1.** Пользуясь формулой Стерджесса, определите интервал группировки сотрудников фирмы по уровню доходов, если общая численность сотрудников составляет 120 человек, а минимальный и максимальный доход соответственно равен 500 и 6500 руб.

**Решение.**

Количество групп равно  $n = 1 + 3,322 \cdot \lg 120 = 8$

Величина интервала  $h = \frac{x_{\max} - x_{\min}}{n} = \frac{6500 - 500}{8} = 750$  руб.

Интервалы выглядят следующим образом:

| № группы | Величина интервала группировки |
|----------|--------------------------------|
| 1        | 500-1250                       |
| 2        | 1250-2000                      |

|   |           |
|---|-----------|
| 3 | 2000-2750 |
| 4 | 2750-3500 |
| 5 | 3500-4250 |
| 6 | 4250-5000 |
| 7 | 5000-5750 |
| 8 | 5750-6500 |

**Пример 2.** Имеются следующие данные о количестве филиалов каждого из двадцати банков в городе.

Количество филиалов в городе у разных банков: 2, 4, 3, 5, 4, 4, 6, 5, 4, 3, 4, 3, 4, 5, 3, 4, 6, 3, 5, 4  
Построить ряд распределения по имеющимся данным. Дать графическое изображение ряда распределения.

**Решение.**

Вариация признака носит дискретный характер, число вариантов дискретного признака невелико, и значения признака у отдельных единиц совокупности повторяются. Поэтому строится дискретный ряд распределения. Для его построения следует перечислить все встречающиеся варианты значений признака и подсчитать частоту повторения.

Дискретный ряд распределения, построенный по данным, выглядит следующим образом

| Количество филиалов в городе организации, $x$ | Число банков (или частота, $f$ ) | Частость, $w$ | Накопленная частота, $S$ |
|-----------------------------------------------|----------------------------------|---------------|--------------------------|
| 2                                             | 1                                | $1/20=0,05$   | 1                        |
| 3                                             | 5                                | $5/20=0,25$   | $1+5 = 6$                |
| 4                                             | 8                                | $8/20=0,40$   | $6+8 = 14$               |
| 5                                             | 4                                | $4/20=0,20$   | $14+4 = 18$              |
| 6                                             | 2                                | $2/20=0,10$   | $18+2 = 20$              |
| Итого                                         | 20                               | 1,00          |                          |

Частость  $w$  рассчитана как отношение соответствующей частоты к общей сумме частот.

По полученному дискретному ряду распределения строится полигон частот.

Для построения кумуляты следует рассчитать накопленные частоты  $S$ . Накопленная частота первой варианты равна частоте первого интервала, т.е. всего 1 банк в городе имеет не больше двух филиалов. Накопленная частота второй варианты равна сумме частот первой и второй вариант (или сумме накопленной частоты первой варианты и частоты второй варианты), т.е. не больше трех филиалов имеют 6 городских банков: у пяти из них по 3 филиала, у одного – 2 филиала. Остальные накопленные частоты определяются аналогично. Накопленная частота последней варианты равна сумме всех частот ряда: все банки в городе имеют не больше 6 филиалов.

**Пример 3.** Имеются следующие данные о размере прибыли двадцати коммерческих банков. Прибыль, млн. руб.:

3,7 4,3 6,7 5,6 5,1 8,1 4,6 5,7 6,4 5,9 5,2 6,2 6,3 7,2 7,9 5,8 4,9 7,6 7,0 6,9

Построить ряд распределения по имеющимся данным. Дать графическое изображение ряда распределения.

**Решение.** Вариация признака носит непрерывный характер, значения признака у отдельных единиц совокупности не повторяются. Поэтому строится интервальный ряд распределения. Для его построения следует определить количество интервалов и величину интервала.

Т.к. количество интервалов заранее не задано, определим его по формуле Стерджесса:  $n=1+3,322 \cdot \lg 20=1+3,322 \cdot 1,3= 5,3$  Дробное число, характеризующее количество интервалов, желательно округлять в меньшую сторону. Т.о.,  $n=5$

Величина интервала  $h=(8,1-3,7)/5=0,88$  Число, характеризующее величину интервала, округляется с той же точностью, что и исходные данные. В нашем случае следует округлить до 0,1:  $h=0,9$ .

Строим интервальный ряд распределения:

| № группы | Группы по размеру прибыли $x$ | Число банков (частота) $f$ | Частость, $w$ | Накопленная частота $S$ |
|----------|-------------------------------|----------------------------|---------------|-------------------------|
| 1        | 3,7 – 4,6                     | 3                          | 0,15          | 3                       |
| 2        | 4,6 – 5,5                     | 3                          | 0,15          | 6                       |
| 3        | 5,5 – 6,4                     | 7                          | 0,35          | 13                      |
| 4        | 6,4 – 7,3                     | 4                          | 0,2           | 17                      |
| 5        | 7,3 – 8,2                     | 3                          | 0,15          | 20                      |
| Итого    |                               | 20                         | 1             |                         |

При подсчете частот воспользуемся принципом «включительно», согласно которому единица совокупности, имеющая значение признака, равное границе двух смежных групп (например, банк с прибылью 4,6 млн. руб.), включается в интервал, где он служит верхней границей (банк с прибылью 4,6 млн. руб. включим в группу с размером прибыли от 3,7 до 4,6 млн. руб.).

Расчет частостей и накопленных частот производится аналогично расчету в дискретных рядах распределения.

По полученным значениям частот строится гистограмма распределения, по накопленным частотам – кумулята.

### 5. Задачи для самостоятельного решения

**Задача 1.** Пользуясь формулой Стерджесса, определите интервалы групп, полученных в результате группировки работников магазина по среднемесячной выработке, если общая численность работников составляет 22 человека, а минимальная и максимальная среднемесячная выработка соответственно равны 100 тыс. руб. и 250 тыс. руб.

**Задача 2.** Имеются следующие данные о числе товарных секций по двадцати магазинам города:

Количество товарных секций в магазине:

|   |   |   |   |
|---|---|---|---|
| 2 | 4 | 3 | 5 |
| 5 | 6 | 4 | 6 |
| 2 | 2 | 4 | 3 |
| 4 | 5 | 5 | 4 |
| 6 | 3 | 3 | 4 |

Построить ряд распределения по имеющимся данным.

Дать графическое изображение ряда распределения.

**Задача 3.** Имеются следующие данные о размере прибыли двадцати коммерческих банков.

Прибыль, млн. руб.:

|     |     |     |     |
|-----|-----|-----|-----|
| 4,7 | 9,1 | 6,2 | 6,8 |
| 5,3 | 5,6 | 7,2 | 5,9 |
| 7,7 | 6,7 | 7,3 | 8,6 |
| 6,6 | 7,4 | 8,2 | 8   |
| 6,1 | 6,9 | 8,9 | 7,9 |

Построить ряд распределения по имеющимся данным. Дать графическое изображение ряда распределения.

## Тема 2

### ОБОБЩАЮЩИЕ СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

Статистическое исследование независимо от его масштабов и целей всегда завершается расчетом и анализом различных по виду и форме выражения статистических показателей.

*Статистический показатель* представляет собой количественную характеристику социально-экономических явлений и процессов в условиях качественной определенности. Качественная определенность показателя заключается в том, что он непосредственно связан с внутренним содержанием изучаемого процесса, его сущностью.

В отличие от признака статистический показатель получается расчетным путем. Это может быть простой подсчет единиц совокупности, суммирование их значений признака, сравнение двух или нескольких величин или более сложные расчеты. Признак – это свойство, присущее единице совокупности. Признак входит в качественное содержание показателя, он существует объективно. Показатель – характеристика группы единиц или совокупности в целом; его построение зависит от цели исследования

Все статистические показатели по охвату единиц совокупности разделяются на индивидуальные и сводные по форме выражения – на *абсолютные, относительные и средние*.

#### 1. Абсолютные показатели

Исходной, первичной формой выражения статистических показателей являются абсолютные величины. Статистические показатели в форме абсолютных величин характеризуют абсолютные размеры изучаемых статистикой процессов и явлений: их массу, площадь, объем, протяженность; отражают их временные характеристики, а также могут представлять объем совокупности, т.е. число составляющих ее единиц. В отличие от математического понятия абсолютной величины, абсолютные показатели в статистике могут быть представлены как положительными, так и отрицательными числами.

*Индивидуальные* абсолютные показатели, как правило, получают непосредственно в процессе статистического наблюдения как результат замера, взвешивания, подсчета и оценки интересующего количественного признака.

*Сводные* объемные показатели, характеризующие объем признака или объем совокупности как в целом по изучаемому объекту, так и по какой-либо его части, получают в результате сводки и группировки индивидуальных значений.

Абсолютные статистические показатели всегда являются именованными числами. В зависимости от социально-экономической сущности исследуемых явлений, их физических свойств они выражаются в натуральных (тонны, килограммы, метры, штуки), условно-натуральных (так, различные виды топлива переводят в условное топливо с определенной теплотой сгорания; перевод в условные единицы осуществляется на основе специальных коэффициентов), стоимостных или трудовых (человеко-дни, и человеко-часы) единицах измерения.

#### 2. Относительные показатели

*Относительный показатель* в статистике – это обобщающий показатель, который дает числовую меру соотношения двух сопоставляемых абсолютных величин.

Величина, с которой производится сравнение (знаменатель дроби), называется *базой сравнения* или *основанием*. В зависимости от базы сравнения относительный показатель может быть представлен в различных долях единицы, процентах, промилле, продецимилле и т.д. По способу получения относительные величины – всегда производные, результат отношения может быть выражен либо в форме коэффициента и процента, либо в форме промилле и продецимилле. Существуют также именованные относительные величины (например, показатель фондоотдачи).

Общие *принципы* построения относительных показателей.

- 1) Сравнимые в относительном показателе абсолютные (или, в свою очередь, относительные) показатели должны быть объективно связаны в реальной жизни.
- 2) При построении относительного статистического показателя сравниваемые исходные показатели могут различаться только одним атрибутом: или видом признака (при одинаковом объекте, периоде времени, плановом или фактическом характере показателей), или временем (при том же признаке, объекте и т.п.), или только фактическим, плановым, нормативным характером показателей (при том же объекте, признаке, периоде времени) и т.д. Нельзя сопоставлять показатели, различные по двум или более атрибутам (например, добычу угля в США в 1980 г. с выплавкой стали в России в 1992 г.).
- 3) Необходимо знать возможные границы существования относительного показателя. Например, если исходные показатели в текущем и базисном периодах имеют разные знаки, то теряет смысл и не может применяться относительная величина динамики

По своему содержанию относительные величины подразделяются на следующие виды:

1). *Относительный показатель динамики (ОПД)* характеризует изменение уровня развития явления во времени. Представляет собой отношение уровня исследуемого процесса или явления за данный период времени (по состоянию на данный момент времени) к уровню этого же процесса или явления в прошлом.

Обозначим уровень показателя через  $y$ :

$y_0$  – уровень показателя в базисном периоде,

$y_1$  – уровень показателя в отчетном периоде

$$ОПД = y_1 / y_0$$

Относительная величина динамики может быть представлена в трех формах: коэффициента (индекса), темпов роста либо прироста.

Показатели динамики могут определяться с использованием постоянной либо переменной базы сравнения. При расчете показателей на постоянной базе каждый уровень сравнивается с одним и тем же базисным уровнем, т.е. вычисляются делением сравниваемого уровня ( $y_i$ ) на уровень, принятый за постоянную базу сравнения,  $y_{o_i}$ :

$$ОПД_{o_i} = y_i : y_{o_i}$$

Исчисляемые при этом показатели называются базисными.

Для расчета показателей на переменной базе каждый последующий уровень сравнивается с предыдущим, т.е. вычисляются делением сравниваемого уровня  $y_i$  на предыдущий уровень  $y_{i-1}$ :

$$ОПД_{y_i} = y_i : y_{i-1}$$

Вычисленные таким образом показатели называются цепными.

Между базисными и цепными относительными показателями динамики имеется взаимосвязь: произведение последовательных цепных относительных показателей динамики равно базисной величине, исчисленной за тот же период, а частное от деления последующего базисного темпа роста на предыдущий равно соответствующему цепному темпу роста.

2). *Относительный показатель планового задания (ОПП)* рассчитывается как отношение уровня, запланированного на предстоящий период, к уровню, фактически сложившемуся в предшествующем периоде.

$$ОПП = y_{пл} / y_0$$

3). *Относительный показатель выполнения задания (ОПВП)* рассчитывается как отношение фактически достигнутого в данном периоде уровня к запланированному.

$$ОПВП = y_1 / y_{пл.}$$

Относительные величины динамики, планового задания и выполнения связаны соотношением

$$ОПД = ОПП * ОПВП$$

4. *Относительные показатели структуры* (ОПС) характеризуют доли, удельные веса составных элементов в общем итоге. Как правило, в форме процентного содержания.

Обозначим через  $Y$  уровень части совокупности,  $\Sigma Y$  – суммарный уровень совокупности

$$ОПС = \frac{Y}{\Sigma Y} * 100$$

Расчет относительных величин структуры за несколько периодов позволяет выявить структурные сдвиги.

Показатели структуры используют для выявления соотношения части и целого.

5. *Относительные показатели координации* (ОПК) характеризуют отношений частей данной совокупности к одной из них, принятой за базу сравнения. В качестве базы сравнения как правило, выбирается та часть, которая имеет наибольший удельный вес или является приоритетной с экономической, социальной или какой-либо другой точки зрения. Относительные величины координации показывают, во сколько раз одна часть совокупности больше другой, либо сколько единиц одной части приходится на 1, 10, 100, 1000 ... единиц другой части.

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

6. *Относительные показатели сравнения (наглядности)* характеризуют результаты сопоставления одноименных абсолютных величин, относящихся к одному и тому же периоду либо моменту времени, но к различным объектам или территориям.

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

7. *Относительные показатели интенсивности* (ОПИ) характеризуют степень распределения или развития данного явления в той или иной среде. Это отношение абсолютного уровня одного показателя, свойственного изучаемой среде, к другому абсолютному показателю, также присущему данной среде, и, как правило, являющемуся для первого показателя факторным признаком. Напр., показатели рождаемости, смертности, естественного прироста, которые рассчитываются как отношение к среднегодовой численности населения данной территории (на 1000 чел.).

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

В отличие от относительных показателей, получаемых в результате сопоставления одноименных показателей и представляемых в виде коэффициентов и процентов, относительные показатели интенсивности являются именованными числами. Относительными показателями интенсивности выступают, напр., показатели выработки продукции в единицу рабочего времени, затрат на единицу продукции, трудоемкости, эффективности использования производственных фондов и т.д.

Разновидностью относительных показателей интенсивности являются *относительные показатели уровня экономического развития*, характеризующие производство продукции в расчете на душу населения и играющие важную роль в оценке развития экономики государства.

### 3. Примеры решения задач

**Пример 1.** Уставный капитал банка в 1998 г. составлял 5,08 млн. руб., а в 2001 г. – 6,15 млн. руб. Найти относительную величину динамики.

**Решение.**

$$ОПД = 6,15 / 5,08 = 1,211$$

т.е. размер уставного фонда вырос за 3 года в 1,211 раза – это коэффициент роста (или индекс роста). В процентном выражении это 121,1% - это темп роста. За три года размер уставного фонда увеличился на 21,1% - это темп прироста

**Пример 2.** По плану на 2000 год предполагалось увеличить производство продукции с 5650 шт. до 6100 шт. В действительности в 2000 году было произведено продукции 5850 шт. Найти относительные величины планового задания, выполнения планового задания.

**Решение.**

$$ОПП = 6100 / 5650 = 1,08$$

т.е. по плану предполагалось увеличить производство продукции в 1,08 раза,  
это - плановый коэффициент роста (плановый индекс роста).

