

Статистика

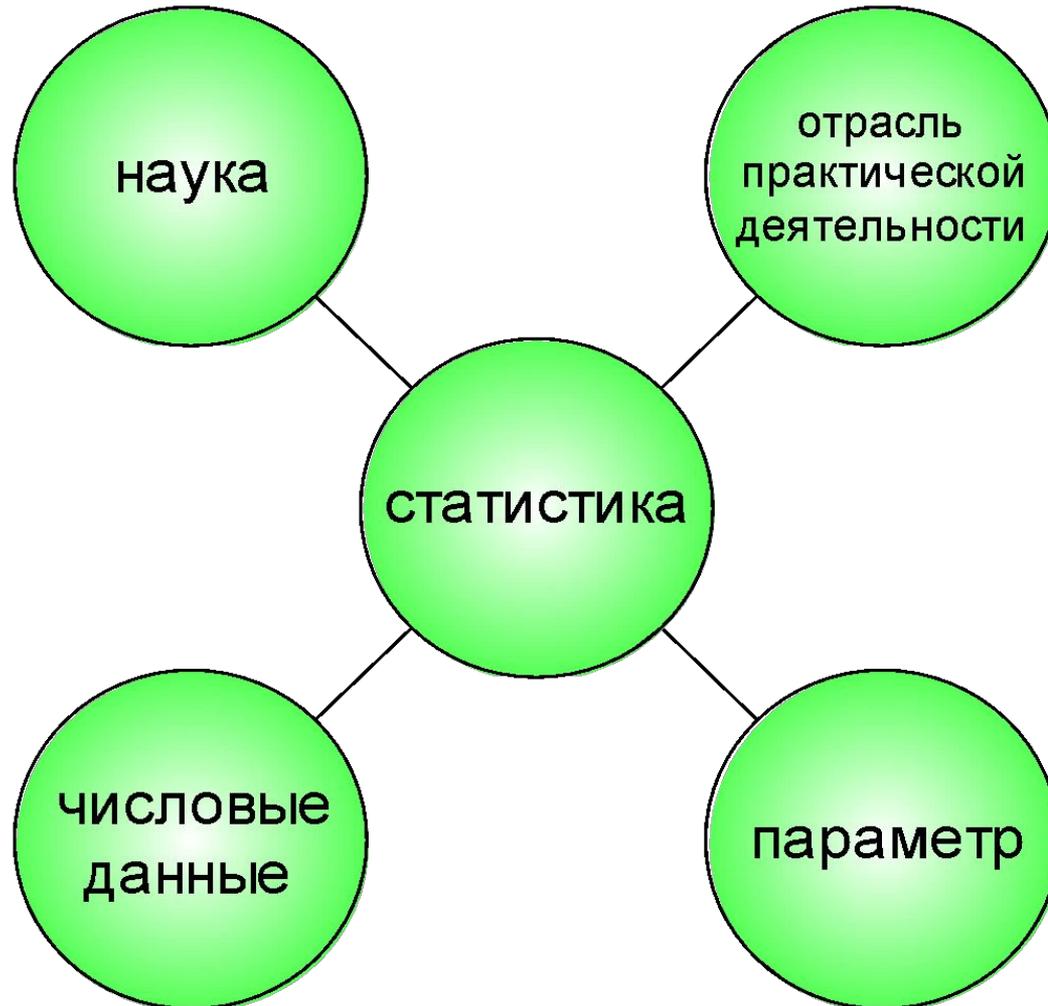
Курс лекций для дистанционного
обучения

К.э.н., доц. Силова Е.С.

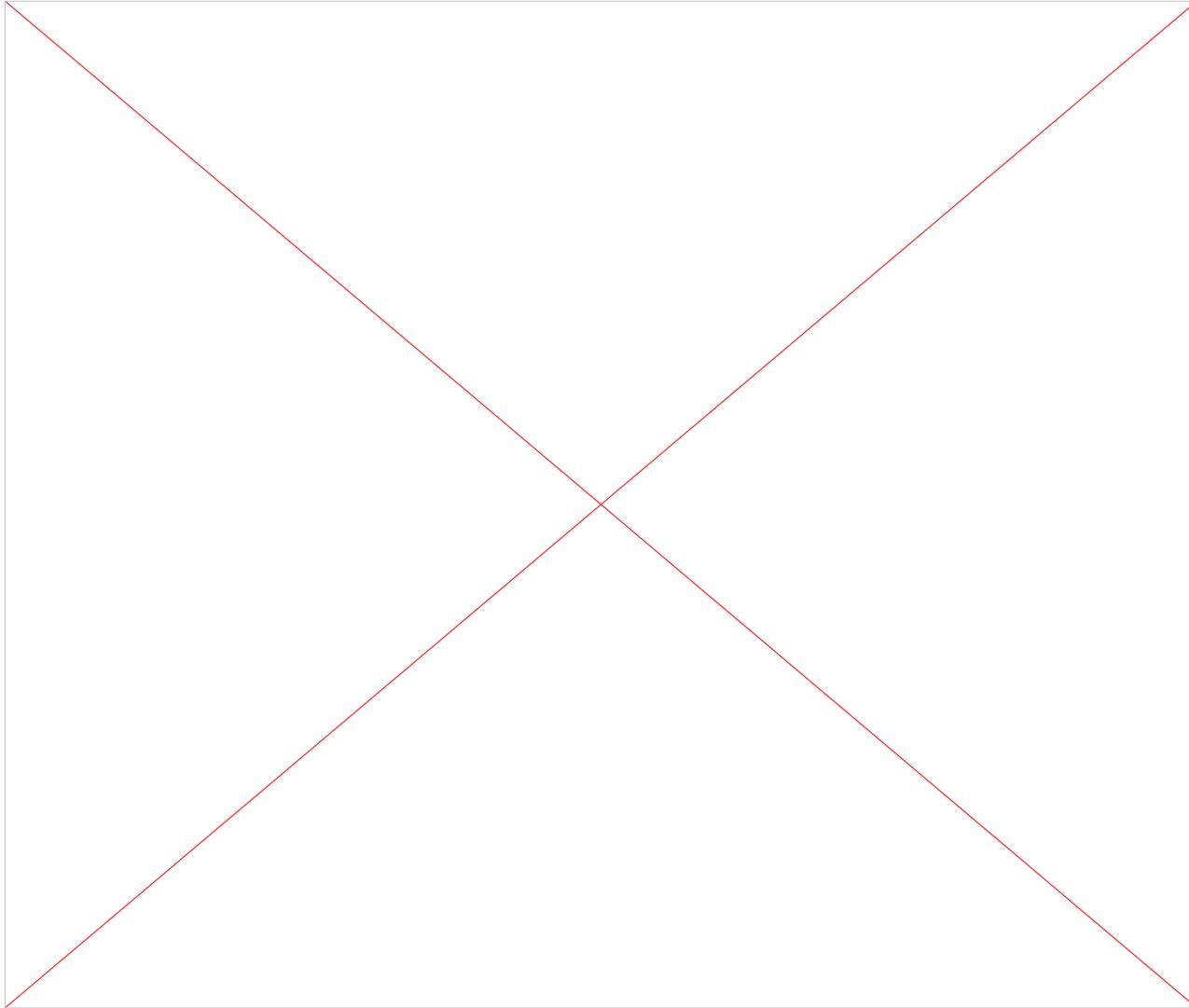
Модуль 1. Теоретические основы статистики

Статистика – общественная наука, которая изучает количественную сторону качественно определенных массовых социально-экономических явлений и процессов, их структуру и распределение, размещение в пространстве, движение во времени, выявляет действующие количественные зависимости, тенденции и закономерности, причем в конкретных условиях места и времени.

