

# ОСНОВЫ СТАТИСТИКИ И БУХГАЛТЕРСКОГО УЧЕТА

# Тема 3

## Сводка и группировка: изучение методов обобщения первоначальных данных

3.1 Суть и организация статистической сводки

3.2 Группировка статистических данных

# 3.1 Суть и организация статистической сводки

**Сводка**, как второй этап статистического исследования, — это научная обработка данных наблюдения для последующего описания статистической совокупности по обобщающим показателям.

**Суть статистической сводки** заключается в объединении единиц совокупности в группы, классы, типы и агрегировании информации о них как в границах групп, так и в целом по совокупности.

**Основная задача сводки** заключается в выявлении типичных черт и закономерностей в совокупности.

**Вся многогранная и сложная работа по статистической сводке состоит из следующих этапов:**

- 1) Формирование цели и задач статистической сводки;
- 2) Формирование групп из единиц наблюдения, определение группировочных признаков, количества групп и величины интервала; решение вопросов, связанных с осуществлением группировки, выделение существенных признаков;
- 3) 3. Осуществление технической стороны сводки, т.е. проверка полноты и качества собранного материала, подсчет результатов и необходимых показателей для характеристики всей совокупности и ее частей.

## Программа статистической сводки устанавливает такие виды работ:

- 1) выбор группировочных признаков;
- 2) установление порядка формирования групп;
- 3) разработка системы статистических показателей для характеристики групп и объекта в целом;
- 4) разработка макетов статистических таблиц для представления результатов сводки;
- 5) выбор способа сводки данных статистического наблюдения.

# СВОДКИ БЫВАЮТ:

## 1. По сложности построения сводки бывают простые и сложные.

*Простая сводка — это подведение итогов первичного материала в целом без любой его систематизации.*

*Сложная сводка соединяет комплекс операций: группировка единиц; подведение групповых и общих итогов; представление результатов сводки в форме статистических таблиц, графиков, рядов распределения.*

## 2. По организации работы разделяют централизованную и децентрализованную сводку.

*При централизованной сводке все данные сосредотачиваются в одном месте (центре), где и обрабатываются. Такой способ сводки используется при переписях населения, одноразовых статистических обследованиях, социологических опросах.*

*При децентрализованной сводке обработка статистической информации осуществляется от низшего до высшего звена управления.*

***В отечественной статистической практике обработка информации происходит преимущественно децентрализованно.***

**На втором этапе статистического исследования** осуществляется распределение совокупности на группы, однородные в том или другом понимании.

*Основой разграничения элементов в совокупности на группы может быть любой признак (атрибутивный или количественный), который имеет качественно отличную характеристику.*

*Такой признак называют **группировочным признаком**.*

Если разграничение элементов совокупности на группы осуществляется по:

- 1) **атрибутивным признакам**, то такой **вид группировки называют классификацией или номенклатурой**. Они разрабатываются международными и национальными статистическими органами и рекомендуются как статистический стандарт.
- 2) **количественным показателям**, то такой вид работ (в отличие от классификаций) называют в узком смысле непосредственно **статистической группировкой**.

## 3.2 Группировка статистических данных

**В зависимости от цели и задач исследования различают такие виды группировки:**

- 1) Типологическая группировка** — это разделение исследуемой качественно разнородной совокупности на классы, социально-экономические типы, однородные группы единиц в соответствии с правилами научной группировки. *Примерами таких видов группировок являются группировка населения по возрасту, группировка предприятий по форме собственности и т.п.*
- 2) Структурные группировки** характеризуют состав однородной совокупности по любым признакам. *К ним принадлежат группировки населения по полу, возрасту, а на производстве — группировка рабочих по производственному стажу, уровню квалификации и др.*
- 3) Аналитическая группировка** — используется для выявления взаимосвязи между признаками. Для этого в основание группировки должен быть положен факторный признак, затем каждая выделенная группа характеризуется средним значением результативного признака. *Примерами таких группировок могут быть группировки, в которых, например, изучаются взаимосвязи между себестоимостью и ее факторами, производительностью труда и ее факторами и т.п.*
- 4) Комбинационная группировка** — когда в основание группировки положено несколько признаков (*например, группировка по возрасту, затем каждая группа дополняется группировкой по полу и т.д.*).

**Группировка по одному признаку будет считаться *простой*, а по двум и более признакам, т.е. с разбиением групп на подгруппы, — *сложной*.**

Принципиальное значение при построении группировок имеет выбор **группировочного признака**, на основе которого выделяют разные типы, группы и подгруппы.

Если группировка осуществляется по *атрибутивному признаку*, то выделяют столько групп, сколько имеется наименований признака.

При составлении группировок на основе *количественных признаков* (дискретных или непрерывных) определяют количество групп и интервалы группировки.

В массовой совокупности оптимальное количество групп с **равными интервалами** приблизительно можно определить по формуле американского ученого Стерджеса:

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

где  $m$  — количество групп;  $n$  — объем совокупности.

**Интервалы**, т.е. промежуток между значениями признака в группе единиц, бывают равные, неравные, открытые и закрытые.

Выбор вида интервала зависит от характера распределения единиц исследуемой совокупности.