В процентном выражении это 108% - это плановый темп роста

т.е. планировалось увеличить пр-во на 8% - это плановый темп прироста

В действительности в 2000 году было произведено продукции 5850 шт. при плане 6100 шт.

$$ОПВП = 5850 / 6100 = 0,959, \text{ или } 95,9 \%$$

т.е. плановое задание было невыполнено на 4,1%

$$\text{Фактический ОПД составил } ОПД = ОПП * ОПВП = 1,08 * 0,959 = 1,035, \text{ или } 103,5\%$$

(или  $ОПД = 5850/5650 = 1,035$ , или 103,5%)

**Пример 3.** Внешнеторговый оборот России в 1997-1998 годах характеризовался следующими данными

| Период       | Внешнетор-говый оборот,<br>всего, млрд. долл. | В том числе |        |
|--------------|-----------------------------------------------|-------------|--------|
|              |                                               | Экспорт     | Импорт |
| 1997 г.      |                                               |             |        |
| I кв.        | 36,7                                          | 21,1        | 15,6   |
| II кв.       | 37,9                                          | 20,4        | 17,5   |
| III кв.      | 40,4                                          | 21,6        | 18,8   |
| IV кв.       | 46,9                                          | 25,1        | 21,8   |
| Итого за год | 161,9                                         | 88,2        | 73,7   |
| 1998 год     |                                               |             |        |
| I кв.        | 36,7                                          | 18,4        | 18,3   |
| II кв.       | 36,4                                          | 18,7        | 17,7   |
| III кв.      | 31,5                                          | 17,8        | 13,7   |
| IV кв.       | 28,7                                          | 19,3        | 9,4    |
| Итого за год | 133,3                                         | 74,2        | 59,1   |

а) Рассчитать относительные величины структуры, характеризующие доли экспорта и импорта во внешнеторговом обороте России.

б) Рассчитать относительные величины координации, характеризующие соотношение экспорта и импорта.

**Решение.** Относительные показатели структуры используют для выявления соотношения части и целого. В нашем случае целое – это внешнеторговый оборот, его части – экспорт и импорт, т.е. требуется сопоставить величины экспорта (импорта) и внешнеторгового оборота в целом.

| Период       | Внешнетор. оборот, всего, млрд. долл. | В том числе |        | Удельный вес, %  |                  | Стоимость импорта на 1000 руб. экспорта |
|--------------|---------------------------------------|-------------|--------|------------------|------------------|-----------------------------------------|
|              |                                       | Экспорт     | Импорт | Экспорта         | Импорта          |                                         |
| 1997 г.      |                                       |             |        |                  |                  |                                         |
| I кв.        | 36,7                                  | 21,1        | 15,6   | 21,1/36,7=57,49  | 15,6/36,7=42,51  | (15,6/21,1)*1000=739                    |
| II кв.       | 37,9                                  | 20,4        | 17,5   | 20,4/37,9=53,83  | 17,5/37,9=46,17  | (17,5/20,4)*1000=858                    |
| III кв.      | 40,4                                  | 21,6        | 18,8   | 21,6/40,4=53,47  | 18,8/40,4=46,53  | (18,8/21,6)*1000=870                    |
| IV кв.       | 46,9                                  | 25,1        | 21,8   | 25,1/46,9=53,52  | 21,8/46,9=46,48  | (21,8/25,1)*1000=869                    |
| Итого за год | 161,9                                 | 88,2        | 73,7   | 88,2/161,9=54,48 | 73,7/161,9=45,52 | (73,7/88,2)*1000=836                    |
| 1998 год     |                                       |             |        |                  |                  |                                         |
| I кв.        | 36,7                                  | 18,4        | 18,3   | 50,14            | 49,86            | 995                                     |
| II кв.       | 36,4                                  | 18,7        | 17,7   | 51,37            | 48,63            | 947                                     |
| III кв.      | 31,5                                  | 17,8        | 13,7   | 56,51            | 43,49            | 770                                     |
| IV кв.       | 28,7                                  | 19,3        | 9,4    | 67,25            | 32,75            | 487                                     |
| Итого за год | 133,3                                 | 74,2        | 59,1   | 55,66            | 44,34            | 796                                     |

Относительные показатели координации характеризуют соотношение частей целого между собой, т.е. требуется сопоставить величины импорта и экспорта между собой.

При расчете относительной величины координации за базу сравнения принимаем величину экспорта как показатель, обладающий большим социально-экономическим значением и большей величиной. Найдем, сколько импорта приходится на 100 р. экспорта.

**Пример 4.** Объем кредитов, выданный банками предприятиям, в области *A* составил 73,2 млн. руб., а в области *B* – 38,8 млн. руб. Рассчитайте относительную величину сравнения.

**Решение.**  $ОПС = 38,8/73,2=0,53$ . Т.о. уровень кредитования банками предприятий в области *B* составляет от уровня области *A* 53%.

**Пример 5.** Производство электроэнергии в области составило 17,2 млрд. квт.-ч. при среднегодовой численности населения 8,4 млн. чел. Определить относительную величину интенсивности, характеризующую производство электроэнергии на душу населения.

**Решение.**  $ОПИ=17,2 / 8,4 = 2,05$  тыс. квт.-ч. на душу населения.

#### 4. Задачи для самостоятельного решения

**Задача 1.** Найти относительные величины динамики, планового задания и выполнения планового задания по следующим данным. Сделать выводы по полученным результатам.. Показать взаимосвязь показателей

|                                          |     |
|------------------------------------------|-----|
| Выпуск продукции в базисном периоде, шт. | 120 |
| Плановое задание, шт.                    | 134 |
| Выпуск в отчетном периоде, шт.           | 127 |

**Задача 2.** Найти относительные величины структуры и координации по данным, характеризующим структуру ВВП страны *A*. Найти относительные величины интенсивности и сравнения.

|                                                                |       |
|----------------------------------------------------------------|-------|
| ВВП страны <i>A</i> , млрд.долл.                               | 508,0 |
| в том числе                                                    |       |
| производство товаров                                           | 185,4 |
| производство услуг                                             | 277,9 |
| Среднегодовая численность населения страны <i>A</i> , млн.чел. | 90,0  |
| ВВП страны <i>B</i> , млрд.долл.                               | 600,0 |

#### Тема 3

## СРЕДНИЕ ВЕЛИЧИНЫ

*Средние* величины – это обобщающие показатели, в которых находят выражение действие общих условий, закономерность изучаемого явления. Сущность средней величины состоит в том, что она отражает общие черты, закономерности, тенденции, присущие данной совокупности, погашая влияние индивидуальных (случайных факторов) и поэтому является обобщающей характеристикой варьирующего признака качественно однородной совокупности.

Признак, по которому находится средняя, называется *осредняемым признаком*, обозначается  $\bar{x}$ .

Все виды средних величин, используемые в статистических исследованиях, подразделяются на 2 категории: степенные и структурные.

### 1. Степенные средние

Наиболее распространены следующие виды степенных средних:

- средняя арифметическая
- средняя гармоническая
- средняя геометрическая
- средняя квадратическая

✓ *Средняя арифметическая* исчисляется в тех случаях, когда объем осредняемого признака образуется как сумма его значений у отдельных единиц изучаемой статистической совокупности.

Некоторые свойства средней арифметической:

1. Сумма отклонений индивидуальных значений признака от их средней величины равна нулю.

$$\sum (x_i - \bar{x}) = 0$$

2. Сумма квадратов отклонений индивидуальных значений признака от их средней величины есть величина минимальная.

$$\sum (x_i - \bar{x})^2 < \sum (x_i - A)^2, \text{ где } A = \bar{x} \pm \varepsilon \text{ (т.е. } A - \text{любое число, отличное от } \bar{x} \text{)}$$

3. Если все частоты разделить на одно и то же число, средняя арифметическая останется без изменений. Т.е. для расчета средней можно воспользоваться не только значениями частот, но и значениями частостей.

Средняя арифметическая простая применяется в случаях, когда имеются отдельные значения признака, т.е. данные не сгруппированы. Равна сумме отдельных значений признака, деленной на число этих значений.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum x}{n}$$

Средняя арифметическая взвешенная в дискретном ряду распределения применяется в случаях, когда данные представлены в виде рядов распределения или группировок. Одни и те же значения признака повторяются несколько раз.

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum xf}{\sum f}$$

где  $f$  - число одинаковых значений признака в рядах распределения, т.е. частота, или вес.

Средняя арифметическая взвешенная зависит не только от значений признака, но и от частот, т.е. от состава совокупности, от ее структуры.

Средняя арифметическая взвешенная в интервальном ряду распределения. В интервальном ряду распределения с закрытыми интервалами варианты осредняемого признака представлены не одним числом, а виде интервала «от - до». Таким образом, каждая группа ряда распределения имеет нижнее и верхнее значение вариантов. Исчисление средней

по сгруппированным данным производится по формуле средней арифметической взвешенной.

$$\bar{x} = \frac{\sum xf}{\sum f}$$

Чтобы применить эту формулу, варианты признака надо выразить одним числом (дискретным). За такое число принимается средняя арифметическая простая из верхнего и нижнего значения интервала. Дальнейший расчет производится обычным методом определения средней арифметической взвешенной.

- ✓ *Средняя гармоническая* – это величина, обратная средней арифметической. Когда статистическая информация не содержит частот по отдельным вариантам, а представлена как их произведение, применяется формула средней гармонической взвешенной. В том случае, когда объемы явлений (т.е. произведения) по каждому признаку равны, применяется средняя гармоническая простая.

Средняя гармоническая простая

$$x_{\text{гарм}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum \frac{1}{x}}$$

Средняя гармоническая взвешенная

$$x_{\text{гарм.}} = \frac{B + B + \dots + B}{\frac{B_1}{x_1} + \frac{B_2}{x_2} + \dots + \frac{B_n}{x_n}} = \frac{\sum B}{\sum \frac{B_i}{x_i}}$$

- ✓ *Средняя геометрическая* – это величина, используемая как средняя из отношений. Этой средней удобно пользоваться, когда уделяется внимание не абсолютным разностям, а отношениям двух чисел, т.е. когда индивидуальные значения признака – относительные величины. Например, средняя геометрическая используется при расчете среднего коэффициента роста.

Средняя геометрическая простая

$$\bar{x} = \sqrt[n]{x_1 * x_2 * \dots * x_n} = \sqrt[n]{\prod x}$$

## 2. Структурные средние

Для характеристики структуры совокупности применяются особые показатели, которые называются структурными средними. К таким показателям относятся мода и медиана.

*Мода* – величина признака (варианта), наиболее часто повторяющаяся в изучаемой совокупности. Другими словами, модой называется то значение признака, которое соответствует максимальной точке теоретической кривой распределения. Мода отражает типичный, наиболее распространенный вариант значения признака.

В дискретном ряду распределения мода – это варианта, которой соответствует наибольшая частота.

В интервальном ряду распределения сначала определяют модальный интервал (т.е. интервал, содержащий моду), которому соответствует наибольшая частота. Конкретное значение моды определяется формулой:

$$Mo = x_{Mo} + i_{Mo} \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}$$

$x_{Mo}$  – начальное значение модального интервала

$i_{Mo}$  – величина модального интервала

$f_{Mo}$  – частота модального интервала

$f_{Mo-1}$  – частота интервала, предшествующего модальному

$f_{Mo+1}$  – частота интервала, следующего за модальным

При этом мода будет несколько неопределенной, т.к. ее значение будет зависеть от величины групп, точного положения границ групп.

**Медиана** – это величина, которая делит численность упорядоченного вариационного ряда на две равные части: одна часть имеет значения, не большие, чем средний вариант, а другая – не меньшие. Справедливо соотношение: сумма абсолютных отклонений членов ряда от медианы есть величина наименьшая

$$\sum |x - Me| < \sum |x - A|, \text{ где } A = Me \pm \varepsilon \text{ (т.е. } A \text{ – любое число, отличное от } Me)$$

Для ранжированного (выстроенного в порядке возрастания или убывания значения признака) ряда с нечетным числом членов медианой является варианта, расположенная в центре ряда. Для ранжированного ряда с четным числом членов медиана рассчитывается как средняя арифметическая из двух вариантов, расположенных в центре ряда.

Для дискретного ряда медиана рассчитывается с помощью накопленных частот: медианой является варианта, которой соответствует накопленная частота, впервые превысившая половину общей суммы частот.

Для интервального ряда с помощью накопленных частот определяют медианный интервал (т.е. интервал, содержащий медиану), которому соответствует накопленная частота, впервые превысившая половину общей суммы частот. Затем конкретное значение медианы рассчитывают по формуле

$$Me = x_{Me} + i_{Me} \frac{\frac{\sum f}{2} - S_{Me-1}}{f_{Me}},$$

где

$x_{Me}$  – начальное значение медианного интервала

$i_{Me}$  – величина медианного интервала

$S_{Me-1}$  – сумма накопленных частот, предшествующих медианному интервалу

$f_{Me}$  – частота медианного интервала

Моду и медиану можно также определить графически.

Мода определяется по полигону (рис.1) или гистограмме (рис.2) распределения. В первом случае мода соответствует наибольшей ординате. Во втором – правую вершину модального прямоугольника соединяют с правым углом предыдущего прямоугольника, а левую вершину – с левым углом последующего прямоугольника. Абсцисса точки пересечения – этих прямых будет модой распределения.

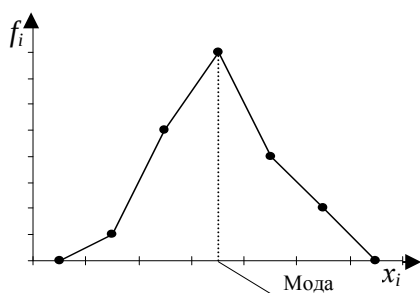


Рис.1

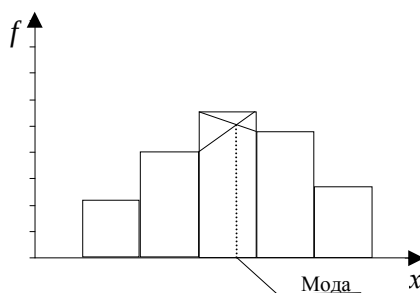


Рис.2

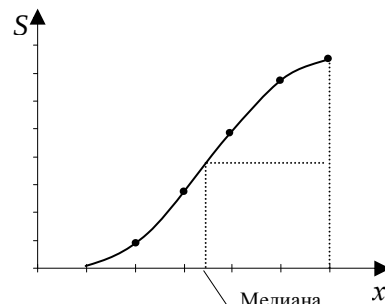


Рис.3

Медиана определяется по кумуляте (рис.3). Для ее определения высоту наибольшей ординаты, которая соответствует общей численности совокупности, делят пополам. Через полученную точку проводят прямую, параллельную оси абсцисс, до пересечения ее с кумулятой. Абсцисса точки пересечения является медианой.

### 3. Примеры решения задач

**Задача 1.** По имеющимся данным о ценах товара в различных фирмах города определить среднюю цену.

4,4 4,3 4,4 4,5 4,3 4,3 4,6 4,2 4,6 4,1

**Решение.**

Поскольку имеются отдельные значения признака, данные не сгруппированы, применим формулу средней арифметической простой.

$$\bar{x} = \frac{\sum x}{n} = \frac{4,4 + 4,3 + 4,4 + 4,5 + 4,3 + 4,3 + 4,6 + 4,2 + 4,6 + 4,1}{10} = 4,4$$

**Пример 2.** Определить среднее количество филиалов банка

| Количество филиалов в городе организации, $x$ | Число банков $f$ | $xf$ | Частость, $w$ | $xw$ |
|-----------------------------------------------|------------------|------|---------------|------|
| 2                                             | 1                | 2    | 0,05          | 0,1  |
| 3                                             | 5                | 15   | 0,25          | 0,75 |
| 4                                             | 8                | 32   | 0,4           | 1,6  |
| 5                                             | 4                | 20   | 0,2           | 1    |
| 6                                             | 2                | 12   | 0,1           | 0,6  |
| Итого                                         | 20               | 81   | 1             | 4,05 |

**Решение.** Данные представлены в виде дискретного ряда распределения, одни и те же значения группировочного признака повторяются несколько раз. Поэтому применим формулу средней арифметической взвешенной. Для расчета заполним столбец  $xf$ , и рассчитаем итог по столбцу.