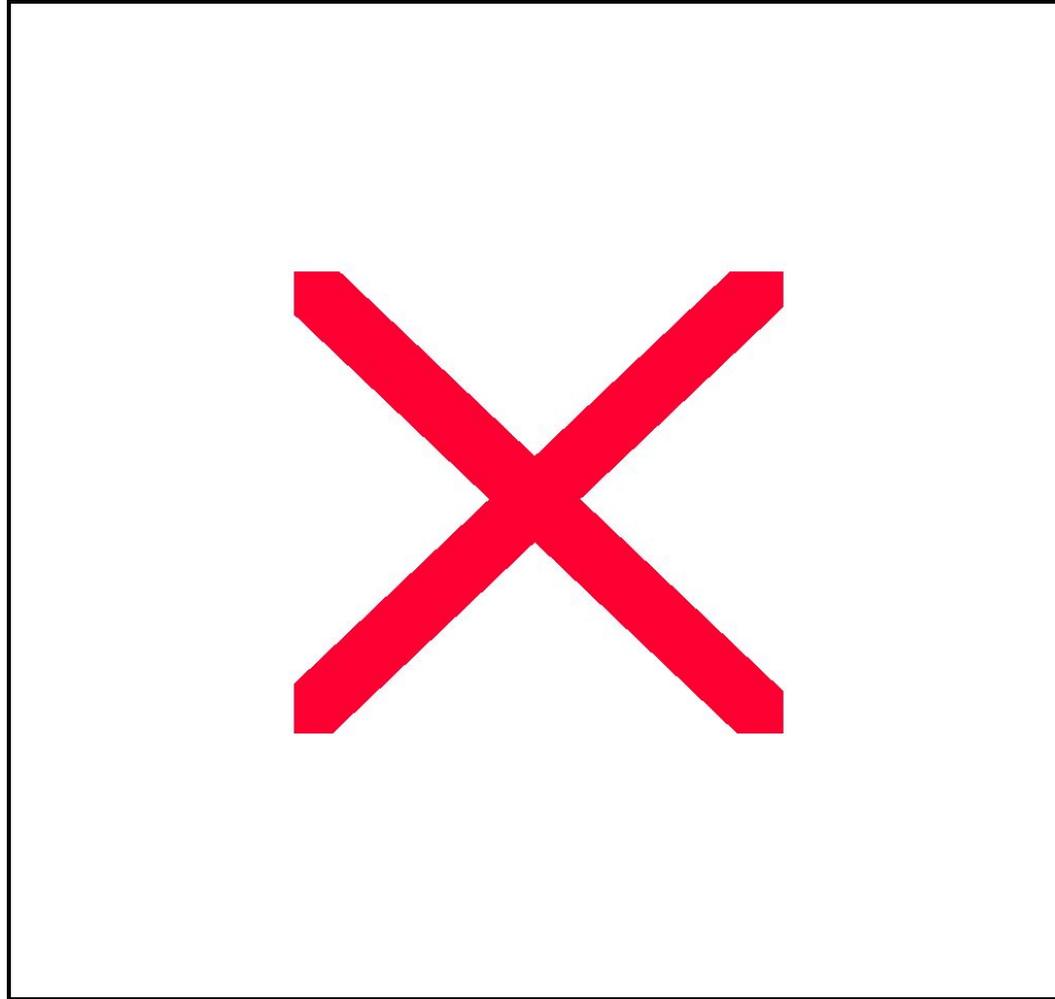
1.1 Статистика как наука и предмет ее изучения