Равные интервалы используют в случаях, когда значения варьируемого признака  $x$  изменяются плавно, постепенно, равномерно.

Ширина интервала  $h$  определяется по формул

где  $x_{\max}$   $x_{\min}$  — наибольшее и наименьшее значение признака в совокупности.

$$h = \frac{x_{\max} - x_{\min}}{m},$$

*Например:*

Прибыльность активов коммерческих банков колеблется в пределах от 5 к 45%.

При принятии количества групп  $m = 4$  ширина интервала  $h = (45 - 5) : 4 = 10$ .

Тогда границы интервалов составляют соответственно: 5—15, 15—25, 25-35, 35-15.

Поскольку границы интервалов совпадают (15 — в первой и второй группе, 25 — во второй и третьей, 35 — в третьей и четвертой), то для исключения неопределенности отнесения граничных значений признака к той или другой группе используют правило: **левое одинаковое число не включает в себя значения признака, правое — включает.**

Все интервалы в данном примере называются закрытыми.

Приведенное распределение прибыльности активов банков может быть представлено в другом виде: до 15, 15-25, 25-35, 35 и более.

Первый и последний интервалы имеют лишь одну границу и называются открытыми.

Неравные интервалы используются в случае, когда диапазон значений признака слишком широкий и распределение совокупности по этому признаку неравномерно.

Например, распределение поселков городского типа по количеству жителей (тыс.

числ.) до 2, 2-4, 4-5, 5-10, 10-100

*Особый вид группировок в статистике представляют ряды распределения, которые являются самым простым способом обобщения статистических данных.*

**Рядом распределения** называют группировку, характеризующую состав (структуру) явления в данный период времени.

В зависимости от того, какой признак (качественный или количественный) положен в основу группировки, **ряды распределения** бывают **атрибутивными (качественными)** или **вариационными (количественными)**.

Примером атрибутивных рядов может быть распределение населения по полу, занятости, национальности, профессии и т.п.

В таблице приведен атрибутивный ряд распределения студентов экономического факультета университета по специальностям:

Название специальности	Число студентов, чел., $f_i$	% от общего количества, $\varphi_i$
Финансы	262	27,3
Учет и аудит	279	29,0

*Окончание табл. 3.1*

Менеджмент организаций	246	25,6
Маркетинг	174	18,1
Всего	961	100

**Элементами (характеристиками) этого ряда распределения являются:**

- 1) значения атрибутивного признака (первая графа таблицы);
- 2) **частоты  $f_i$**  — численные характеристики отдельных значений признака, т.е. числа, которые показывают, как часто встречается то или другое значение признака в ряду (вторая графа);
- 3) **частоты  $\varphi_i$**  — это частоты, выраженные в относительных величинах (коэффициентах или процентах), что приведено в третьей графе таблицы.

# Вариационные ряды распределения бывают дискретными и интервальными.

1. **Дискретные вариационные ряды** основаны на величинах признаков, которые имеют целые значения (например, тарифный разряд рабочих, количество марок автомобилей и т.п.).

Количество детей, $x_i$	Количество семей, $f_i$	В % от общего количества семей, $\varphi_i$
1	63	56,2
2	48	42,8
3	1	1,0
Всего	112	100

*Таблицей иллюстрируется дискретный ряд распределения семей в населенном пункте по количеству детей*

2. В интервальных вариационных рядах группировочный признак может принимать любое значение (целое, дробное) в пределах каждого интервала (например, распределение заработной платы работающих в организации, распределение основных фондов предприятия и т.п.).

Размер капитала, млн д. е., $x_i$	Прибыль, млн д. е., $f_i$	В % к общей прибыли, $\varphi_i$
3–6	9	34,6
6–9	13	50,0
9–12	4	15,4
Всего	26	100

*В таблице приведен интервальный ряд распределения коммерческих банков в населенном пункте по величине прибыли.*

**К основным элементам вариационных рядов распределения относятся:**

- a) варианты  $x_i$  — численные значения количественного признака в группировке (первая графа таблицы), которые могут быть положительными, отрицательными, абсолютными, относительными;
- b) частоты  $f_i$  — значение отдельных вариантов (вторая графа);
- c) частоты  $\varphi_i$  (третья графа).

Кроме отмеченных элементов вариационных рядов, к их характеристикам могут быть отнесены также такие показатели:

- a) плотность частоты  $f_i/h$  ( $h$  — ширина интервала);
- b) накопленная частота  $F = \sum \varphi_i$

Такие показатели используются при анализе построенных рядов распределения.

**Использование характеристик вариационного интервального ряда на примере распределения сельских хозяйств области:**

Количество голов, $x_i$	Количество хозяйств, $f_i$ %	Плотность распределения, $f_i/h \times 100\%$	Накопленная частота, $F = \sum_{i=1}^k \varphi_i$
До 300	4,4	1,47	4,4

*Окончание табл. 3.4*

300–599	11,3	3,77	15,7
600–999	24,5	6,12	40,2
1 000–3 000	53,7	2,69	93,9
Более 3 000	6,1	0,303	100
Всего	100	x	x