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{81}{20} = 4,05$$

Используя свойства средней арифметической, для расчета вместо частот можно использовать значения частостей.

$$\bar{x} = \frac{\sum xw}{\sum w} = \frac{4,05}{1} = 4,05$$

**Пример 3.** Рассчитать средний размер прибыли банка.

| № группы | Размер прибыли, $x$ | Число банков (частота) $f$ | $x'$ | $x'f$ |
|----------|---------------------|----------------------------|------|-------|
| 1        | 3,7 - 4,6           | 3                          | 4,15 | 12,45 |
| 2        | 4,6 - 5,5           | 3                          | 5,05 | 15,15 |
| 3        | 5,5 - 6,4           | 7                          | 5,95 | 41,65 |
| 4        | 6,4 - 7,3           | 4                          | 6,85 | 27,4  |
| 5        | 7,3 - 8,2           | 3                          | 7,75 | 23,25 |
| Итого    |                     | 20                         |      | 119,9 |

**Решение.** Варианты осредняемого признака (размера прибыли) представлены не одним числом, а виде интервала «от - до». Для расчета по формуле средней арифметической взвешенной исчисляются середины интервалов  $x'$ . Дальнейший расчет производится обычным методом определения средней арифметической взвешенной.

$$\bar{x} = \frac{\sum x'f}{\sum f} = \frac{119,9}{20} = 5,995 \text{ млн. руб.}$$

При расчете можно, так же, как в предыдущем случае, воспользоваться значениями частостей.

**Пример 4.** По трем обменным пунктам известен курс доллара и выручка от продажи валюты. Рассчитать средний курс доллара по этим обменным пунктам.

| Номер обменного пункта | Валютный курс<br>$x$ | Выручка от продажи валюты<br>$B$ |
|------------------------|----------------------|----------------------------------|
| 1                      | 28,70                | 232,47                           |
| 2                      | 28,68                | 298,27                           |
| 3                      | 28,73                | 149,40                           |
| Итого                  |                      | 680,14                           |

**Решение.**

Статистическая информация не содержит частот по отдельным вариантам, а представлена как их произведение, поскольку выручка от продажи валюты – это произведение валютного курса ( $x$ ) на объем продаж. Поэтому применим формулу средней гармонической взвешенной.

$$\bar{x} = \frac{\text{Общ.выручка}}{\text{Объемпродаж}}$$

$$x_{\text{гарм.}} = \frac{\sum B_i}{\sum \frac{B_i}{x_i}} = \frac{232.47 + 298.27 + 149.40}{\frac{232.47}{28.70} + \frac{298.27}{28.68} + \frac{149.40}{28.73}} = 28.70 \text{ руб.}$$

**Пример 5.** Двое рабочих в течение рабочего дня заняты изготовлением одинаковых деталей. Один рабочий тратит на изготовление детали 3 минуты, другой – 6 мин. Определить средние затраты времени на изготовление детали.

**Решение.**

На первый взгляд, следует применить формулу средней арифметической простой, но в течение рабочего дня ими было изготовлено разное число деталей.

Средние затраты времени на 1 деталь должны определяться по формуле

$$\bar{x} = \frac{\text{все}_\text{затраты}_\text{времени}}{\text{все}_\text{изготовленные}_\text{детали}}$$

Затраты времени представляют собой произведение количества изготовленных деталей ( $f$ ) и времени на изготовление одной детали ( $x$ ). Поскольку затраты рабочего времени ( $xf$ ) у обоих рабочих равны (рабочий день), то применим формулу средней гармонической простой.

$$\text{Итак, } \bar{x} = \frac{\text{все}_\text{затраты}_\text{времени}}{\text{все}_\text{изготовленные}_\text{детали}}$$

$$x_{\text{гарм}} = \frac{n}{\sum \frac{1}{x}} = \frac{1+1}{\frac{1}{3} + \frac{1}{6}} = 4 \text{ мин.}$$

**Пример 6.** По имеющимся данным о ценах товара в различных фирмах города определить моду и медиану.

а) 4,4 4,3 4,4 4,5 4,3 4,3 4,6 4,2 4,6

б) 4,4 4,3 4,4 4,5 4,3 4,3 4,6 4,2 4,6 4,1

**Решение.** В обоих случаях данные не сгруппированы.

а) в данной совокупности чаще всего повторяется значение 4,3, поэтому  $Mo=4,3$

Для определения медианы надо провести ранжирование:

4,2 4,3 4,3 4,3 4,4 4,4 4,5 4,6 4,6

В данном ряду нечетное число членов, варианта, расположенная посередине, является медианой.  $Me=4,4$

б) в данной совокупности чаще всего повторяется значение 4,3, поэтому  $Mo=4,3$

Для определения медианы проведем ранжирование:

4,1 4,2 4,3 4,3 4,3 4,4 4,4 4,5 4,6 4,6

В данном ряду четное число членов (10), поэтому медиана рассчитывается как средняя арифметическая из двух вариантов, расположенных в центре ряда, т.е.  $Me=(4,3+4,4)/2=4,35$

**Пример 7.** По имеющимся данным определить моду и медиану

| Количество филиалов в городе организации, $x$ | Число банков<br>$f$ | Накопленные частоты $S$ |
|-----------------------------------------------|---------------------|-------------------------|
| 2                                             | 1                   | 1                       |
| 3                                             | 5                   | 6                       |
| 4                                             | 8                   | 14                      |
| 5                                             | 4                   |                         |
| 6                                             | 2                   |                         |
| Итого                                         | 20                  |                         |

**Решение.** Данные представлены в виде дискретного ряда распределения.

Наибольшая частота  $f=8$  соответствует варианту  $x=4$ , поэтому  $Mo = 4$ .

Для нахождения медианы следует рассчитать накопленные частоты.  $S=14$ , впервые превысившая 10 (половину общей суммы частот), соответствует варианту  $x=4$ . Значит,  $Me=4$ .

**Пример 8.** По имеющимся данным определить моду и медиану

| № группы | Размер прибыли, $x$ | Число банков<br>(частота) $f$ | Накопленные частоты $S$ |
|----------|---------------------|-------------------------------|-------------------------|
| 1        | 3,7 - 4,6           | 3                             | 3                       |
| 2        | 4,6 - 5,5           | 3                             | 6                       |
| 3        | 5,5 - 6,4           | 7                             | 13                      |
| 4        | 6,4 - 7,3           | 4                             |                         |
| 5        | 7,3 - 8,2           | 3                             |                         |
| Итого    |                     | 20                            |                         |

**Решение.** Данные представлены в виде интервального ряда распределения.

Для расчета моды требуется сначала определить модальный интервал: наибольшая частота  $f=7$  соответствует интервалу 5,5 - 6,4. Значит, это модальный интервал. Конкретное значение моды определяется по формуле:

$$Mo = x_{Mo} + i_{Mo} \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})} = 5,5 + 0,9 * \frac{7-3}{(7-3) + (7-4)} = 6,01$$

Для расчета медианы определим медианный интервал. Для этого рассчитаем накопленные частоты, пока они не превысят половину суммы частот (т.е. 10).  $S=13$  соответствует интервалу 5,5-6,4, значит, это медианный интервал. Конкретное значение медианы найдем по формуле:

$$Me = x_{Me} + i_{Me} \frac{\sum f - S_{Me-1}}{f_{Me}} = 5,5 + 0,9 * \frac{10-6}{7} = 6,01$$

**Задачи для самостоятельного решения.**

**Задача 1.** По имеющимся данным найти среднюю выработку рабочего, структурные средние

| № рабочего                 | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9   | 10 | 11 |
|----------------------------|----|----|----|----|----|----|----|----|-----|----|----|
| Дневная выработка рабочего | 70 | 73 | 68 | 75 | 75 | 74 | 83 | 81 | 100 | 73 | 80 |

**Задача 2.** По имеющимся данным найти среднее количество секций в магазине, структурные средние.

| Количество товарных секций в магазине | Количество магазинов |
|---------------------------------------|----------------------|
| 2                                     | 3                    |

|   |   |
|---|---|
| 3 | 4 |
| 4 | 6 |
| 5 | 4 |
| 6 | 3 |

**Задача 3.** По имеющимся данным найти средний размер прибыли банка, структурные средние

| Размер<br>прибыли | Число<br>банков |
|-------------------|-----------------|
| 4,7 - 5,6         | 3               |
| 5,6 - 6,5         | 3               |
| 6,5 - 7,4         | 7               |
| 7,4 - 8,3         | 4               |
| 8,3 - 9,2         | 3               |

**Задача 4.** Бригада операторов компьютерного набора из трех человек должна выполнить набор книги в 500 страниц. Первый оператор тратит на набор одну страницу 15 мин., другой – 10 мин., третий – 20 мин. Определить, сколько времени им потребуется.

**Задача 5.** Три предприятия производят одноименную продукцию. По данным о себестоимости единицы изделия и общих издержках производства определить среднюю себестоимость единицы изделия.

| Предприятие | Себестоимость<br>единицы изделия | Общие издержки<br>производства |
|-------------|----------------------------------|--------------------------------|
| А           | 28                               | 23200                          |
| Б           | 27,5                             | 29800                          |
| В           | 28,3                             | 14900                          |

## Тема 4

### ПОКАЗАТЕЛИ ВАРИАЦИИ

Различие индивидуальных значений признака внутри изучаемой совокупности называется *вариацией признака*. Средняя величина – это абстрактная обобщающая характеристика признака изучаемой совокупности, но она не показывает строения совокупности. Для характеристики совокупностей и исчисленных средних величин важно знать, какая вариация признака скрывается за средними. В некоторых случаях отдельные значения признака близко примыкают к средней арифметической и мало от нее отличаются, в таких случаях средняя хорошо представляет всю совокупность. В других случаях, наоборот, отдельные значения далеко отстоят от средней, и средняя плохо представляет совокупность. Колеблемость отдельных значений, степень их близости к средней характеризуют показатели вариации.

#### 1. Абсолютные и средние показатели вариации.

Наиболее простой показатель вариации – *размах вариации*, определяемый как разность между наибольшим ( $x_{max}$ ) и наименьшим ( $x_{min}$ ) значениями вариант

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

Этот показатель прост в вычислении и указывает на общие размеры вариации, но он не дает представления о степени колеблемости внутри совокупности, т.к. улавливает только крайние отклонения.

Различие всех единиц изучаемой совокупности учитывает *среднее линейное отклонение*. Среднее линейное отклонение есть средняя арифметическая из отклонений индивидуальных значений от средней (без учета знака этих отклонений):

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n}$$

или 
$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i}$$

На практике меру вариации более объективно отражает показатель дисперсии. *Дисперсия* – это средняя арифметическая квадратов отклонений каждого значения признака от средней арифметической. Другими словами, это средний квадрат отклонений. Дисперсия вычисляется по формуле:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

или 
$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}$$

Корень квадратный из дисперсии представляет собой *среднее квадратическое отклонение*. Достоинством этого показателя является то, что он выражается в тех же единицах измерения, что и признак.

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

или 
$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}}$$

Дисперсия и среднее квадратическое отклонение являются основными обобщающими показателями вариации. Среднее квадратическое отклонение является мерилем надежности средней. Чем меньше этот показатель, тем лучше средняя арифметическая отражает всю представляемую совокупность.

## 2. Относительные показатели вариации

Относительные показатели вариации позволяют сравнивать характер рассеивания в различных совокупностях, например, при сравнении разноименных совокупностей, при различных значениях средней. Расчет относительных показателей вариации осуществляют как отношение абсолютного показателя вариации к средней арифметической. Как правило, они рассчитываются в процентах.

*Коэффициент осцилляции* отражает относительную колеблемость крайних значений вокруг средней

$$V_R = \frac{R}{\bar{x}} \times 100\%.$$

*Линейный коэффициент вариации* характеризует долю усредненного значения абсолютных отклонений от средней величины

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

*Коэффициент вариации* – наиболее распространенный показатель колеблемости, используемый для оценки типичности средней.

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \times 100\%$$

Чем больше разброс значений признака вокруг средней, тем больше коэффициент вариации и тем менее представительна средняя. Как правило, считают, что если  $V_{\sigma} > 33\%$ , то это говорит о большой колеблемости признака в совокупности, и совокупность неоднородна.

## 3. Правило сложения дисперсий

Определить влияние отдельных факторов, характеризующих колеблемость индивидуальных значений признака можно при помощи группировок, подразделив изучаемую совокупность на группы, однородные по изучаемому признаку. При этом можно исчислить следующие виды дисперсий: общую дисперсию, внутригрупповые дисперсии, среднюю из внутригрупповых дисперсий и межгрупповую дисперсию.

*Внутригрупповые дисперсии* ( $\sigma_1, \sigma_2, \dots$ ) отражают случайную вариацию, т.е. часть вариации, происходящую влиянием неучтенных факторов и не зависящую от признака-фактора, положенного в основание группировки.

*Средняя из внутригрупповых дисперсий* ( $\overline{\sigma^2}$ ) – это средняя арифметическая взвешенная из внутригрупповых дисперсий.

*Межгрупповая дисперсия* ( $\delta^2$ ) – это средний квадрат отклонений групповых средних от общей средней. Характеризует систематическую вариацию, т.е. различия в величине изучаемого (результативного) признака за счет признака-фактора, положенного в основание группировки.

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

*Общая дисперсия* ( $\sigma^2$ ) характеризует вариацию признака, которая зависит от всех условий в данной совокупности.

Между указанными видами дисперсий существует соотношение: общая дисперсия равна сумме величин средней из внутригрупповых дисперсий и межгрупповой дисперсии. Формула правила сложения дисперсий:

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

Правило сложения дисперсий позволяет выявить зависимость результативного признака от определяющих факторов путем соотношения межгрупповой и общей дисперсии:

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

Здесь  $\eta^2$  - коэффициент детерминации, который показывает долю вариации результативного признака, объясненную влиянием вариации факторного признака.

#### 4. Дисперсия альтернативного признака

Альтернативные признаки – это признаки, которыми обладают одни единицы совокупности и не обладают другие (например, работники либо имеют высшее образование, либо не имеют, т.е. это два взаимоисключающих варианта). При статистическом выражении колеблемости альтернативного признака, наличие признака обозначается 1, а доля единиц совокупности, обладающих данным признаком, обозначается  $p$ . Отсутствие признака обозначается 0, доля единиц, не обладающих данным признаком, -  $q$ . Очевидно,  $p+q=1$ .

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{1 * p + 0 * q}{p + q} = \frac{p}{1} = p$$

Отсюда,

$$\sigma_p^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} = \frac{(1-p)^2 * p + (0-p)^2 * q}{p + q} = \frac{q^2 p + p^2 q}{1} = pq(p + q) = pq$$

т.е.  $\sigma_p^2 = pq$

Т.о., дисперсия альтернативного признака равна произведению доли единиц, обладающих признаком, и доли единиц, не обладающих им.

#### 5. Характеристика закономерности рядов распределения

В вариационных рядах существует определенная связь в изменении частот и значений варьирующего признака: с увеличением варьирующего признака величина частот вначале возрастает до определенной величины, а затем уменьшается. Такого рода изменения называются закономерностями распределения.

Положение кривой на оси абсцисс и ее рассеивание являются двумя наиболее существенными свойствами кривой. Другими словами, фактическая форма кривой для любого распределения зависит от значений  $\bar{x}$  и  $\sigma$ . Наряду с ними существует ряд других важных свойств кривой распределения: степень асимметрии, высоко- или низковооршинность, которые в совокупности характеризуют форму, или тип, кривой распределения. Выяснение общего характера распределения предполагает оценку степени его однородности, а также вычисления показателей асимметрии и эксцесса.

Распределение является *симметричным*, если частоты двух любых вариантов, равноотстоящих в обе стороны от центра распределения, равны между собой. Для симметричного распределения средняя арифметическая, мода и медиана равны между собой:

$$\bar{x} = Me = Mo.$$

Чем больше разница между средней арифметической и модой (медианой), тем больше асимметрия ряда.

Коэффициент асимметрии исчисляется по формуле

$$A_s = \frac{\bar{x} - Mo}{\sigma}$$

Коэффициент асимметрии изменяется от  $-3$  до  $+3$ . Если  $A_s > 0$ , то кривая распределения имеет длинный правый «хвост», т.е. налицо правосторонняя асимметрия. При этом выполняется соотношение  $Mo < Me < \bar{x}$ .

Если  $A_s < 0$ , то асимметрия левосторонняя, кривая распределения имеет длинный левый «хвост». При этом  $\bar{x} > Me > Mo$ .

На практике асимметрия считается значительной, если коэффициент асимметрии превышает по модулю 0,25.