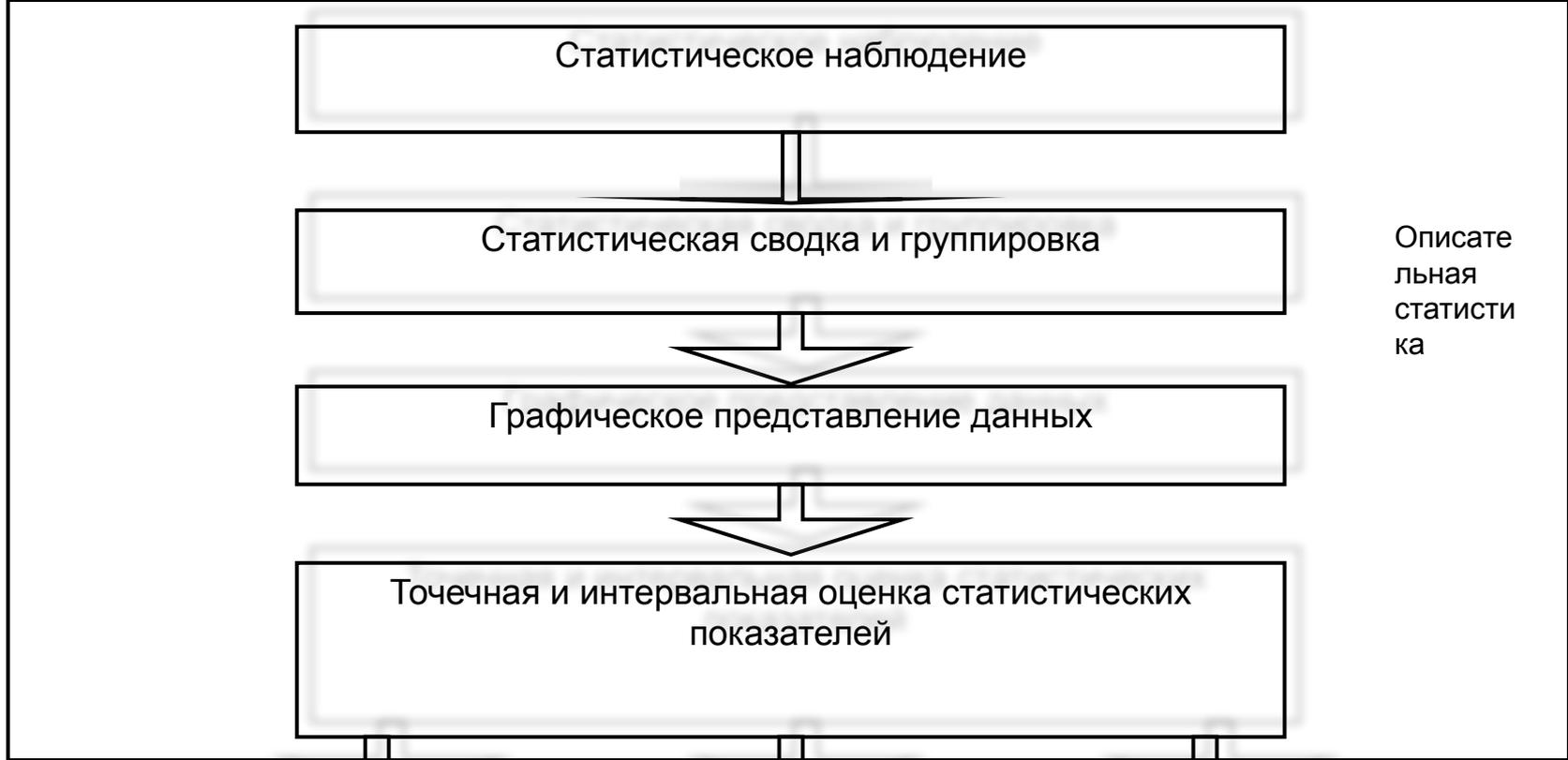


1.1 Статистика как наука и предмет ее изучения

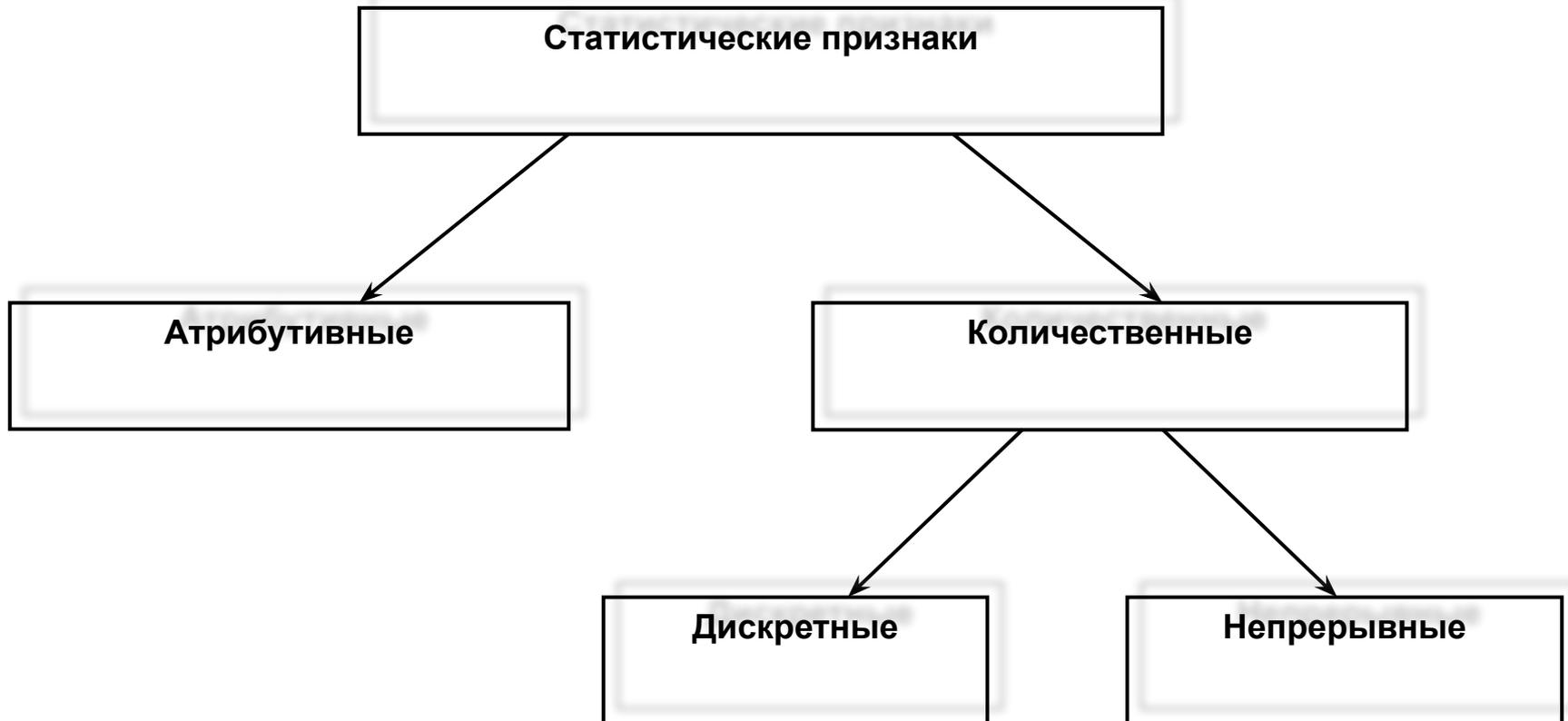


1.1 Статистика как наука и предмет ее изучения

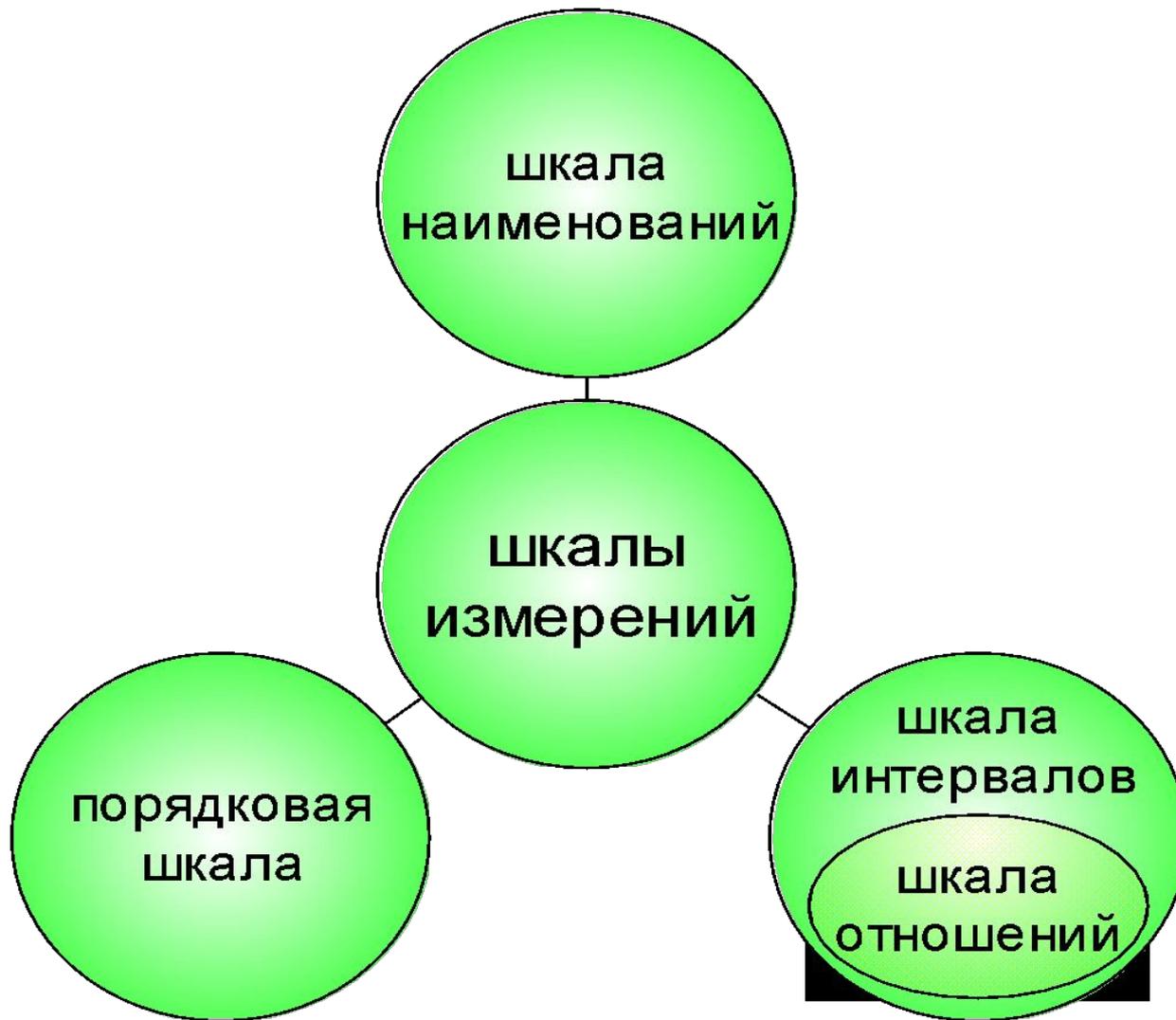




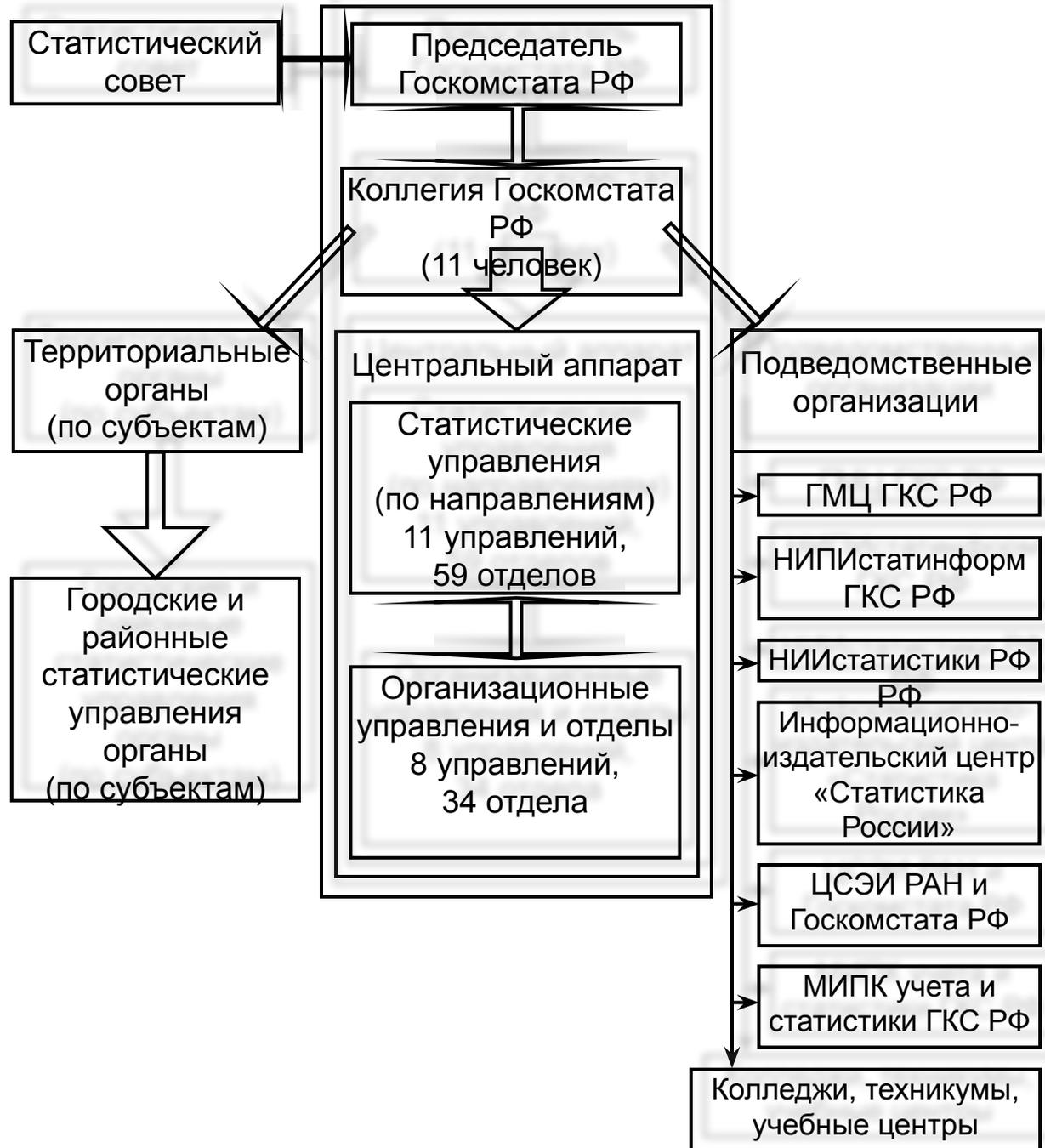
1.1 Статистика как наука и предмет ее изучения



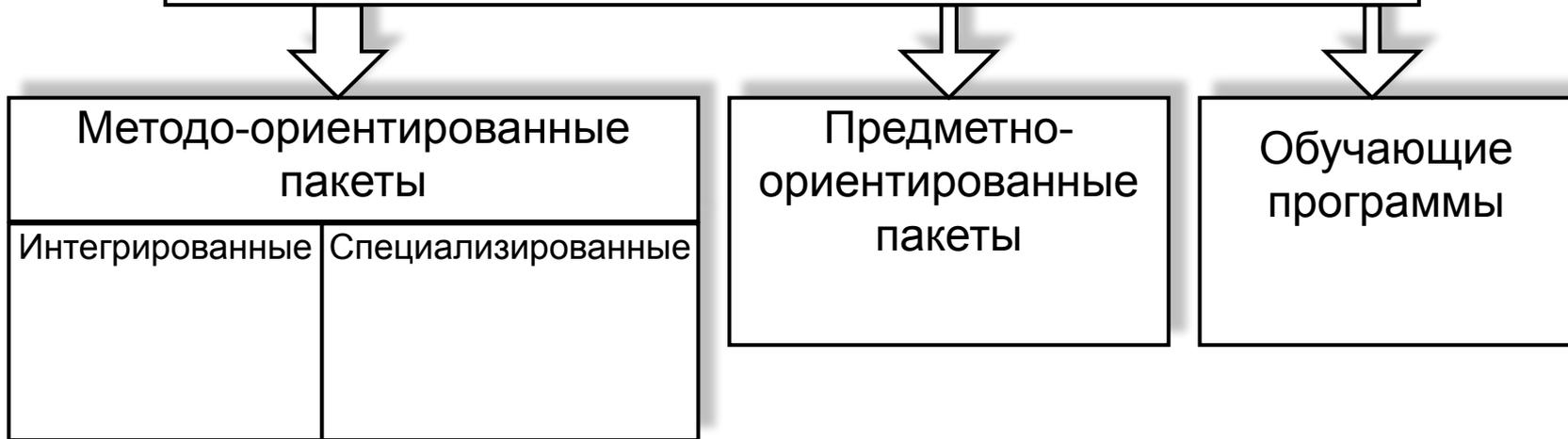
1.1 Статистика как наука и предмет ее изучения



- Организационная структура Госкомстата РФ



Виды статистических программных продуктов



- В настоящее время наибольшее распространение получили методо-ориентированные пакеты.

Модуль 2. Статистическое наблюдение

Статистическое наблюдение – это массовое, планомерное, научно организованное наблюдение за явлениями социальной и экономической жизни, которое заключается в регистрации отобранных признаков у каждой единицы совокупности.