Экссесс представляет собой вершины распределения вверх или вниз от вершины нормального распределения. Коэффициент эксцесса рассчитывается по формуле

$$E_x = \frac{\bar{m}_4}{\sigma^4} - 3,$$

где  $\bar{m}_4$  - центральный момент четвертого порядка,  $\bar{m}_4 = \frac{\sum(x - \bar{x})^4}{n}$ , или  $\bar{m}_4 = \frac{\sum(x - \bar{x})^4 f}{\sum f}$ .

При нормальном распределении  $\frac{\bar{m}_4}{\sigma^4} = 3$ , эксцесс нормального распределения равен 0.

Обычно, если эксцесс положителен, то распределение островершинное, если отрицательный – то плосковершинное.

## 6. Примеры решения задач

**Пример 1.** По имеющимся данным о ценах товара в различных фирмах города рассчитать абсолютные и относительные показатели вариации:

4,4 4,3 4,4 4,5 4,3 4,3 4,6 4,2 4,6 4,1

**Решение.**

Абсолютные показатели вариации.

$$R = x_{\max} - x_{\min} = 4,8 - 4,1 = 0,7$$

Для расчета остальных показателей вариации заполним в таблице дополнительные расчетные графы, зная, что  $\bar{x} = 4,4$

| Цены товара в разных фирмах, $x$ | $ x_i - \bar{x} $ | $(x_i - \bar{x})^2$ |
|----------------------------------|-------------------|---------------------|
| 4,1                              | 0,3               | 0,09                |
| 4,2                              | 0,2               | 0,04                |
| 4,3                              | 0,1               | 0,01                |
| 4,3                              | 0,1               | 0,01                |
| 4,3                              | 0,1               | 0,01                |
| 4,4                              | 0                 | 0                   |
| 4,4                              | 0                 | 0                   |
| 4,5                              | 0,1               | 0,01                |
| 4,6                              | 0,2               | 0,04                |
| 4,8                              | 0,4               | 0,16                |
| <i>Итого</i>                     | 1,4               | 0,37                |

Поскольку имеются отдельные значения признака, данные не сгруппированы, применим невзвешенные формулы показателей вариации:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{1,4}{10} = 0,14$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{0,37}{10} = 0,037$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{0,037} = 0,192$$

Относительные показатели вариации:

$$V_R = \frac{R}{\bar{x}} \times 100\% = \frac{0,7}{4,4} \times 100\% = 15,9\%$$

$$V_{\bar{d}} = \frac{\bar{d}}{\bar{x}} \times 100\% = \frac{0,14}{4,4} \times 100\% = 3,18\%$$

$$V_{\sigma} = \frac{\sigma}{\bar{x}} \times 100\% = \frac{0,184}{4,4} \times 100\% = 4,18\%$$

Колеблемость признака в совокупности небольшая, совокупность можно считать однородной по данному признаку.

**Пример 2.** По имеющимся данным рассчитать абсолютные и относительные показатели вариации:

| Количество филиалов в городе организации, $x$ | Число банков $f$ | $ x_i - \bar{x} $ | $ x_i - \bar{x}  f$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 f$ |
|-----------------------------------------------|------------------|-------------------|---------------------|---------------------|-----------------------|
| 2                                             | 1                | 2                 | 2                   | 4                   | 4                     |
| 3                                             | 5                | 1                 | 5                   | 1                   | 5                     |
| 4                                             | 8                | 0                 | 0                   | 0                   | 0                     |
| 5                                             | 4                | 1                 | 4                   | 1                   | 4                     |
| 6                                             | 2                | 2                 | 4                   | 4                   | 8                     |
| <i>Итого</i>                                  | 20               |                   | 15                  |                     | 21                    |

**Решение.**

$$R = x_{\max} - x_{\min} = 6 - 2 = 4$$

Для расчета остальных показателей вариации заполним в таблице дополнительные расчетные графы.

Поскольку данные представлены в виде дискретного ряда распределения, применим взвешенные формулы показателей вариации.

Для удобства расчетов округлим значение  $\bar{x} = 4,05$  до  $\bar{x} = 4$

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i} = \frac{15}{20} = 0,75$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{21}{20} = 1,05$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{1,05} = 1,025$$

Относительные показатели вариации:

$$V_R = \frac{R}{x} \times 100\% = \frac{4}{4} \times 100\% = 100\%$$

$$V_{\bar{d}} = \frac{\bar{d}}{x} \times 100\% = \frac{0,75}{4} \times 100\% = 18,7\%$$

$$V_{\sigma} = \frac{\sigma}{x} \times 100\% = \frac{1,025}{4} \times 100\% = 25,6\%$$

Колеблемость признака в совокупности достаточно высокая, но  $V_{\sigma} < 33\%$ , поэтому совокупность можно считать однородной по данному признаку.

**Пример 3.**

Имеются следующие данные о выработке рабочих и их квалификации.

| Выработка | Рабочие 3 разряда | Рабочие 4 разряда |
|-----------|-------------------|-------------------|
| 101       | 5                 |                   |
| 102       | 4                 |                   |
| 103       | 3                 | 1                 |
| 104       | 1                 | 2                 |
| 105       |                   | 4                 |
| 106       |                   | 3                 |

Определить, влияет ли фактор квалификации рабочего на его выработку, рассчитать коэффициент детерминации.

**Решение.**

Для расчета коэффициента детерминации воспользуемся правилом сложения дисперсий. Дополним таблицу дополнительными расчетными графами.

| Выработка,<br>$x$ | Рабочие<br>3<br>разряда,<br>$f$ | $xf$ | $ x_i - \bar{x} $ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 f$ | Рабочие<br>4<br>разряда,<br>$f$ | $xf$ | $ x_i - \bar{x} $ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 f$ |
|-------------------|---------------------------------|------|-------------------|---------------------|-----------------------|---------------------------------|------|-------------------|---------------------|-----------------------|
| 101               | 5                               | 505  | 1                 | 1                   | 5                     |                                 |      |                   |                     |                       |
| 102               | 4                               | 408  | 0                 | 0                   | 0                     |                                 |      |                   |                     |                       |
| 103               | 3                               | 309  | 1                 | 1                   | 3                     | 1                               | 103  | 2                 | 4                   | 4                     |
| 104               | 1                               | 104  | 2                 | 4                   | 4                     | 2                               | 208  | 1                 | 1                   | 2                     |
| 105               |                                 |      |                   |                     |                       | 4                               | 420  | 0                 | 0                   | 0                     |
| 106               |                                 |      |                   |                     |                       | 3                               | 318  | 1                 | 1                   | 3                     |
| Итого             | 13                              | 1326 |                   |                     | 12                    | 10                              | 1049 |                   |                     | 9                     |

1) Для расчета внутригрупповых дисперсий рассчитаем сначала внутригрупповые средние (по формуле средней взвешенной)

$$\bar{x}_1 = \frac{\sum xf}{\sum f} = \frac{1326}{13} = 102$$

$$\bar{x}_2 = \frac{\sum xf}{\sum f} = \frac{1049}{10} = 104,9 \approx 105$$

Внутригрупповые дисперсии:

$$\sigma_1^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{12}{13} = 0,92$$

$$\sigma_2^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{9}{10} = 0,9$$

2) Средняя из внутригрупповых дисперсий рассчитывается как средняя арифметическая взвешенная из внутригрупповых дисперсий, где весами выступает численность групп:

$$\bar{\sigma}^2 = \frac{\sum \sigma_i^2 f}{\sum f} = \frac{0,92 * 13 + 0,9 * 10}{13 + 10} = 0,91$$

3) Для расчета межгрупповой дисперсии сначала определим общую среднюю как среднюю арифметическую взвешенную из групповых средних:

$$\bar{x} = \frac{\sum \bar{x}_i f}{\sum f} = \frac{102 * 13 + 105 * 10}{13 + 10} = 103,3$$

Среднюю можно также вычислить обычным способом.

$$\bar{x} = \frac{\sum (x_i - \bar{x})^2 f}{\sum f} = \frac{(102 - 103,3)^2 * 13 + (105 - 103,3)^2 * 10}{13 + 10} = 2,2$$

Как видим, межгрупповая дисперсия, характеризующая различия в величине результативного признака (выработки) за счет факторного признака (квалификации), значительно превышает внутригрупповые дисперсии, которые отражают случайную вариацию под влиянием неучтенных факторов.

4) Общую дисперсию найдем по правилу сложения дисперсий

$$\sigma^2 = \bar{\sigma}^2 + \delta^2 = 0,91 + 2,2 = 3,11$$

Общую дисперсию можно также вычислить обычным способом.

5) Долю вариации результативного признака (выработки) под влиянием факторного (квалификации) показывает коэффициент детерминации:

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{2,2}{3,11} = 0,707$$

Таким образом, различия в величине выработке рабочих на 70,7% объясняются различиями в их квалификации, а на 29,3% - влиянием прочих факторов.

**Пример 4.** По имеющимся данным о ценах товара в различных фирмах города рассчитать показатель асимметрии распределения:

4,4 4,3 4,4 4,5 4,3 4,3 4,6 4,2 4,6 4,1

**Решение.**

$$A_s = \frac{\bar{x} - Mo}{\sigma}$$

Зная, что

$$\bar{x} = 4,4$$

$$Mo = 4,3$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{0,037} = 0,192,$$

вычислим

$$A_s = \frac{4,4 - 4,3}{0,192} = 0,52$$

Значение показателя асимметрии говорит о наличии значительной правосторонней асимметрии.

## ВЫБОРОЧНЫЙ МЕТОД В ЭКОНОМИКО-СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ

Статистическое исследование может осуществляться по данным несплошного наблюдения, основная цель которого состоит в получении характеристик изучаемой совокупности по обследованной ее части. Одним из наиболее распространенных в статистике методов, применяющих несплошное наблюдение, является выборочный метод.

### 1. Понятие о выборочном исследовании

Под *выборочным* понимается метод статистического исследования, при котором обобщающие показатели изучаемой совокупности устанавливаются по некоторой ее части на основе положений случайного отбора. При выборочном методе обследованию подвергается сравнительно небольшая часть всей изучаемой совокупности (обычно до 5-10%, реже до 15-25%). При этом подлежащая изучению статистическая совокупность, из которой производится отбор части единиц, называется *генеральной совокупностью*. Отобранная из генеральной совокупности некоторая часть единиц, подвергающаяся обследованию, называется *выборочной совокупностью* (или *выборкой*).

Значение выборочного метода состоит в том, что при меньшей численности обследуемых единиц проведение исследования осуществляется с меньшими затратами и в более короткие сроки, повышая оперативность статистической информации.

Поскольку изучаемая статистическая совокупность состоит из единиц с варьирующими признаками, то состав выборочной совокупности может в той или иной степени отличаться от состава генеральной совокупности. Это объективно возникающее расхождение между характеристиками выборки и генеральной совокупности составляет *ошибку выборки*. Она зависит от ряда факторов: степени вариации изучаемого признака, численности выборки, методов отбора единиц в выборочную совокупность, принятого уровня достоверности результата исследования.

Способы определения ошибки выборки при различных приемах формирования выборочной совокупности и распространение характеристик выборки на генеральную совокупность составляют основное содержание статистической методологии выборочного метода.

### 2. Характеристики выборочной совокупности и их распространение на генеральную совокупность.

При использовании выборочного метода в социально-экономических исследованиях обычно применяют два основных вида обобщающих показателей: относительную величину альтернативного признака и среднюю величину количественного признака.

Относительная величина альтернативного признака характеризует долю (удельный вес) единиц в статистической совокупности, которые отличаются от всех других единиц этой совокупности только наличием (отсутствием) изучаемого признака. Например, доля нестандартных изделий во всей партии товара, удельный вес продавцов в общей численности работников магазина и т.п.

Средняя величина количественного признака – это обобщающая характеристика варьирующего признака, который имеет различные значения у отдельных единиц статистической совокупности. Например, средний вес изделия, средняя выработка одного продавца и т.д.

В генеральной совокупности доля единиц, обладающих изучаемым признаком, называется *генеральной долей* (обозначается  $P$ ), а средняя величина варьирующего признака – *генеральной средней* (обозначается  $\bar{x}$ ).

В выборочной совокупности долю изучаемого признака называют *выборочной долей*  $w$ , а среднюю величину в выборке – *выборочной средней*  $\tilde{x}$ .

Выборочная доля определяется из отношения единиц, обладающих изучаемым признаком,  $m$  к общей численности единиц выборочной совокупности  $n$ :

$$w = \frac{m}{n}$$

Основная задача выборочного исследования – на основе характеристик выборочной совокупности  $w$  и  $\tilde{x}$  получить достоверные суждения о показателях доли  $P$  и средней  $\bar{x}$  в генеральной совокупности.

Возможные расхождения между характеристиками выборочной и генеральной совокупностей измеряются средней ошибкой выборки  $\mu$ . В математической статистике доказывается, что значения  $\mu$  определяются по формуле

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

где  $\sigma_{ген}^2$  – генеральная дисперсия. Но при проведении выборочных обследований она, как правило, неизвестна. На практике для определения  $\mu$  обычно используется дисперсия выборочной совокупности  $\sigma^2$ .

При этом для показателя доли альтернативного признака дисперсия определяется по формуле дисперсии альтернативного признака, т.е.

$$\sigma_w^2 = w(1-w)$$

Следует иметь в виду, что приведенная выше формула расчета средней ошибки выборки  $\mu$  применяется лишь при *повторном* отборе, когда каждая попавшая в выборку единица после фиксации значения изучаемого признака должна быть возвращена в генеральную совокупность, где ей опять представляется возможность попасть в выборку. Но на практике выборочные обследования проводятся обычно по схеме *бесповторного* отбора, при котором повторное попадание в выборку одних и тех же единиц исключено.

Поскольку при бесповторном отборе численность генеральной совокупности  $N$  в ходе выборки сокращается, то в формулу расчета  $\mu$  включают дополнительный множитель  $1 - \frac{n}{N}$ .

Формула средней ошибки выборки принимает следующий вид:

$$\begin{aligned} \mu &= \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} && \text{- общий вид;} \\ \mu_w &= \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} && \text{- для выборочной доли} \\ \mu_x &= \sqrt{\frac{\sigma_x^2}{n} \left(1 - \frac{n}{N}\right)} && \text{- для выборочной средней} \end{aligned}$$

Значения средней ошибки выборки для выборочной доли и выборочной средней необходимы для установления возможных значений генеральной доли  $P$  и генеральной средней  $\bar{x}$ . Пределы значений этих показателей определяются по формулам:

$$\begin{aligned} P &= w \pm \mu_w \\ \bar{x} &= \bar{x} \pm \mu_x \end{aligned}$$

В математической статистике доказывается, что пределы значений характеристик генеральной совокупности  $P$  и  $\bar{x}$  отличаются от характеристик выборочной совокупности  $w$  и  $\tilde{x}$  на величину  $\pm \mu$  с вероятностью 0,683. Т.е. в 683 случаях из тысячи генеральные характеристики будут находиться в установленных пределах, в остальных 317 случаях они могут выйти за эти пределы.

Вероятность суждения можно повысить, если расширить пределы отклонений, увеличив среднюю ошибку выборки в  $t$  раз. Таким образом, показатели генеральной совокупности по показателям выборки определяются по формулам:

$$P = w \pm t \cdot \mu_w$$

$$\bar{x} = \tilde{x} \pm t \cdot \mu_x$$

Величина  $t \cdot \mu$  называется *предельной ошибкой выборки*  $\Delta$ . Т.е.

$$\Delta_w = t \cdot \mu_w$$

$$\Delta_x = t \cdot \mu_x$$

Множитель  $t$  называется *коэффициентом доверия* и определяется в зависимости от того, с какой вероятностью надо гарантировать результаты выборочного обследования. Конкретные значения коэффициента доверия  $t$  для различных степеней вероятности определяются с помощью функции А.М.Ляпунова. На практике пользуются готовыми таблицами этой функции:

| $t$ | Вероятность | $t$ | Вероятность |
|-----|-------------|-----|-------------|
| 0,0 | 0,0000      | 2,0 | 0,9545      |
| 1,0 | 0,6827      | 2,5 | 0,9876      |
| 1,5 | 0,8664      | 3,0 | 0,9973      |

### 3. Оптимальная численность выборки

При организации выборочного наблюдения прежде всего следует иметь в виду, что размер ошибки выборки прежде всего зависит от численности выборки  $n$ . Уменьшение средней ошибки выборки всегда связано с увеличением объема выборки, но не в прямой пропорции. Из формулы расчета средней ошибки выборки  $\mu$  следует, что  $\mu$  обратно пропорционально  $\sqrt{n}$ , т.е. при увеличении выборки в 4 раза ее ошибки уменьшаются лишь вдвое.