Теория статистического наблюдения

Этапы статистического наблюдения:

- **Подготовка статистического наблюдения ;**
- **Проведение массового сбора данных;**
- **Подготовка данных к автоматизированной обработке;**
- **Разработка рекомендаций по совершенствованию статистического наблюдения.**

Источники статистической информации

Источники статистической информации

Внутренние источники

Отчетность предприятий

Бухгалтерская отчетность

Статистическая отчетность

Регистр предприятий и организаций

Переписи и обследования

Внешние источники

Административные источники

Отчет об исполнении госбюджета

Налоговая статистика

Денежная и банковская статистика

Платежный баланс

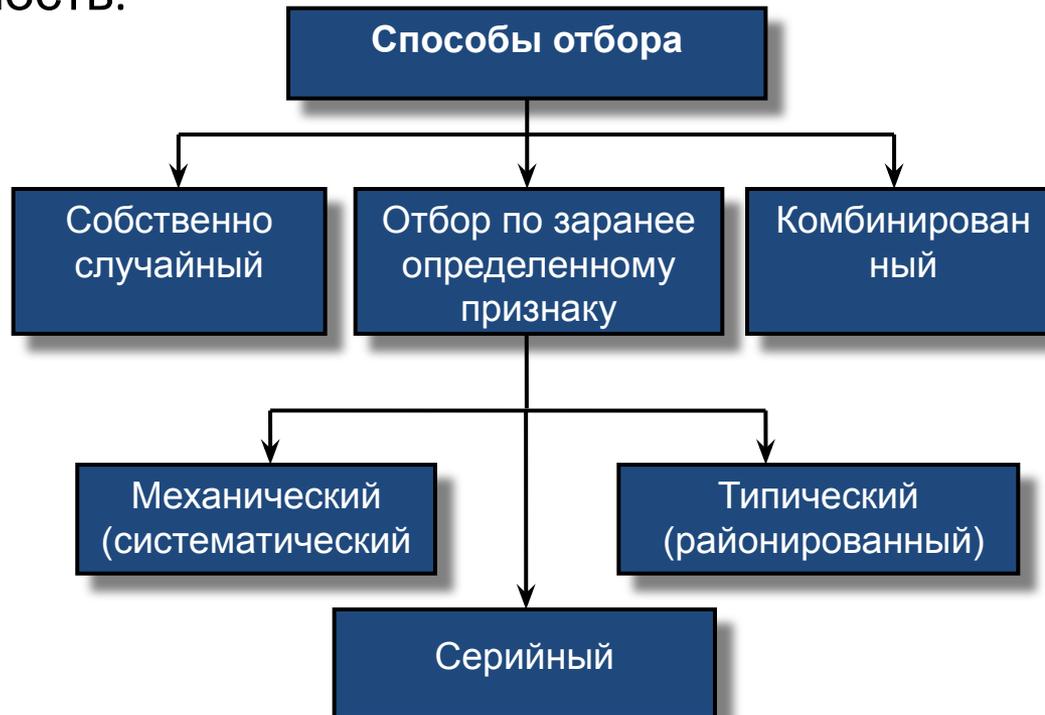
Таможенная статистика

Виды статистического наблюдения

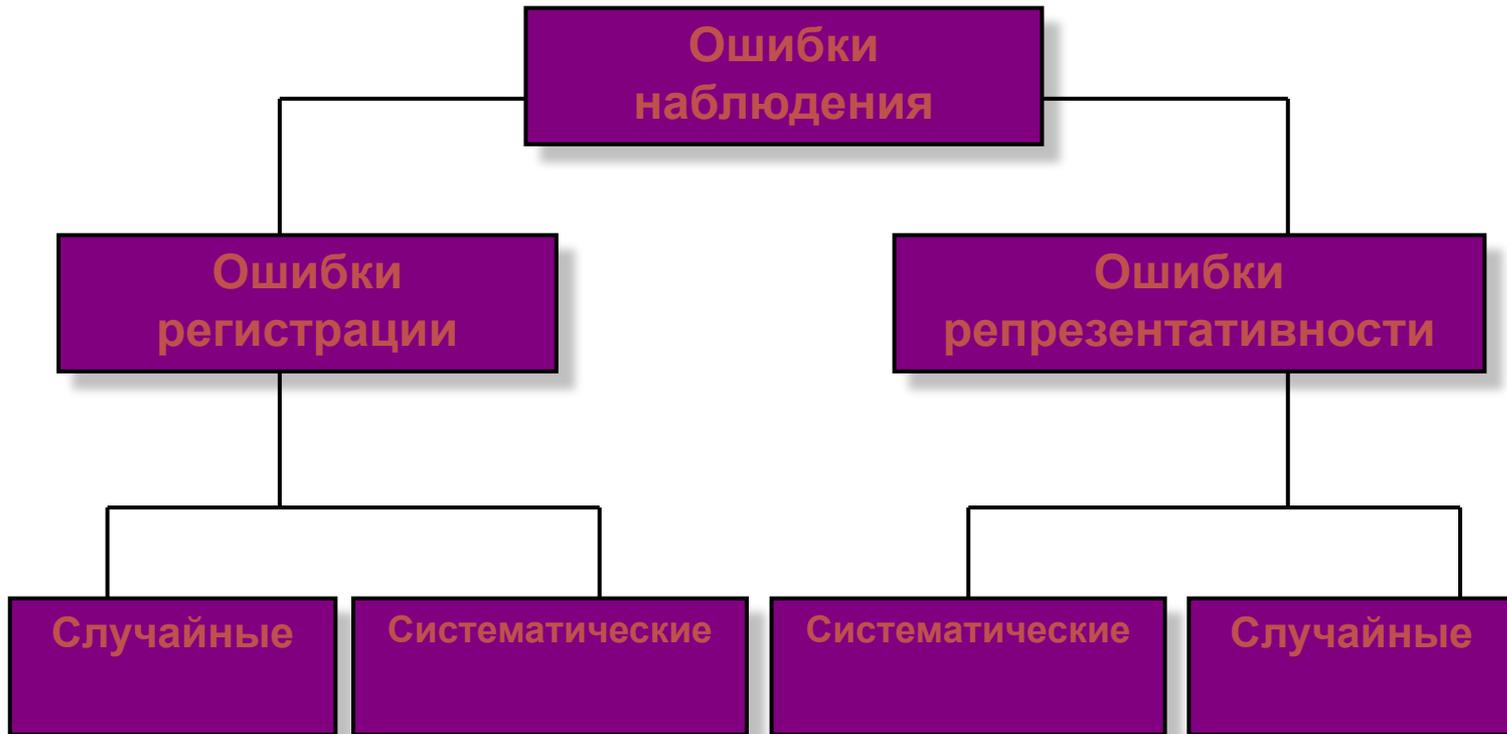


Выборочное наблюдение

- Под **выборочным наблюдением** понимается такое несплошное наблюдение, при котором отбор единиц в изучаемую совокупность производится случайным образом.
- Качество результата выборочного наблюдения определяется **системой организации** отбора единиц в выборочную совокупность.



Точность наблюдения



Модуль 3. Статистическая сводка и группировка

- 3.1. Содержание, задачи и виды сводки
- 3.2. Сущность, задачи и этапы проведения статистической группировки
- 3.3. Принципы построения статистической таблицы
- 3.4. Виды статистических группировок
- 3.5. Определение числа групп и границ интервалов
- 3.6. Сопоставимость статистических группировок.
Вторичная группировка
- 3.7. Основные классификации, группировки и номенклатуры в экономической статистике
- 3.8. Многомерные статистические группировки

3.1. Содержание, задачи и виды сводки

Сводка – это комплекс последовательных операций по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, присущих изучаемому явлению в целом.