Рассмотрим формулу предельной ошибки выборки для случая повторной выборки:

$$\Delta_x = t \cdot \mu_x = t \sqrt{\frac{\sigma^2}{n}} = \frac{t\sigma}{\sqrt{n}}$$

Отсюда:

$$n = \frac{t^2 \sigma^2}{\Delta_x^2}$$

Численность выборки для бесповторного отбора определяется аналогично:

$$n = \frac{N \cdot t^2 \sigma^2}{N \cdot \Delta_x^2 + t^2 \sigma^2}$$

Используемая в формулах величина  $\Delta_x$  - это абсолютная величина предельной ошибки выборки. На практике нередко задается величина не абсолютной предельной ошибки, а величина относительной погрешности выраженная в процентах к средней:

$$\Delta_{\text{относ}} = \frac{\Delta_x}{\bar{x}} \cdot 100\%,$$

откуда

$$\Delta_x = \frac{\Delta_{\text{относ}} \cdot \bar{x}}{100\%}$$

Для оценки неизвестной величины  $\sigma^2$  (дисперсии в генеральной совокупности) используются следующие способы:

- пробное обследование небольшого объема
- использование данных прошлых выборочных обследований, проводившихся в аналогичных целях
- если распределение признака в генеральной совокупности можно отнести к нормальному закону распределения, то  $\sigma \approx R/6$ , где  $R$  – размах вариации.

#### 4. Примеры решения задач

**Пример 1.** Проведено выборочное обследование партии заготовок деталей. При механическом бесповторном отборе 2,5 % изделий получены следующие данные о распределении образцов по весу.

| Исходные данные |               | Расчетные показатели |               |                   |                       |                         |
|-----------------|---------------|----------------------|---------------|-------------------|-----------------------|-------------------------|
| Вес изделия, г. | Число изделий | Середина интервала   | $xf$          | $x_i - \tilde{x}$ | $(x_i - \tilde{x})^2$ | $(x_i - \tilde{x})^2 f$ |
| до 1000         | 22            | 987,5                | 21725         | -52,5             | 2756,25               | 60637,5                 |
| 1000-1025       | 77            | 1012,5               | 77962,5       | -27,5             | 756,25                | 58231,25                |
| 1025-1050       | 183           | 1037,5               | 189862,5      | -2,5              | 6,25                  | 1143,75                 |
| 1050-1075       | 85            | 1062,5               | 90312,5       | 22,5              | 506,25                | 43031,25                |
| 1075-1100       | 23            | 1087,5               | 25012,5       | 47,5              | 2256,25               | 51893,75                |
| свыше 1100      | 10            | 1112,5               | 11125         | 72,5              | 5256,25               | 52562,5                 |
| <i>Итого</i>    | <i>400</i>    |                      | <i>416000</i> |                   |                       | <i>267500</i>           |

При условии, что к нестандартной продукции относятся заготовки весом до 1000 г. и свыше 1100 г. определить пределы значения удельного веса стандартной продукции и среднего веса изделия для всей партии с вероятностью 0,954.

**Решение.**

По условию  $n = 400$ . Найдем  $N = 400 \cdot 100\% / 2,5\% = 16000$  шт.

Установим обобщающие показатели выборочной совокупности.

Расчет выборочной доли  $w$ .

Число стандартных единиц в выборке  $m = 400 - (22+10) = 368$ , общее число единиц в выборке  $n = 400$ .

$$w = \frac{m}{n} = \frac{368}{400} = 0,92, \text{ т.е. удельный вес стандартных изделий в выборке } 92\%$$

Расчет выборочной средней  $\tilde{x}$ . Вычислим  $\tilde{x}$  по формуле средней взвешенной  $\tilde{x} = \frac{\sum xf}{\sum f}$ .

Для этого определим середины интервалов. Середины крайних (открытых) интервалов определим, исходя из гипотезы равнонаполненности интервалов, т.е. принимаем границы первого интервала от 975 до 1000 г., последнего – от 1100 до 1125 г.

$$\text{Средний вес изделия в выборке составляет } \tilde{x} = \frac{\sum xf}{\sum f} = \frac{41600}{400} = 1040 \text{ г.}$$

Установим средние ошибки выборки для обобщающих характеристик выборочной совокупности, пользуясь формулами для бесповторного отбора:

Для выборочной доли.

$$\mu_w = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,92(1-0,92)}{400} \left(1 - \frac{400}{16000}\right)} = 0,0133, \text{ т.е. средняя ошибка выборки}$$

для доли стандартной продукции составляет 1,33%

Для выборочной средней.

$$\text{Сначала требуется вычислить } \sigma^2 = \frac{\sum (x_i - \tilde{x})^2 f}{\sum f} = \frac{267500}{400} = 668,75$$

$$\mu_x = \sqrt{\frac{\sigma_x^2}{n} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{668,75}{400} \left(1 - \frac{400}{16000}\right)} = \sqrt{1,63} \approx 1,27 \text{ г., т.е. средняя ошибка выборки для}$$

средней величины составляет 1,27 г.

Установим предельные значения для характеристик генеральной совокупности, учитывая, что вероятности 0,954 соответствует значение коэффициента доверия  $t=2$ :

Для генеральной доли

$$P = w \pm t \cdot \mu_w = 92 \pm 2 \cdot 1,33 (\%), \quad \text{или} \quad 89,34\% \leq P \leq 94,66\%$$

Для генеральной средней

$$\bar{x} = \tilde{x} \pm t \cdot \mu_x = 1040 \pm 2 \cdot 1,27 (\text{г}), \quad \text{или} \quad 1037,46 \text{ г.} \leq \tilde{x} \leq 1042,52 \text{ г.}$$

Итак, с вероятностью 95,4% доля стандартных изделий в партии находится в пределах от 89,34% до 94,66%, а средний вес изделия – в пределах от 1037,46 до 1042,52

**Пример 2.** По данным пробного обследования среднее квадратическое отклонение веса нарезных батонов составило 15,4 г. Необходимо установить оптимальный объем выборки из партии нарезных батонов (2000 шт.), чтобы с вероятностью 0,997 предельная ошибка выборки не превысила 3% веса 500-граммового батона.

**Решение.** Итак, по условию

$$\sigma = 15,4 \text{ г.}$$

$$\Delta_{\text{относ}} = 3\%$$

$$N = 2000 \text{ шт.}$$

$$\bar{x} = 500 \text{ г.}$$

Заданную относительную ошибку выборки выразим абсолютной величиной:

$$\Delta_x = \frac{\Delta_{\text{относ}} \cdot \bar{x}}{100\%} = \frac{500 \cdot 3}{100} = 15 \text{ г.}$$

Значение коэффициента доверия, соответствующее вероятности 0,997,  $t=3$

Подставляем значения в формулу для бесповторного отбора:

$$n = \frac{N \cdot t^2 \sigma^2}{N \cdot \Delta_x^2 + t^2 \sigma^2} = \frac{2000 \cdot 3^2 \cdot 15,4^2}{2000 \cdot 15^2 + 3^2 \cdot 15,4^2} \approx 10 \text{ шт.}$$

Итак, для соблюдения указанных условий требуется провести обследование 10 батонов.

## 5. Задачи для самостоятельного решения

**Задача 1.** Для определения среднегодового стажа работы рабочих завода произведена десяти процентная бесповторная выборка.

|                   |      |     |     |     |      |       |
|-------------------|------|-----|-----|-----|------|-------|
| Стаж работы, годы | До 2 | 2-4 | 4-6 | 6-8 | 8-10 | 10-12 |
| Число рабочих     | 20   | 80  | 100 | 60  | 30   | 10    |

Определить с вероятностью 0,954:

1. Пределы, в которых находится средний стаж работы всех рабочих предприятия
2. Пределы, в которых находится доля рабочих со стажем до 6 лет.

## Тема 6

### СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СВЯЗИ

#### 1. Основные понятия и предпосылки корреляционно-регрессионного анализа

Большинство статистических исследований ставит своей целью выявление взаимозависимостей между признаками. Все статистические методы прогнозирования базируются на факте существования таких зависимостей, иначе прогноз стал бы невозможным. Признаки по их значению для изучения взаимосвязи делятся на два класса: факторные, или факторы – признаки, обуславливающие изменения других, связанных с ними, признаков, и результативные – признаки, изменяющиеся под действием факторных признаков.

Между общественными явлениями существует два типа связи: функциональная и корреляционная.

При *функциональной* связи изменение независимых переменных приводит к получению точно определенных значений зависимой переменной.

*Корреляционной* связью называется важнейший частный случай статистической связи, состоящий в том, что разным значениям одной переменной соответствуют различные *средние* значения другой переменной. В статистике принято различать следующие варианты зависимостей:

1. *парная* корреляция – связь между двумя признаками (результативным и факторным или двумя факторными)
2. *частная* корреляция – зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков.
3. *множественная* корреляция – зависимость результативного и двух или более факторных признаков, включенных в исследование

По направлению различают *прямую* связь, при которой с увеличением (уменьшением) значений факторного признака происходит увеличение (уменьшение) значений результативного, и *обратную* связь, при которой значения факторного признака изменяются под воздействием факторного в противоположном направлении.

*Корреляционно-регрессионный анализ* как общее понятие включает в себя измерение тесноты, направления связи и установление аналитической формы связи. *Корреляционный анализ* имеет своей задачей количественное определение тесноты связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи). *Регрессионный анализ* заключается в определении аналитического выражения связи, в котором изменение результативного признака обусловлено влиянием одного или нескольких факторов, а множество всех прочих факторов, также оказывающих влияние на зависимую величину, принимается за постоянные и средние значения. Корреляция и регрессия тесно связаны между собой: первая оценивает силу статистической связи, вторая исследует ее форму.

*Предпосылки корреляционно-регрессионного анализа.*

1. Наличие данных по достаточно большой совокупности явлений. Это общее условие всякого статистического исследования. Обычно считается, что число наблюдений должно быть в 5-6 (а лучше – не менее чем в 10 раз) больше числа факторов. Большое число наблюдений позволяет закону больших чисел, действуя в полную силу, обеспечить эффективное взаимопогашение случайных отклонений от закономерного характера связи признаков.

2. Качественная однородность тех единиц, которые подвергаются изучению методами корреляционно-регрессионного анализа.

3. При выполнении вышеуказанных требований далее необходимо провести количественную оценку однородности исследуемой совокупности по комплексу признаков. Одним из возможных вариантов такой оценки является расчет относительных показателей вариации (традиционно широкое применение для этих целей получил коэффициент вариации).

4. При ограничении числа факторов, вводимых в модель, наряду с качественным анализом целесообразно использовать и количественные оценки, позволяющие конкретно охарактеризовать влияние факторов на результирующий показатель. Включаемые в исследование факторы должны быть независимы друг от друга, так как наличие тесной связи между ними свидетельствует о том, что они характеризуют одни и те же стороны изучаемого явления и дублируют друг друга.

5. Целесообразным является изучение формы распределения исследуемых признаков, т.к. все основные положения теории корреляции разрабатывались применительно к предположению о нормальном характере распределения исследуемых признаков. Это условие связано с применением метода наименьших квадратов (МНК) при расчете параметров корреляции: только при нормальном распределении МНК дает оценку параметров, отвечающую принципам максимального правдоподобия. На практике эта предпосылка выполняется приближенно. Однако при значительном отклонении распределения признаков от нормального закона возникают проблемы с оценкой надежности рассчитанных по выборочным данным коэффициентов корреляции.

В соответствии с сущностью корреляционной связи ее изучение имеет две цели:

1. измерение тесноты связи двух или более признаков между собой
2. измерение параметров уравнения, выражающего зависимость средних величин результирующего признака от значений одного или нескольких факторных признаков;

## **2. Измерение степени тесноты корреляционной связи**

### **в случае парной зависимости**

Показатели тесноты связи используются для решения следующих задач:

1. Вопрос о необходимости изучения данной связи и целесообразности ее практического применения.
2. Вопрос о степени различий тесноты связи для конкретных условий.
3. Для выявления решающих факторов, воздействующих главным образом на формирование величины результирующего признака.

Теснота связи при линейной зависимости измеряется с помощью линейного коэффициента корреляции Пирсона:

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

Значение линейного коэффициента корреляции важно для исследования социально-экономических явлений и процессов, распределение которых близко к нормальному. Он принимает значения в интервале  $-1 \leq r \leq 1$ . Отрицательные значения указывают на обратную связь, положительные – прямую. При  $r=0$  линейная связь отсутствует. Чем ближе  $r$  по абсолютной величине к 1, тем теснее связь между признаками. При  $r=\pm 1$  связь функциональная.

Квадрат коэффициента корреляции  $r^2$  представляет собой *коэффициент детерминации*, который показывает долю вариации результирующего признака, объясненную влиянием вариации факторного признака.

Для оценки существенности (значимости) линейного коэффициента корреляции используется тот факт, что величина  $t_{расч} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$  при условии отсутствия связи в

генеральной совокупности распределена по закону Стьюдента с  $(n-2)$  степенями свободы (где  $n$  – объем выборки). Полученную  $t_{расч}$  сравнивают табличным значением. Коэффициент корреляции признается значимым при уровне значимости  $\alpha$ , если  $t_{расч} > t_{табл}$ . В этом случае практически невероятно, что найденное значение коэффициента корреляции обусловлено только случайными совпадениями. Уровень значимости  $\alpha$  показывает вероятность принятия ошибочного решения, например, при  $\alpha=0,05$  в среднем пяти случаях из ста есть риск сделать ошибочное заключение о значимости коэффициента корреляции (в социально-экономических исследованиях обычно  $\alpha=0,1$ ,  $\alpha=0,05$  или  $\alpha=0,01$ ).

### 3. Вычисление параметров уравнения регрессии

Задачи регрессионного анализа:

1. установление формы зависимости
2. определение функции регрессии
3. использование уравнения для оценки неизвестных значений зависимой переменной

Важнейшим этапом построения регрессионной модели является установление математической функции, которая лучше других выражает реальные связи между анализируемыми признаками. Выбор типа функции может опираться на теоретические знания об изучаемом явлении, опыт предыдущих аналогичных исследований, или осуществляться эмпирически – перебором и оценкой функций разных типов и т.п.

Уравнение однофакторной парной линейной корреляционной связи имеет вид:

$$y = a_0 + a_1 x,$$

где  $y$  – теоретические значения результативного признака, полученные по уравнению регрессии;

$a_0, a_1$  – параметры уравнения регрессии

Параметры уравнения  $a_0, a_1$  находят посредством МНК, при котором в качестве решения принимается точка минимума суммы квадратов отклонений эмпирических данных  $y_i$  от теоретических  $y$ , рассчитанных по модели, т.е.

$$\sum (y_i - y)^2 \rightarrow \min$$

Для нахождения минимума данной функции, ее частные производные приравнивают нулю и получают систему нормальных уравнений:

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

Решая систему в виде, получают значения параметров уравнения.

Параметр  $a_1$  называется *коэффициентом регрессии*. Его можно найти также по формуле:

$$a_1 = r_{xy} \frac{\sigma_y}{\sigma_x}$$

Коэффициент регрессии  $a_1$  показывает, насколько в среднем изменяется величина результативного признака (в его единицах измерения) при изменении факторного признака на единицу.

Параметр  $a_0$  показывает усредненное влияние прочих факторов на результативный признак. Параметр  $a_0$  связан с коэффициентом регрессии  $a_1$  соотношением

$$a_0 = \bar{y} - a_1 \bar{x}$$

Коэффициент регрессии  $a_1$  применяется также для расчета *коэффициента эластичности*, который показывает, на сколько процентов изменится величина результативного признака при изменении факторного признака на 1%:

$$\mathcal{E}_x = a_1 \frac{\bar{x}}{\bar{y}}$$

### 4. Примеры решения задач

**Пример 1.** Имеется следующая информация по 10 однотипным торговым предприятиям о возрасте типового оборудования (в годах) и затратах на его ремонт (в тыс. руб.).