3.1. Содержание, задачи и виды сводки

Основные этапы сводки:

- 1) Определение задач сводки;
- 2) Разработка системы статистических показателей;
- 3) Определение видов таблиц;
- 4) Систематизация и группировка полученных данных;
- 5) Расчет обобщающих показателей по всей совокупности и по группам.

3.1. Содержание, задачи и виды сводки

Виды статистической сводки:



3.2. Сущность, задачи и этапы проведения статистической группировки

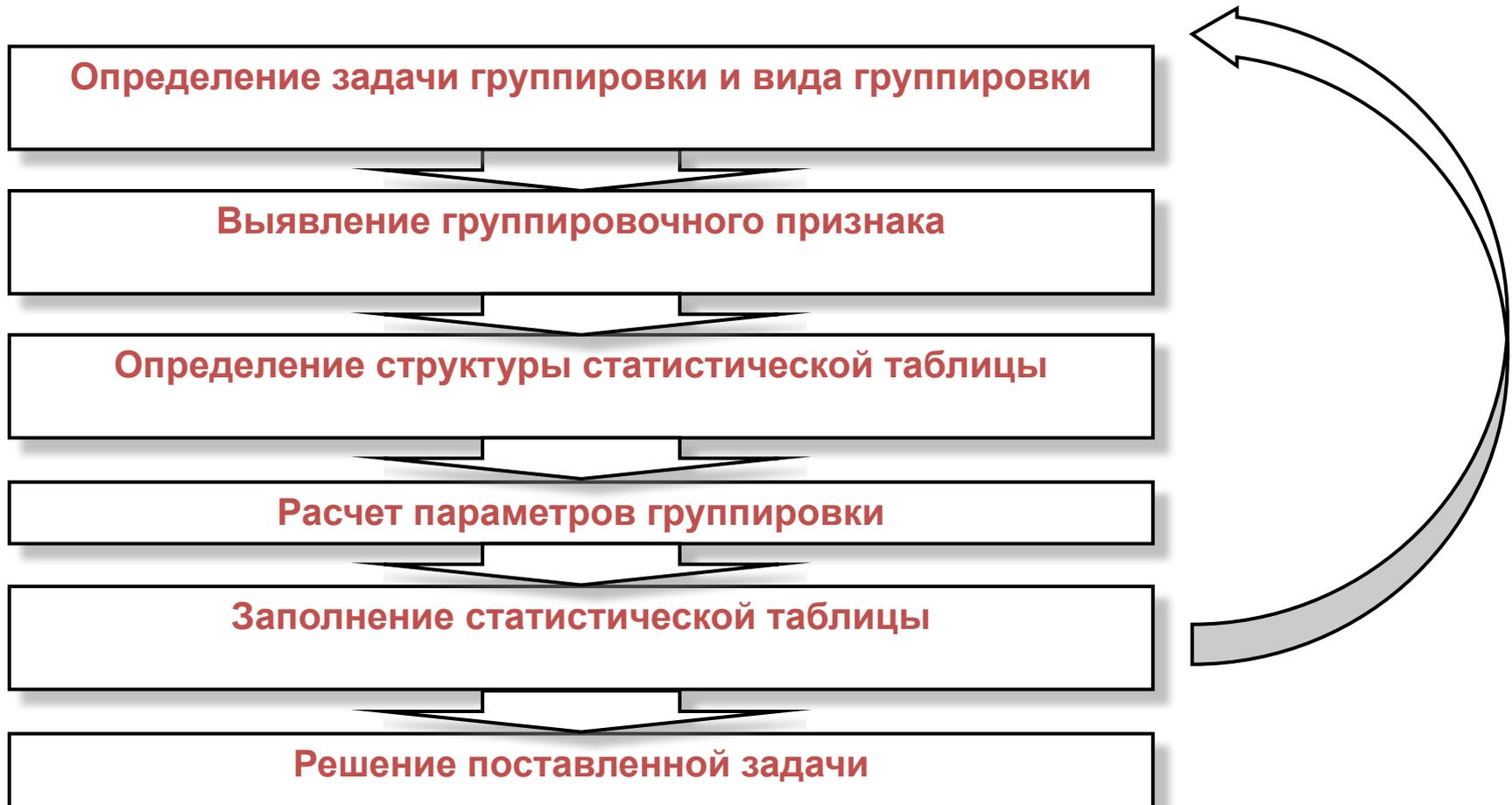
Группировка – это разделение множества единиц изучаемой совокупности на группы по определенным существенным для них признакам.

3.2. Сущность, задачи и этапы проведения статистической группировки

В отличие от группировки, классификация – систематизированное распределение явлений и объектов на определенные группы, классы, разряды на основании их сходства и различия.

3.2. Сущность, задачи и этапы проведения статистической группировки

Алгоритм выполнения группировки



3.2. Сущность, задачи и этапы проведения статистической группировки

Задачи, решаемые группировкой:

- Выделение социально-экономических типов явлений;
- Изучение структуры явления и структурных сдвигов, происходящих в нем;
- Выявление взаимосвязи и зависимости между явлениями.

3.2. Сущность, задачи и этапы проведения статистической группировки

Группировочный признак (основание группировки) – признак, по которому проводится разбивка единиц совокупности на отдельные группы.

3.3. Принципы построения статистической таблицы

Статистическая таблица – это таблица, содержащая сводную числовую характеристику исследуемой совокупности по одному или нескольким существенным признакам, взаимосвязанным логикой экономического анализа.

3.3. Принципы построения статистической таблицы

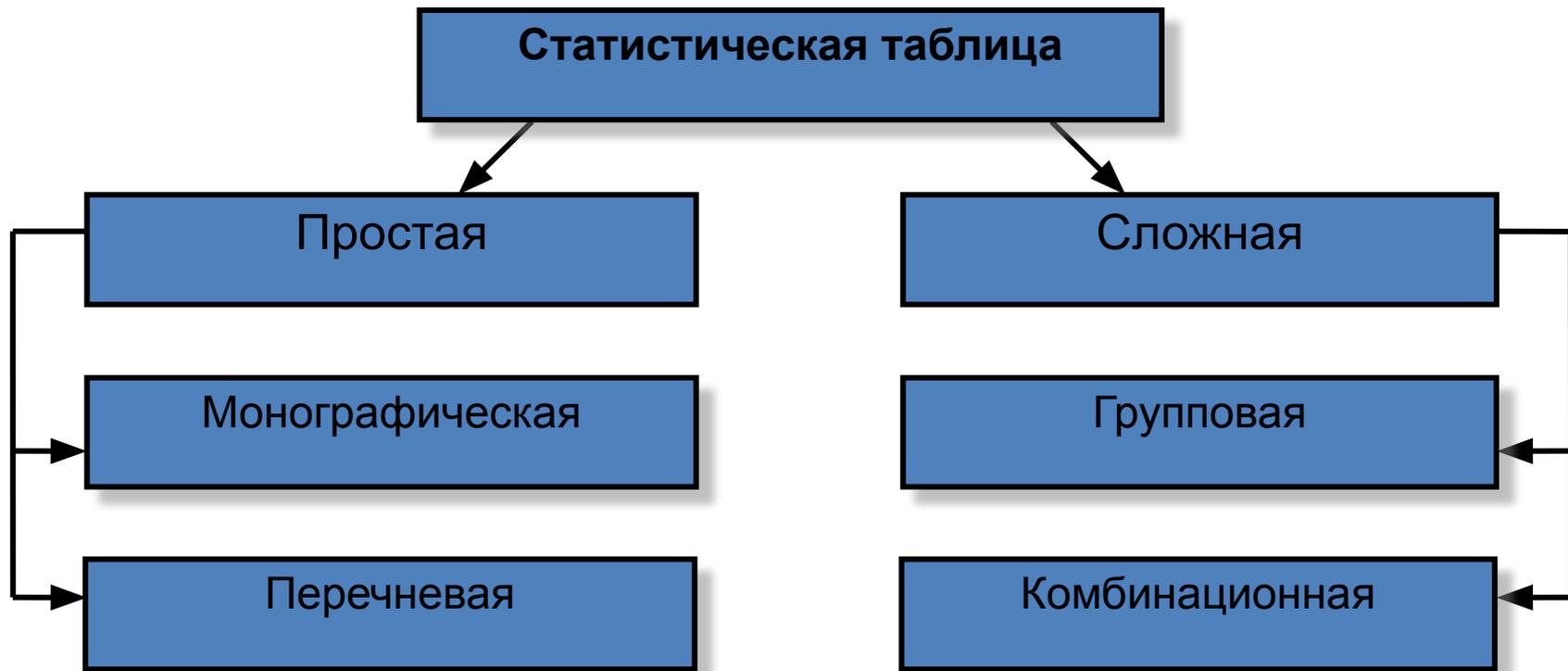
Общий вид статистической таблицы

Название таблицы (общий заголовок)*

| Содержание строк | Наименование граф (верхние заголовки) | | | | | |
|---|---------------------------------------|---|---|---|---|----------------|
| А | 1 | 2 | 3 | 4 | 5 | ... |
| Наименование строк (боковые заголовки) | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| Итоговая строка * Примечание таблице. | | к | | | | Итоговая графа |

3.3. Принципы построения статистической таблицы

Виды статистических таблиц по структуре подлежащего

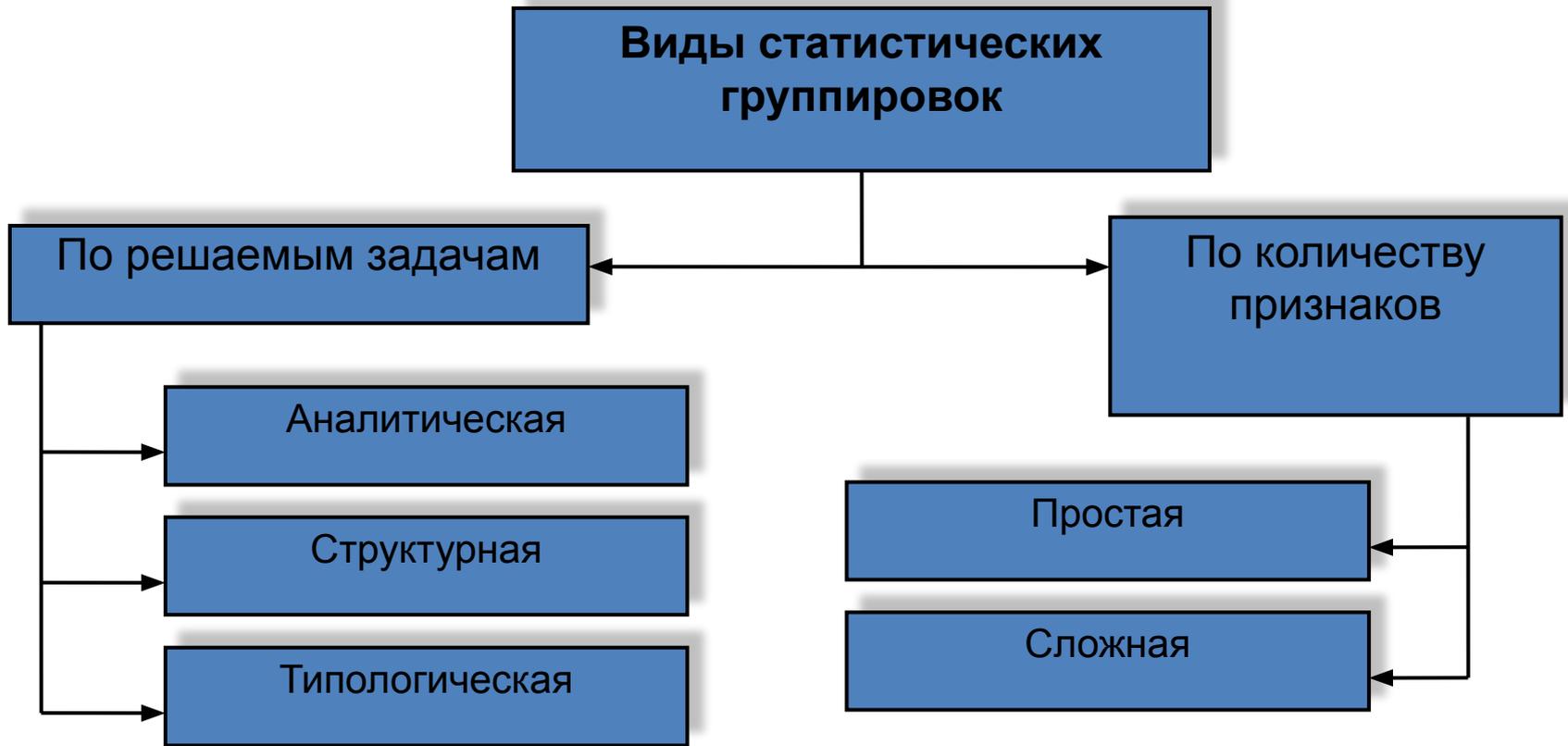


3.3. Принципы построения статистической таблицы

Простая разработка сказуемого

| № _{п/п} | Предприятие | Принято, чел. | Уволено, чел | Среднесписочное число работников-мужчин, чел. | Среднесписочное число работников-женщин, чел. |
|------------------|-------------|---------------|--------------|---|---|
| 1. | Альфа | 10 | 7 | 25 | 10 |
| 2. | Гамма | 5 | 7 | 20 | 22 |
| ... | ... | ... | ... | ... | ... |

3.4. Виды статистических группировок



3.4. Виды статистических группировок

Образец типологической группировки

Распределение федеральных округов по числу регионов

| № п/п | Федеральный округ | Число регионов, ед. | Процент к итогу |
|-------|--------------------|---------------------|-----------------|
| 1 | Дальневосточный ФО | 10 | 11,2 |
| 2 | Приволжский ФО | 15 | 16,9 |
| 3 | Северо-Западный ФО | 11 | 12,4 |
| 4 | Сибирский ФО | 16 | 18,0 |
| 5 | Уральский ФО | 6 | 6,7 |
| 6 | Центральный ФО | 18 | 20,2 |
| 7 | Южный ФО | 13 | 14,6 |
| | Итого | 89 | 100,0 |

3.4. Виды статистических группировок

Образец структурной группировки

Распределение регионов по доходам бюджетов

| № п/п | Группа по доходам бюджета (млрд. руб.) | Число регионов в группе | Процент к итогу |
|-------|--|-------------------------|-----------------|
| 1 | Менее 0,9 | 5 | 5,62 |
| 2 | От 0,9 до 2,2 | 11 | 12,36 |
| 3 | От 2,2 до 5,4 | 30 | 33,71 |
| 4 | От 5,4 до 13,3 | 26 | 29,21 |
| 5 | От 13,3 до 33,0 | 11 | 12,36 |
| 6 | От 33,0 до 80,0 | 4 | 4,49 |
| 7 | 80,0 и более | 2 | 2,25 |
| | Итого | 89 | 100,00 |