Среднее значение возраста типового оборудования составило 7 лет, среднее квадратическое отклонение равно 2,43.

Среднее значение затрат на ремонт составило 2,7 тыс. руб, среднее квадратическое отклонение равно 1,3.

Среднее произведение значений признаков равно 21,71.

Оценить тесноту связи показателей, построить адекватную регрессионную модель.

**Решение.** Возраст оборудования – факторный признак ( $x$ ), влияющий на затраты на ремонт ( $y$ ). Итак,  $\bar{x}=7$ ,  $\bar{y}=2,7$ ,  $\overline{xy}=21,71$ ,  $\sigma_x=2,43$ ,  $\sigma_y=1,3$

Оценка тесноты связи

Рассчитаем коэффициент корреляции  $r = \frac{\overline{xy} - \bar{x} \times \bar{y}}{\sigma_x \times \sigma_y} = 0,89$

Значение коэффициента корреляции свидетельствует о возможном наличии сильной прямой связи между признаками.

Значимость коэффициента корреляции проверяется с помощью распределения Стьюдента.

$$t_{расч} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,89 * \sqrt{10-2}}{\sqrt{1-0,89^2}} = 3,69$$

С учетом уровня значимости  $\alpha=0,05$  и 8 степеней свободы табличное значение  $t_{табл}=2,3$ . Поскольку  $t_{расч} > t_{табл}$ , с вероятностью 0,95 можно утверждать, что между признаками существует сильная прямая связь.

Значение коэффициента детерминации  $r^2=0,89^2=0,792$  свидетельствует о том, что 79,2% общей вариации затрат на ремонт оборудования объясняется изменением возраста оборудования (а оставшиеся 20,8% - другими причинами).

Вычисление параметров уравнения регрессии

$$a_1 = r_{xy} \frac{\sigma_y}{\sigma_x} = 0,89 * \frac{1,3}{2,43} = 0,476$$

$$a_0 = \bar{y} - a_1 \bar{x} = 2,7 - 0,476 * 7 = -0,632$$

Подставляя значение найденных параметров в уравнение

$$\hat{y} = a_0 + a_1 x$$

получаем уравнение регрессии:

$$\hat{y} = -0,632 + 0,476 * x$$

Найденное значение коэффициента регрессии  $a_1 = 0,476$  говорит о том, что увеличение возраста оборудования в среднем на 1 год приводит к увеличению затрат на ремонт в среднем на 0,476 тыс.руб.

Коэффициент эластичности позволяет выразить эту взаимосвязь в процентах:

$$\mathcal{E}_x = a_1 \frac{\bar{x}}{\bar{y}} = 0,476 * \frac{7}{2,7} = 1,23$$

При увеличении возраста оборудования на 1% затраты на ремонт возрастают на 1,23%.

## 5. Задачи для самостоятельного решения.

**Задача 1.** По следующим данным оценить тесноту связи показателей, построить адекватную регрессионную модель, рассчитать коэффициент эластичности, сделать выводы.

$$\bar{x} = 17 \quad \bar{y} = 15,3 \quad \overline{xy} = 268,6 \quad \sigma_x = 3,4 \quad \sigma_y = 2,8$$

## Тема 7

### РЯДЫ ДИНАМИКИ И ИХ СТАТИСТИЧЕСКИЙ АНАЛИЗ

#### 1. Понятие о статистических рядах динамики

Основная цель статистического изучения динамики социально-экономических явлений состоит в выявлении и измерении закономерностей их развития во времени. Это достигается посредством построения и анализа статистических рядов динамики (или динамических рядов, или временных рядов).

*Рядами динамики* называются статистические данные, отображающие развитие изучаемого явления во времени. В каждом ряду динамики имеются два основных элемента:

1. показатель времени  $t$ . В качестве показателей времени в рядах динамики выступают либо определенные даты (моменты) времени, либо отдельные периоды (годы, кварталы, месяцы, сутки).
2. соответствующие им уровни развития изучаемого явления  $y$ . Уровнями ряда динамики называются отдельные наблюдения этого ряда. Уровни рядов динамики отображают количественную оценку (меру) развития во времени изучаемого явления. Они могут выражаться абсолютными, относительными и средними величинами.

В зависимости от характера изучаемого явления уровни рядов динамики могут относиться или к определенным датам (моментам) времени, либо к отдельным периодам. В соответствии с этим выделяют:

- *моментные* ряды динамики, которые отображают состояние изучаемых явлений на определенные даты (моменты) времени. Особенностью моментного ряда динамики является то, что в его уровни могут входить одни и те же единицы изучаемой совокупности. Посредством моментных рядов динамики изучаются показатели, отображающие состояние изучаемых явлений на отдельные даты, например, состояние кадров, товарные запасы, наличие основных фондов и т.д.

Пример моментного ряда динамики:

| Дата                   | 1.01.2001г. | 1.04.2001г. | 1.07.2001 г. | 1.10.2001 г. | 1.01.2002 г. |
|------------------------|-------------|-------------|--------------|--------------|--------------|
| Число работников, чел. | 192         | 190         | 195          | 198          | 200          |

- *интервальные* ряды динамики, которые отображают итоги развития (функционирования) изучаемых явлений за отдельные периоды (интервалы) времени. Особенностью интервального ряда динамики является то, что каждый уровень интервального ряда складывается из данных за более короткие интервалы. Свойство суммирования уровней за последовательные интервалы времени позволяет получать ряды динамики более укрупненных периодов.

Пример интервального ряда динамики:

| Год                                       | 1987  | 1988  | 1989  | 1990   | 1991   |
|-------------------------------------------|-------|-------|-------|--------|--------|
| Объем розничного товарооборота, тыс. руб. | 885.7 | 932.6 | 980.1 | 1028.7 | 1088.4 |

#### 2. Показатели динамики социально-экономических явлений.

Для количественной оценки динамики социально-экономических явлений применяются: абсолютные приросты, темпы роста и прироста, темпы наращивания и др.

Для расчета показателей рядов динамики *на постоянной базе* каждый уровень ряда сравнивается с одним и тем же базисным уровнем. Исчисляемые при этом показатели называются *базисными*.

Для расчета показателей динамики *на переменной базе* каждый последующий уровень ряда сравнивается с предыдущим. Вычисленные таким образом показатели динамики называются *цепными*.

✓ *Абсолютный прирост* - определяется в разностном сопоставлении двух уровней ряда динамики в единицах измерения исходной информации.

*Базисный абсолютный прирост*  $\Delta y_{\delta}$  исчисляется как разность между сравниваемым уровнем  $y_i$  и уровнем, принятым за постоянную базу сравнения  $y_o$ :

$$\Delta y_{\delta i} = y_i - y_o$$

*Цепной абсолютный прирост*  $\Delta y_{\psi}$  – разность между сравниваемым уровнем  $y_i$  и уровнем, который ему предшествует,  $y_{i-1}$ :

$$\Delta y_{\psi i} = y_i - y_{i-1}$$

Между базисными и цепными абсолютными приростами существует связь: сумма базисных абсолютных приростов  $\sum \Delta y_{\psi i}$  равна базисному абсолютному приросту последнего периода ряда динамики  $\Delta y_{\delta n}$ .

✓ *Темп роста* – характеризует отношение двух уровней ряда и может выражаться в виде коэффициента или в процентах:

*Базисные темпы роста*  $Tr_{\delta}$  исчисляются делением сравниваемого уровня  $y_i$  на уровень, принятый за постоянную базу сравнения,  $y_o$ :

$$Tr_{\delta i} = \frac{y_i}{y_o}$$

*Цепные темпы роста*  $Tr_{\psi}$  исчисляются делением сравниваемого уровня  $y_i$  на предыдущий уровень  $y_{i-1}$ :

$$Tr_{\psi i} = \frac{y_i}{y_{i-1}}$$

Между базисными и цепными темпами роста имеется взаимосвязь: произведение последовательных цепных темпов роста равно базисному темпу роста, а частное от деления последующего базисного темпа роста на предыдущий равно соответствующему цепному темпу роста.

✓ *Темп прироста* характеризует абсолютный прирост в относительных величинах. Исчисленный в процентах темп прироста показывает, на сколько процентов изменился сравниваемый уровень с уровнем, принятым за базу сравнения.

*Базисный темп прироста*  $Tn_{\delta}$  вычисляется делением сравниваемого базисного абсолютного прироста  $\Delta y_{\delta i}$  на уровень, принятый за постоянную базу сравнения  $y_o$ :

$$Tn_{\delta i} = \frac{\Delta y_{\delta i}}{y_o}$$

*Цепной темп прироста*  $Tn_{\psi}$  – это отношение сравниваемого цепного абсолютного прироста  $\Delta y_{\psi i}$  к предыдущему уровню  $y_{i-1}$ :

$$Tn_{\psi i} = \frac{\Delta y_{\psi i}}{y_{i-1}}$$

Между показателями темпа роста и прироста имеется взаимосвязь:

$$Tn_i (\%) = Tr_i (\%) - 100\% \quad (\text{при выражении темпа роста в процентах}).$$

$$Tn_i = Tr_i - 1 \quad (\text{при выражении темпа роста в коэффициентах}).$$

### 3. Средние показатели в рядах динамики

Для получения обобщающих показателей динамики социально-экономических явлений определяются средние величины: средний уровень, средний абсолютный прирост, средний темп роста и прироста и др.

✓ *Средний уровень ряда динамики* характеризует типическую величину абсолютных уровней.

В интервальных рядах динамики средний уровень  $\bar{y}$  определяется делением суммы уровней на их число  $n$ :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{y_1 + y_2 + \dots + y_n}{n}$$

В моментном ряду динамики с равностоящими датами времени средний уровень определяется по формуле средней хронологической:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1}$$

В моментном ряду динамики с неравноотстоящими датами времени средний уровень определяется по формуле:

$$\bar{y} = \frac{\sum t_i y_i}{\sum t_i} = \frac{t_1 y_1 + t_2 y_2 + \dots + t_n y_n}{t_1 + t_2 + \dots + t_n}, \quad \text{где}$$

$y_i$  – уровни ряда динамики, сохранившиеся без изменения в течение промежутка времени  $t_i$ .

✓ *Средний абсолютный прирост* представляет собой обобщенную характеристику индивидуальных абсолютных приростов ряда динамики. Для определения среднего абсолютного прироста  $\overline{\Delta y}$  сумма цепных абсолютных приростов  $\sum \Delta y_{цi}$  делится на их число: (где  $n$  – число уровней ряда)

$$\overline{\Delta y} = \frac{\sum \Delta y_{цi}}{n-1}.$$

Основываясь на взаимосвязи цепных и базисных абсолютных приростов, средний абсолютный прирост можно определить и по абсолютным уровням ряда динамики:

$$\overline{\Delta y} = \frac{y_n - y_0}{n-1} = \frac{\Delta y_{бn}}{n-1}.$$

✓ *Средний темп роста* – обобщающая характеристика индивидуальных темпов роста ряда динамики. Для определения среднего темпа роста  $Tr$  применяется формула средней геометрической:

$$\overline{Tr} = \sqrt[m]{Tr_{ц1} \cdot Tr_{ц2} \cdot \dots \cdot Tr_{цm}},$$

где  $Tr_{ц1}, Tr_{ц2}, \dots, Tr_{цn-1}$  – индивидуальные (цепные) темпы роста (в коэффициентах),  $m$  – число индивидуальных темпов роста ( $m=n-1$ , где  $n$  – число уровней ряда).

Основываясь на взаимосвязи между цепными и базисными темпами роста средний темп роста можно определить по формуле и по абсолютным уровням ряда динамики

$$\overline{Tr} = \sqrt[n-1]{\frac{y_n}{y_0}} = \sqrt[n-1]{Tr_{б}}$$

где  $n$  – число уровней ряда

✓ *Средний темп прироста  $\overline{Tn}$*  можно определить на основе взаимосвязи между темпами роста и прироста. При наличии данных о средних темпах роста  $\overline{Tr}$  для получения средних темпах прироста  $\overline{Tn}$  используется зависимость:

$$\overline{Tn} = \overline{Tr} - 1 \quad (\text{при выражении темпа роста в коэффициентах})$$

#### 4. Выявление и количественная оценка основной тенденции развития (тренда). Изучение периодических колебаний.

Одной из важнейших задач статистики является определение в рядах динамики общей тенденции развития явления. На развитие явления во времени оказывают влияние факторы, различные по характеру и силе воздействия. Одни из них оказывают практически постоянное воздействие и формируют в рядах динамики определенную тенденцию развития. Воздействие же других факторов может быть кратковременным или носить случайный характер.

Основная тенденция (*тренд*) – изменение, определяющее общее направление развития, это систематическая составляющая долговременного действия.

Задача - выявить общую тенденцию в изменении уровней ряда, освобожденную от действия различных случайных факторов. Методы выявления тренда:

1) *Метод укрупнения интервалов* основан на укрупнении периодов времени, к которым относятся уровни ряда динамики (одновременно уменьшается количество интервалов). Средняя, исчисленная по укрупненным интервалам, позволяет выявить направление и характер (ускорение или замедление роста) основной тенденции развития, в то время как слишком малые интервалы между наблюдениями приводят к появлению ненужных деталей в динамике процесса, засоряющих общую тенденцию.

2) *Метод скользящей средней*. Сущность его заключается в том, что исчисляется средняя уровень из определенного числа (обычно нечетного) первых по счету уровней ряда, затем - из такого же числа уровней, но начиная со второго по счету, далее - начиная с третьего и т.д. Таким образом, средняя как бы “скользит” по ряду динамики, передвигаясь на один срок. Недостатком сглаживания ряда является укорачивание сглаженного ряда по сравнению с фактическим, а, следовательно, потеря информации.

3) *Аналитическое выравнивание ряда динамики* используется для того, чтобы дать количественную модель, выражающую основную тенденцию изменения уровней динамического ряда во времени. Общая тенденция развития рассчитывается как функция времени:

$$\hat{y}_t = f(t), \text{ где}$$

$\hat{y}_t$  - уровни динамического ряда, вычисленные по соответствующему аналитическому уравнению на момент времени  $t$ .

Определение теоретических (расчетных) уровней  $\hat{y}_t$  производится на основе так называемой адекватной математической модели, которая наилучшим образом отображает (аппроксимирует) основную тенденцию ряда динамики. Простейшими моделями (формулами), выражающими тенденцию развития, являются:

$$\hat{y}_t = a_0 + a_1 t \quad - \text{линейная функция}$$

$$\hat{y}_t = a_0 \cdot a_1^t \quad - \text{показательная функция}$$

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 \quad - \text{степенная функция-кривая второго порядка (парабола)}$$

Расчет параметров функции обычно производится методом наименьших квадратов, в котором в качестве решения принимается точка минимума суммы квадратов отклонений между теоретическими и эмпирическими уровнями:

$$\sum (\hat{y}_t - y_t)^2 \rightarrow \min$$

где  $\hat{y}_t$  - выравненные (расчетные) уровни,  $y_t$  - фактические уровни.

Параметры  $a_i$ , удовлетворяющие этому условию, могут быть найдены решением системы нормальных уравнений. На основе найденного уравнения тренда вычисляются выравненные уровни. Т.о., выравнивание ряда динамики заключается в замене фактических уровней  $y_i$  плавно изменяющимися уровнями  $\hat{y}_i$ , наилучшим образом аппроксимирующими статистические данные.

*Периодические колебания* являются результатом влияния природно-климатических условий, общих экономических факторов, а также многочисленных и разнообразных факторов, которые часто являются регулируемы. В широком понимании к сезонным относят все явления, которые обнаруживают в своем развитии четко выраженную закономерность внутригодовых изменений, т.е. более или менее устойчиво повторяющиеся из года в год колебания уровней. Динамический ряд в этом случае называют *сезонным* рядом динамики.

Метод изучения и измерения сезонности заключается в построении специальных показателей, которые называются индексами сезонности. *Индексами сезонности* являются процентные отношения фактических внутригрупповых уровней к теоретическим уровням, выступающим в качестве базы сравнения. Для расчета индекса сезонности исходные данные берут за несколько лет и:

1. для каждого месяца рассчитывается средняя величина уровня
2. затем вычисляют среднемесячный уровень для всего ряда за несколько лет
3. определяют показатель сезонной волны - индекс сезонности  $i_s$  как процентное отношение средних для каждого месяца к общему среднемесячному уровню ряда, %:

$$I_s = (\bar{y}_i / \bar{y}) * 100,$$

где  $\bar{y}_i$  - средний уровень для каждого месяца,  $\bar{y}$  - среднемесячный уровень для всего ряда

Для наглядного изображения сезонной волны индексы сезонности изображают в виде графика.

#### 4. Примеры решения задач

**Пример 1.** По данным о величине уставного капитала банка рассчитать показатели динамики. Показать взаимосвязь показателей.

| Год  | Уставной капитал, млн. руб. |
|------|-----------------------------|
| 1998 | 5,08                        |
| 1999 | 5,5                         |
| 2000 | 5,9                         |
| 2001 | 6,15                        |

**Решение.**

1) Базисные абсолютные приросты  $\Delta y_{\delta i} = y_i - y_o$ :

1999 г.  $5,5 - 5,08 = 0,42$  млн.р.

2000 г.  $5,9 - 5,08 = 0,82$  млн.р.

2001 г.  $6,15 - 5,08 = 1,07$  млн.р.

2) Цепные абсолютные приросты  $\Delta y_{\eta i} = y_i - y_{i-1}$

1999 г.  $5,5 - 5,08 = 0,42$  млн.р.

2000 г.  $5,9 - 5,5 = 0,4$  млн.р.

2001 г.  $6,15 - 5,9 = 0,25$  млн.р.

3) Взаимосвязь базисных и цепных абсолютных приростов  $\Delta y_{\delta n} = \sum \Delta y_{\eta i}$

$1,07 = 0,42 + 0,4 + 0,25$  (млн.р.)

4) Базисные темпы роста  $Tr_{\delta i} = \frac{y_i}{y_o}$

1999 г.  $5,5 / 5,08 = 1,083 = 108,3\%$

2000 г.  $5,9 / 5,08 = 1,161 = 116,1\%$

2001 г.  $6,15/5,08=1,211=121,1\%$

5) Цепные темпы роста  $Tr_{u_i} = \frac{y_i}{y_{i-1}}$

1999 г.  $5,5/5,08=1,083 = 108,3\%$

2000 г.  $5,9/5,5 = 1,073 = 107,3\%$

2001 г.  $6,15/5,9=1,042 = 104,2\%$

6) Взаимосвязь базисных и цепных темпов роста

$1,211=1,083*1,073*1,042$

7) Базисные темпы прироста  $Tn_{\delta_i} = \frac{\Delta y_{\delta_i}}{y_o}$

1999 г.  $0,42/5,08= 0,083 = 8,3 \%$

2000 г.  $0,82/5,08= 0,163 = 16,1\%$

2001 г.  $1,07/5,08= 0,211 = 21,1\%$

8) Цепные темпы прироста  $Tn_{u_i} = \frac{\Delta y_{u_i}}{y_{i-1}}$

1999 г.  $0,42/5,08 = 0,083 = 8,3\%$

2000 г.  $0,4/5,5 = 0,073 = 7,3\%$

2001 г.  $0,25/5,9 = 0,042 = 4,2\%$

9) Взаимосвязь базисных темпов роста и прироста  $Tn_i(\%) = Tr_i(\%) - 100\%$  или  $Tn_i = Tr_i - 1$

|         |                        |                 |
|---------|------------------------|-----------------|
| 1999 г. | $8,3\%=108,3\%-100\%$  | $0,083=1,083-1$ |
| 2000 г. | $16,1\%=116,1\%-100\%$ | $0,161=1,161-1$ |
| 2001 г. | $21,1\%=121,1\%-100\%$ | $0,211=1,211-1$ |

10) Взаимосвязь цепных темпов роста и прироста  $Tn_i(\%) = Tr_i(\%) - 100\%$  или  $Tn_i = Tr_i - 1$

|         |                       |                 |
|---------|-----------------------|-----------------|
| 1999 г. | $8,3\%=108,3\%-100\%$ | $0,083=1,083-1$ |
| 2000 г. | $7,3\%=107,3\%-100\%$ | $0,073=1,073-1$ |
| 2001 г. | $4,2\%=104,2\%-100\%$ | $0,042=1,042-1$ |

11) Средний уровень ряда вычисляется по формуле  $\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + \frac{1}{2}y_n}{n-1}$ , т.к. исходные

данные – это моментный ряд с равноотстоящими датами

$$\bar{y} = \frac{\frac{1}{2} * 5,08 + 5,5 + 5,9 + \frac{1}{2} * 6,15}{4-1} = 5,67 \text{ млн.р.}$$

12) Средний абсолютный прирост

$$\overline{\Delta y} = \frac{\sum \Delta y_{u_i}}{n-1}, \quad \overline{\Delta y} = \frac{0,42 + 0,4 + 0,25}{4-1} = 0,36 \text{ млн.р.,}$$

$$\text{или } \overline{\Delta y} = \frac{\Delta y_{\delta_n}}{n-1} \quad \overline{\Delta y} = \frac{1,07}{4-1} = 0,36 \text{ млн.р.}$$

13) Средний темп роста

$$\overline{Tr} = \sqrt[n]{Tr_{u1} \cdot Tr_{u2} \cdot \dots \cdot Tr_{um}}, \quad \overline{Tr} = \sqrt[3]{1,083 \cdot 1,073 \cdot 1,042} = 1,066 = 106,6\%$$

$$\text{или } \overline{Tr} = \sqrt[n]{Tr_{\delta}} = \sqrt[3]{1,211} = 1,066 = 106,6\%$$

14) Средний темп прироста  $\overline{Tn} = \overline{Tr} - 1$ , или  $\overline{Tn} = \overline{Tr} - 100\%$

$$\overline{Tn} = 1,066 - 1 = 0,066,$$

$$\text{или } \overline{Tn} = 106,6\% - 100\% = 6,6\%$$

### 5. Задачи для самостоятельного решения

**Задача 1.** По данным о величине уставного капитала банка рассчитать показатели динамики, средние показатели ряда динамики. Показать взаимосвязь показателей.

| Годы                                 | 1991 | 1992 | 1993 | 1994 | 1995 |
|--------------------------------------|------|------|------|------|------|
| Производство тракторов<br>(тыс. шт.) | 45,0 | 47,8 | 50,4 | 55,3 | 58,2 |

**Задача 2.** По данным, характеризующим численность работающих в организации на первое число каждого месяца определить показатели динамики, средние показатели ряда динамики. Показать взаимосвязь показателей.

| Дата                      | 01.01 | 01.02 | 01.03 | 01.04 | 01.05 | 01.06 | 01.07 |
|---------------------------|-------|-------|-------|-------|-------|-------|-------|
| Численность<br>работающих | 224   | 229   | 232   | 236   | 229   | 230   | 234   |

## ИНДЕКСЫ И ИХ ИСПОЛЬЗОВАНИЕ В ЭКОНОМИКО-СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЯХ

### 1. Понятие, виды, свойства и основные задачи применения индексов в экономико-статистических исследованиях

*Статистический индекс* – это относительная величина сравнения сложных совокупностей и отдельных их единиц. При этом под сложной понимается такая статистическая совокупность, отдельные элементы которой по отдельности не подлежат суммированию.

Основным элементом индексного отношения является *индексируемая величина*, под которой понимается значение признака статистической совокупности, изменение которой является объектом изучения. Так, при изучении изменения цен индексируемой величиной является цена единицы товара  $p$ ; при изучении изменения физического объема товарной массы – данные о количестве товаров в натуральных измерителях  $q$ .

Способы построения индексов зависят от содержания изучаемых показателей, методологии расчета исходных статистических показателей, имеющихся в распоряжении исследователя статистических данных и целей исследования.

По степени охвата элементов совокупности различают индивидуальные и сводные (общие) индексы. *Индивидуальные* индексы (обозначаются буквой  $i$ ) характеризуют изменение только одного элемента совокупности. *Сводный (общий)* индекс (обозначается  $I$ ) отражает изменение по всей совокупности элементов сложного явления.

В зависимости от содержания и характера индексируемой величины различают индексы *количественных* (объемных) показателей (например, индекс физического объема продукции) и индексы *качественных* показателей (например, индексы цен, себестоимости).

В зависимости от методологии расчета различают *агрегатные* индексы и *средние из индивидуальных индексов* (или *преобразованную форму индексов*). Последние в свою очередь делятся на средние арифметические и средние гармонические.

Если сравнивают друг с другом не два момента (периода) времени, а более, то выделяют *цепную* и *базисную* систему индексов.

Индексы обладают синтетическими и аналитическими свойствами. Синтетические свойства состоят в том, что посредством индексного метода производится соединение (агрегирование) в целое разнородные единиц статистической совокупности. Аналитические свойства состоят в том, что посредством индексного метода определяется влияние факторов на изменение изучаемого показателя. Таким образом, с помощью индексных показателей решаются следующие основные задачи:

- 1) характеристика общего изменения сложного экономического показателя или формирующих его отдельных показателей-факторов;
- 2) выделение в изменении сложного показателя влияния одного из факторов путем элиминирования влияния других факторов.

Формулы для расчета индексов приведены далее на примере индексируемых цен ( $p$ ), физического объема продукции ( $q$ ), товарооборота ( $pq$ ), изменяющихся во времени.

### 2. Индивидуальные индексы и общие индексы в агрегатной форме

Динамика одноименных явлений изучается с помощью индивидуальных индексов

$$i_q = \frac{q_1}{q_0} \quad \text{- индивидуальный индекс физического объема продукции}$$

$$i_p = \frac{p_1}{p_0} \quad \text{- индивидуальный индекс цен}$$

$$i_{pq} = \frac{p_1 q_1}{p_0 q_0} \quad \text{- индивидуальный индекс товарооборота}$$

где подстрочное обозначение «0» соответствует уровню базисного периода (с которым сравнивают) или момента времени, «1» - уровню отчетного (сравниваемого) периода или момента времени.

Изменения совокупностей, состоящих из элементов, непосредственной несопоставимых (например, различных видов продукции), изучают с помощью сводных (общих) индексов. По методам расчета их подразделяют на агрегатные индексы и средние из индивидуальных.

Основной формой сводных (общих) индексов являются агрегатные индексы. В числителе и знаменателе общих индексов в агрегатной форме содержатся соединенные наборы (агрегаты) элементов изучаемой совокупности.

$$1) I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} \quad - \text{агрегатный индекс товарооборота}$$

где  $pq$  – индексируемое сложное явление.

Разница между числителем и знаменателем индекса составляет абсолютное изменение товарооборота:

$$\Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0$$

Это изменение товарооборота является результатом действия двух факторов: изменения физического объема продукции и изменения уровня цен.

Влияние изменения количества выпущенной продукции на изменение общего товарооборота отражается *агрегатным индексом физического объема*  $I_q$ . Влияние изменения цен выражается *агрегатным индексом цен*  $I_p$ .

$$2) I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} \quad - \text{агрегатный индекс физического объема продукции}$$

где  $q$  – индексируемая величина,

$p_0$  – соизмеритель, или вес, который фиксируется на уровне одного и того же периода. В практике статистики индексы количественных показателей исчисляются с базисными весами, а индексы качественных показателей - с отчетными весами. В данном случае вес фиксируется на уровне базисного периода

Разница между числителем и знаменателем индекса означает абсолютное изменение товарооборота (прирост или снижение) за счет изменения физического объема продукции:

$$\Delta_{pq}^q = \sum q_1 p_0 - \sum q_0 p_0$$

$$3) I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} \quad - \text{агрегатный индекс цен}$$

где  $p$  – индексируемая величина,

$q_1$  – соизмеритель, или вес, который фиксируется на уровне одного и того же периода (в данном случае – на уровне отчетного периода).

Разница между числителем и знаменателем индекса означает абсолютное изменение товарооборота (прирост или снижение) за счет изменения цен, или экономию (перерасход) потребителя за счет изменения цен.

$$\Delta_{pq}^p = \sum p_1 q_1 - \sum p_0 q_1$$

Между рассмотренными сводными индексами в агрегатной форме существует взаимосвязь:

$$I_{pq} = I_p * I_q,$$

кроме того,

$$\Delta_{pq} = \Delta_{pq}^q + \Delta_{pq}^p$$

### 3. Общие индексы в преобразованной форме (в форме средних из индивидуальных индексов).

Если неизвестна индексируемая величина за отчетный период или базисный период, но известна величина соответствующего индивидуального индекса, то используется преобразованная форма индекса. Сводный индекс тогда рассматривается как средняя величина соответствующих индивидуальных индексов, и рассчитать его можно как среднюю арифметическую или среднюю гармоническую.

Средняя арифметическая применяется, если есть данные для знаменателя, а числитель нужно получить путем преобразований. Средняя гармоническая применяется, если есть данные для числителя, а знаменатель надо получить путем преобразований.

Суть этого преобразования заключается в том, что на основе формул индивидуальных индексов в формулу сводного индекса вместо, например,  $p_0$  подставляется  $\frac{p_1}{i_0}$ , или вместо

$p_1$  подставляется  $i_p * p_0$ .

Индексы в форме средней арифметической:

$$I_{pq} = \frac{\sum i_{pq} * p_0 * q_0}{\sum p_0 * q_0} \quad - \text{сводный индекс товарооборота}$$

$$I_q = \frac{\sum i_q * q_0 * p_0}{\sum q_0 * p_0} \quad - \text{сводный индекс физического объема продукции}$$

$$I_p = \frac{\sum i_p * p_0 * q_1}{\sum p_0 * q_1} \quad - \text{сводный индекс цен}$$

Индексы в форме средней гармонической:

$$I_{pq} = \frac{\sum p_1 * q_1}{\sum \frac{p_1 * q_1}{i_{pq}}} \quad - \text{сводный индекс товарооборота}$$

$$I_q = \frac{\sum q_1 * p_0}{\sum \frac{q_1 * p_0}{i_q}} \quad - \text{сводный индекс физического объема продукции}$$

$$I_p = \frac{\sum p_1 * q_1}{\sum \frac{p_1 * q_1}{i_p}} \quad - \text{сводный индекс цен}$$

Значимость преобразованной формы индексов состоит в том, что количественный учет в современных условиях осуществляется не везде. Реализация товаров учитывается, как правило, в стоимостном выражении. В то же время для определения общих индексов цен в агрегатной форме необходимы данные о количестве отдельных товаров в натуральных измерителях. Индексы же в преобразованной форме используют в качестве весов осредняемых индивидуальных индексов реальные экономические категории, такие как:

$q_1 p_1$  и  $q_0 p_0$  - фактический товарооборот текущего и базисного периодов;

$z_1 q_1$  и  $z_0 q_0$  - фактические затраты на производство продукции в текущем и базисном периодах (здесь  $z$  – себестоимость единицы продукции)

и т.д.

В связи с этим в практике статистических расчетов широкое распространение получили расчет сводного индекса физического объема в форме средней арифметической и расчет сводного индекса цен (а также других качественных показателей: себестоимости, фондоотдачи, производительности труда и др.) в форме средней гармонической.

#### 4. Индексы переменного и постоянного состава и структурных сдвигов.

Индексный метод широко применяется для изучения динамики средних величин и выявления факторов, влияющих на динамику средних. С этой целью исчисляется система взаимосвязанных индексов: переменного, постоянного состава и структурных сдвигов.

*Индекс переменного состава*  $I_{\text{пер}}$  представляет собой отношение двух взвешенных средних величин, характеризующее изменение индексируемого (осредняемого) показателя.

$$I_{\text{пер}} = \frac{\bar{p}_1}{\bar{p}_0} = \frac{\sum p_1 * q_1}{\sum q_1} : \frac{\sum p_0 * q_0}{\sum q_0}$$

Величина этого индекса характеризует изменение средней взвешенной за счет влияния двух факторов: осредняемого показателя у отдельных единиц совокупности и структуры изучаемой совокупности.

*Индекс постоянного (фиксированного) состава*  $I_{\text{фикс}}$  представляет собой отношение средних взвешенных с одними и теми же весами (т.е. при постоянной структуре).

$$I_{\text{фикс}} = \frac{\bar{p}_1}{\bar{p}'} = \frac{\sum p_1 * q_1}{\sum q_1} : \frac{\sum p_0 * q_1}{\sum q_1}$$

Индекс постоянного состава учитывает изменение только индексируемой величины и показывает средний размер изменения изучаемого показателя у единиц совокупности.

*Индекс структурных сдвигов*  $I_{\text{стр}}$  характеризует влияние изменения структуры изучаемого явления на динамику среднего уровня индексируемого показателя.

$$I_{\text{стр}} = \frac{\bar{p}'}{\bar{p}_0} = \frac{\sum p_0 * q_1}{\sum q_1} : \frac{\sum p_0 * q_0}{\sum q_0}$$

Под структурными изменениями понимается изменение доли отдельных групп единиц совокупности к общей их численности.

Система взаимосвязанных индексов при анализе динамики средних величин имеет вид:

$$I_{\text{пер}} = I_{\text{фикс}} * I_{\text{стр}}$$

## 5. Примеры решения задач

**Пример 1.** Имеются данные по предприятию

| Изделие | Выпуск продукции, шт. $q$ |         | Цена единицы продукции, руб. $p$ |         |
|---------|---------------------------|---------|----------------------------------|---------|
|         | 2000г.                    | 2001 г. | 2000г.                           | 2001 г. |
| А       | 22000                     | 28000   | 2.0                              | 1.8     |
| Б       | 7000                      | 12000   | 6.0                              | 5.0     |
| В       | 2000                      | 5000    | 20.0                             | 18.0    |

Определить:

- 1) индивидуальные индексы физического объема продукции, цен и товарооборота по каждому изделию;
- 2) общий индекс товарооборота, агрегатные индексы физического объема и цен; абсолютные приросты товарооборота за счет изменения объемов производства, цен, за счет совместного действия обоих факторов;
- 3) показать взаимосвязь показателей.

**Решение.**

1) *Индивидуальные индексы физического объема*

$$i_{qA} = 28000/22000 = 127\% \text{ (рост на 27\%)}$$

$$i_{qB} = 12000/7000 = 171\% \text{ (рост на 71\%)}$$

$$i_{qV} = 5000/2000 = 250\% \text{ (рост в 2,5 раза)}$$

*Индивидуальные индексы цен*

$$i_{pA} = 1,8/2 = 0,9 = 90\% \text{ (снижение на 10\%)}$$

$$i_{pB} = 5/6 = 0,83 = 83\% \text{ (снижение на 17\%)}$$

$$i_{pV} = 18/20 = 0,9 = 90\% \text{ (снижение на 10\%)}$$

*Индивидуальные индексы товарооборота*

$$i_{pqA} = (28000 \cdot 1,8) / (22000 \cdot 2,0) = 114,5\% \text{ (рост на 14,5\%)}$$

$$i_{pqB} = (12000 \cdot 5,0) / (7000 \cdot 6,0) = 142,9\% \text{ (рост на 42,9\%)}$$

$$i_{pqB} = (5000 \cdot 18,0) / (2000 \cdot 20,0) = 225\% \text{ (рост 2,25\%)}$$

2) Изменение по предприятию в целом (по трем изделиям) индивидуальным индексом оценить нельзя, т.к. совокупность неоднородная. Поэтому воспользуемся сводным индексом. Сводный индекс общего товарооборота

$$I_{pq} = \frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{28000 \cdot 1,8 + 12000 \cdot 5 + 5000 \cdot 18}{22000 \cdot 2 + 7000 \cdot 6 + 2000 \cdot 20} = \frac{200400}{126000} = 159\%$$

Объем общего товарооборота вырос на 59%. В абсолютном выражении изменение товарооборота составляет:

$$\Delta_{pq} = \sum p_1 q_1 - \sum p_0 q_0 = 200,4 - 126 = 74,4 \text{ т.р.}$$

Этот рост достигнут за счет изменения двух факторов: изменения уровня цен и изменения количества продукции.

Агрегатный индекс физического объема

Поскольку данный индекс является индексом количественного показателя (объема продукции), вычислим его, применяя базисные веса, т.е. при расчете используем уровень цен базисного периода

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{28000 \cdot 2 + 12000 \cdot 6 + 5000 \cdot 20}{126000} = \frac{228000}{126000} = 181\%$$

Наблюдается рост физического объема продукции на 81%, в абсолютном выражении прирост физического объема продукции равен

$$\Delta_{pq}^q = \sum q_1 p_0 - \sum q_0 p_0 = 228 - 126 = 102 \text{ т.р.}$$

Агрегатный индекс цен

Поскольку данный индекс является индексом качественного показателя (цен), вычислим его, применяя отчетные веса, т.е. при расчете используем объем производства отчетного периода

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{200400}{228000} = 88\%$$

Цены снизились на 12%, экономия потребителя за счет изменения цен составила

$$\Delta_{pq}^p = \sum p_1 q_1 - \sum p_0 q_1 = 200,4 - 228 = -27,6 \text{ т.р. (знак «-» указывает на экономию, знак «+» - на перерасход денежных средств потребителя)}$$

3) Взаимосвязь показателей

$$I_{pq} = I_p \cdot I_q \quad 1,59 = 1,81 \cdot 0,88$$

$$\Delta_{pq} = \Delta_{pq}^q + \Delta_{pq}^p \quad 74,4 = 102 - 27,6 \text{ т.р.}$$

Общий вывод: Рост физического объема продукции на 81% обеспечил прирост товарооборота на 102 т.р. Одновременное снижение цен на 12% уменьшило товарооборот на 27,6 т.р. Совместное действие факторов обусловило рост товарооборота на 59%, или 74.4 т.р.

**Пример 2.** Имеются следующие данные

| Изделие | Цена единицы в базисном периоде $p_0$ | Выпуск в базисном периоде, шт $q_0$ | Изменение физического объема продукции в отчетном периоде по сравнению с базисным |
|---------|---------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------|
| А       | 110                                   | 12000                               | 1,10                                                                              |
| Б       | 16                                    | 15000                               | 1,15                                                                              |

Определить индивидуальные и общий индекс физического объема продукции

**Решение.**

По условию, индивидуальные индексы физического объема продукции по изделиям А и Б составили

$$i_{qA}=1.10$$

$$i_{qB}=1.15$$

Сводный индекс физического объема продукции  $I_q$  определим как среднюю арифметическую из двух индивидуальных индексов  $i_{qA}$  и  $i_{qB}$ . Исходные данные позволяют рассчитать  $I_q$  по формуле средней арифметической.

$$I_q = \frac{\sum q_1 * p_0}{\sum q_0 * p_0} = \frac{\sum i_q * q_0 * p_0}{\sum q_0 * p_0} = \frac{1.10 * 12000 * 110 + 1.15 * 15000 * 16}{12000 * 110 + 15000 * 16} = \frac{1728000}{156000} = 1.108,$$

Физический объем выпускаемой продукции вырос на 10,8%

**Пример 3.** Имеются следующие данные об издержках производства продукции по предприятию

| Изделие | Общие издержки производства (тыс. руб.) $z * q$ |                 | Изменение себестоимости единицы продукции в % к базисному периоду |
|---------|-------------------------------------------------|-----------------|-------------------------------------------------------------------|
|         | Базисный период                                 | Отчетный период |                                                                   |
| А       | 150,0                                           | 174,6           | +3                                                                |
| Б       | 289,0                                           | 323,0           | -5                                                                |

Определить среднее изменение себестоимости в отчетном периоде по сравнению с базисным.

**Решение.**

По условию, индивидуальные индексы себестоимости продукции по изделиям А и Б составили

$$i_{zA}=1.03$$

$$i_{zB}=0.95$$

Сводный индекс себестоимости  $I_z$  определим как среднюю арифметическую из двух индивидуальных индексов  $i_{zA}$  и  $i_{zB}$ . Исходные данные позволяют рассчитать  $I_z$  по формуле средней гармонической

$$I_z = \frac{\sum z1 * q1}{\sum z0 * q1} = \frac{\sum z1 * q1}{\sum \frac{z1 * q1}{i_z}} = \frac{174.6 + 323.0}{\frac{174.6}{1.03} + \frac{323.0}{0.95}} = \frac{497.6}{509.5} = 0.976 = 97.6\%$$

В среднем по предприятию себестоимость снизилась на 2,4%.

**Пример 4.** Имеются данные о производстве однородной продукции на двух предприятиях

| Предприятие | Выпуск, шт. $q$ |                 | Себестоимость единицы продукции $z$ |                 |
|-------------|-----------------|-----------------|-------------------------------------|-----------------|
|             | Базисный период | Отчетный период | Базисный период                     | Отчетный период |
| №1          | 18              | 20              | 5,0                                 | 4,5             |
| №2          | 22              | 30              | 4,6                                 | 3,8             |

Определить изменение средней себестоимости:

- 1) общее
- 2) за счет изменения себестоимости единицы продукции
- 3) за изменения структуры выпуска продукции
- 4) показать взаимосвязь системы индексов

**Решение.**

На изменение средних издержек влияют два фактора:

- себестоимость единицы продукции на каждом предприятии
- структура выпуска продукции

Необходимо учитывать как совместное влияние этих факторов, так и их отдельное влияние.

1) совместное влияние факторов на изменение средних издержек производства учитывает индекс переменного состава.

Он представляет собой соотношение двух средних величин, т.е. здесь учитываются и структурные изменения в составе совокупности, и изменение качественного признака у отдельных объектов.

$$I_{zme} = \frac{\bar{z}_1}{\bar{z}_0} = \frac{\sum z_1 * q_1}{\sum q_1} : \frac{\sum z_0 * q_0}{\sum q_0} = \frac{4.5 * 20 + 3.8 * 30}{20 + 30} : \frac{5.0 * 18 + 4.6 * 22}{18 + 22} = \frac{204}{50} : \frac{191.2}{40} = \frac{4.08(\bar{z}_1)}{4.78(\bar{z}_0)} = 85.4\%$$

Средняя себестоимость 1 изделия снизилась на 14,6% за счет совместного действия двух факторов

В абсолютном выражении это

$$\Delta^{zq} = (4,08 - 4,78) = -70 \text{ коп}$$

Т.е. средняя себестоимость 1 изделия снизилась на 70 коп.

2) изменение за счет качественного признака учитывает индекс фиксированного (постоянного) состава

$$I_{\text{фикс}} = \frac{\bar{z}_1}{z'} = \frac{\sum z_1 * q_1}{\sum q_1} : \frac{\sum z_0 * q_1}{\sum q_1} = 4.087 : \frac{5.0 * 20 + 4.6 * 30}{20 + 30} = \frac{4.087}{4.76} = 0.857 = 85.7\%$$

Средняя себестоимость снизилась на 14,3% за счет изменения себестоимости единицы продукции на каждом предприятии

В абсолютном выражении это

$$\Delta^z = (4,08 - 4,76) = -68 \text{ коп}$$

3) изменение структуры выпуска продукции (т.е. изменение доли предприятий в общем выпуске продукции) учитывает индекс структурных сдвигов.

$$I_{zcm} = \frac{\bar{z}'}{z_0} = \frac{\sum z_0 * q_1}{\sum q_1} : \frac{\sum z_0 * q_0}{\sum q_0} = \frac{4.76}{4.78} = 99,6\%$$

Средняя себестоимость снизилась на 0,4% за счет изменения структуры выпуска продукции.

В абсолютном выражении это

$$\Delta^q = (4,76 - 4,78) = -2 \text{ коп.}$$

4) Взаимосвязь системы индексов:

$$I_{\text{пер}} = I_{\text{фикс}} * I_{\text{стр.}}$$

$$0,854 = 0,857 * 0,996$$

Взаимосвязь абсолютных изменений:

$$\Delta^{zq} = \Delta^z + \Delta^q$$

$$-70 = -68 - 2$$

Общий вывод: если бы произошедшие изменения себестоимости продукции не сопровождалось структурными изменениями в ее выпуске, то средняя себестоимость снизилась бы на 14,3% (на 68 коп.). Изменение структуры выпуска продукции отдельных предприятий в общем объеме выпуска вызвало снижение себестоимости на 0,4% (2 коп.). Одновременное воздействие обоих факторов снизило среднюю себестоимость продукции на 14,6%, или 70 коп.

## 6. Задачи для самостоятельного решения.

**Задача 1.** Имеются данные по предприятию

| Изделие | Выпуск продукции, тыс. шт. |         | Цена единицы продукции, руб. |         |
|---------|----------------------------|---------|------------------------------|---------|
|         | 2000г.                     | 2001 г. | 2000г.                       | 2001 г. |
| А       | 23                         | 31      | 3                            | 2,8     |
| Б       | 8                          | 13      | 7                            | 6       |
| В       | 3                          | 6       | 21                           | 19      |

Определить:

- 1) индивидуальные индексы физического объема продукции, цен и товарооборота по каждому изделию;
- 2) общий индекс товарооборота, агрегатные индексы физического объема и цен; абсолютные приросты товарооборота за счет изменения объемов производства, цен, за счет совместного действия обоих факторов;
- 3) показать взаимосвязь показателей.

**Задача 2.** Имеются следующие данные

| Изделие | Цена единицы в базисном периоде | Выпуск в базисном периоде, тыс. шт. | Изменение физического объема продукции в отчетном периоде по сравнению с базисным |
|---------|---------------------------------|-------------------------------------|-----------------------------------------------------------------------------------|
| А       | 100                             | 12                                  | 1,15                                                                              |
| Б       | 12                              | 150                                 | 1,2                                                                               |

Определить индивидуальные и общий индекс физического объема продукции

**Задача 3.** Имеются следующие данные об издержках производства продукции по предприятию

| Изделие | Общие издержки производства (тыс. руб.) |                 | Изменение себестоимости единицы продукции в % к базисному периоду |
|---------|-----------------------------------------|-----------------|-------------------------------------------------------------------|
|         | Базисный период                         | Отчетный период |                                                                   |
| А       | 170                                     | 186             | +4                                                                |
| Б       | 300                                     | 320             | -2                                                                |

Определить среднее изменение себестоимости в отчетном периоде по сравнению с базисным.

**Задача 4.** Имеются данные о производстве однородной продукции на двух предприятиях

| Предприятие | Выпуск, шт.     |                 | Себестоимость единицы продукции |                 |
|-------------|-----------------|-----------------|---------------------------------|-----------------|
|             | Базисный период | Отчетный период | Базисный период                 | Отчетный период |
| №1          | 38              | 45              | 4,0                             | 3,5             |
| №2          | 40              | 50              | 3,6                             | 2,8             |

Определить изменение средней себестоимости:

- 1) общее
- 2) за счет изменения себестоимости единицы продукции
- 3) за изменения структуры выпуска продукции

## Список рекомендуемой литературы

### *Основная*

1. Общая теория статистики: статистическая методология в изучении коммерческой деятельности: Учебник / Под ред. А.А.Спирина, О.Э.Башиной. М.: Финансы и статистика, 2000.
2. Общая теория статистики: Учебник / Под ред. М.Р.Ефимовой. М.: ИНФРА-М, 2001.
3. Теория статистики: Учебник / Под ред. Р.А.Шмойловой. М.: Финансы и статистика, 1998.
4. Статистика: курс лекций / Под ред. В.Г. Ионина. - Новосибирск: Изд-во НГАЭиУ, М.: ИНФРА-М, 1998.
5. Экономическая статистика: Учебник / Под ред. Ю. Н. Иванова. - М.: Финансы и статистика, 1998.
6. Гусаров В.М. Статистика: Учебное пособие для вузов. – М.: ЮНИТИ-ДАНА, 2001 г.- 463 с.
7. Сборник задач по общей теории статистики: Учеб. Пособие для вузов/ Под ред. Серга Л.К. – М.: Филинь, 1999. - 360 с.
8. Практикум по статистике: учеб. пособие для вузов/ Под ред. В.М.Симчеры – М.: ЗАО «Финстатинформ», 1999

### *Дополнительная*

1. Социальная статистика: Учебник / Под ред. И.И. Елисеевой. - М.: Финансы и статистика, 1997.
2. Российский статистический ежегодник/ Госкомстат России. М.: Финансы и статистика.
3. Вестник статистики (ежемесячный журнал).
4. Статистика финансов: Учебник для вузов / Под ред. В.Н.Самина. - М.: Финансы и статистика, 2000. - 816 с.
5. Боровиков В.П. Популярное введение в программу STATISTICA.-М.: Компьютер Пресс, 1998.- 267 с.