3.4. Виды статистических группировок

Образец аналитической группировки

Аналитическая группировка регионов
по Федеральным округам и средним доходам бюджета

| №п/ п | Федеральный округ | Число регионов | Сумма доходов бюджета по ФО, млрд. руб. | Средние по ФО доходы бюджета, млрд. руб. |
|----------|--------------------|-------------------|---|--|
| 1 | Уральский ФО | 6 | 177,8 | 29,6 |
| 2 | Центральный ФО | 18 | 310,0 | 17,2 |
| 3 | Приволжский ФО | 15 | 204,2 | 13,6 |
| 4 | Северо-Западный ФО | 11 | 101,0 | 9,2 |
| 5 | Сибирский ФО | 16 | 107,4 | 6,7 |
| 6 | Дальневосточный ФО | 10 | 58,7 | 5,9 |
| 7 | Южный ФО | 13 | 72,4 | 5,6 |
| | Итого | 89 | 1031,5 | 22,9 |

3.5. Определение числа групп и границ интервалов

Факторы, влияющие на число групп:

- Задачи исследования;
- Вид группировочного признака;
- Численность совокупности;
- Степень вариации признака в совокупности.

3.5. Определение числа групп и границ интервалов

Определение числа групп: формула Стерджесса

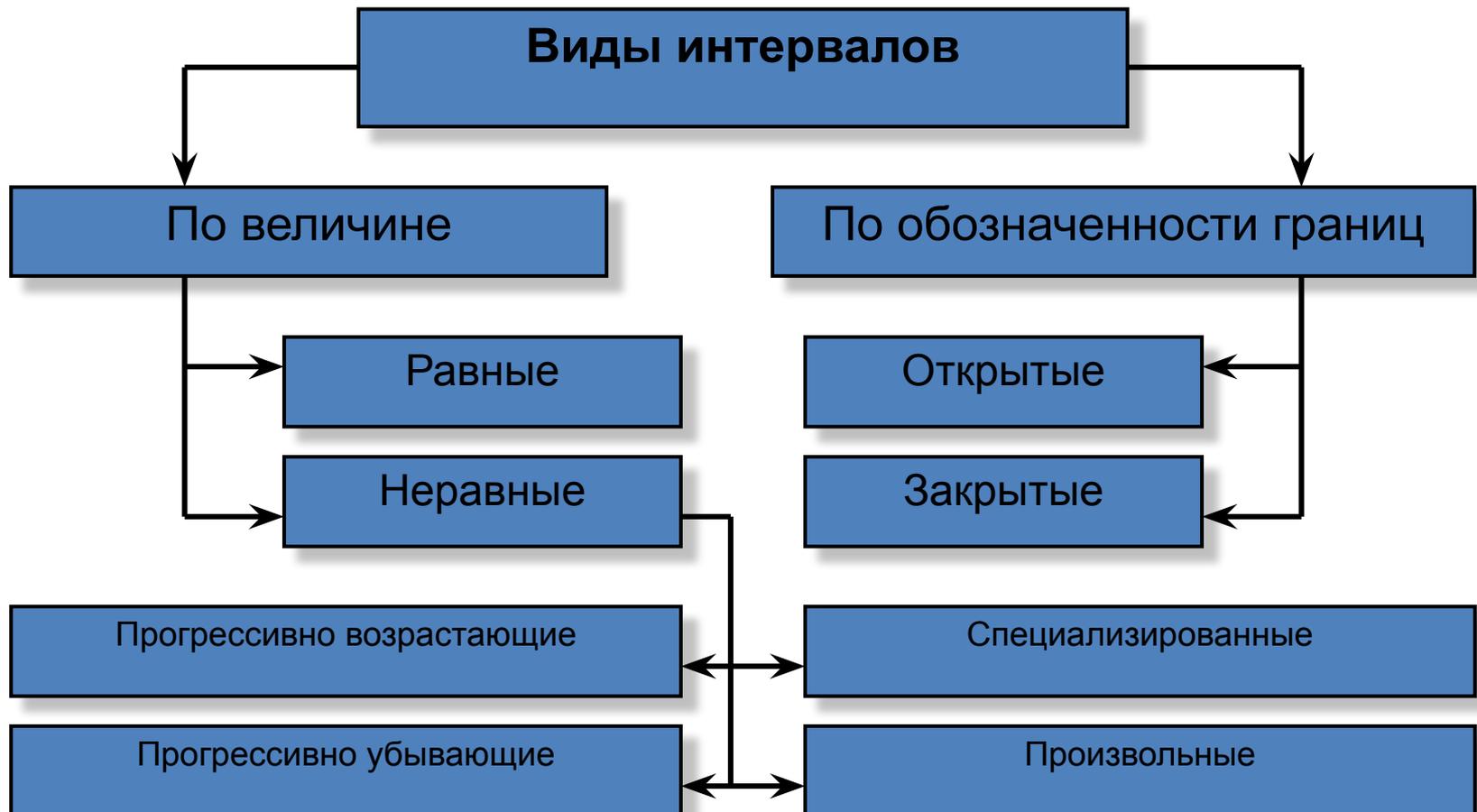
$$m = 1 + 3,322 \lg N,$$

где m – число групп;

N – число единиц совокупности.

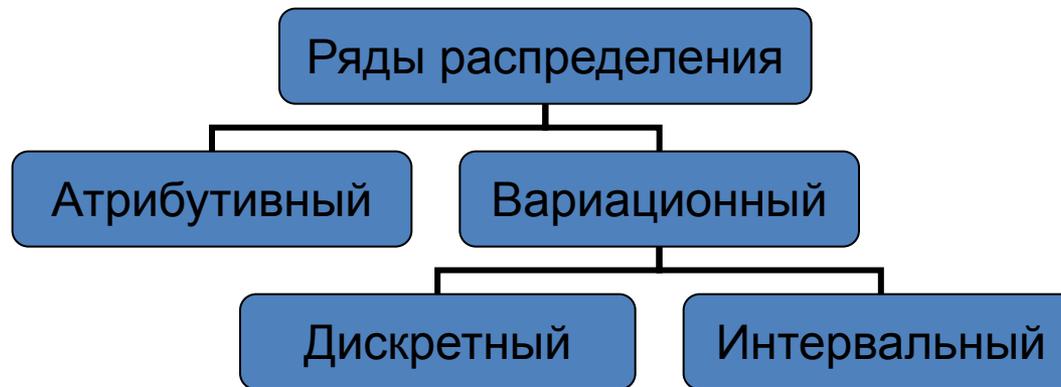
3.5. Определение числа групп и границ интервалов

Виды интервалов



Модуль 4. Модуль 4. Сущность, виды и графическое изображение ряда распределения

Виды рядов распределения



Сущность, виды и графическое изображение ряда распределения

- **Атрибутивные ряды** распределения строятся на основе атрибутивного признака и достаточно часто используются в экономике. Они получаются в результате проведения типологической группировки и характеризуют распределение разнородной совокупности.
- Ряды распределения, построенные по количественному признаку, называют **вариационными** рядами.
- Соответственно, в зависимости от того, какой признак лежит в основе ряда распределения: дискретный или непрерывный, различают **дискретные** и **интервальные** вариационные ряды.

Сущность, виды и графическое изображение ряда распределения

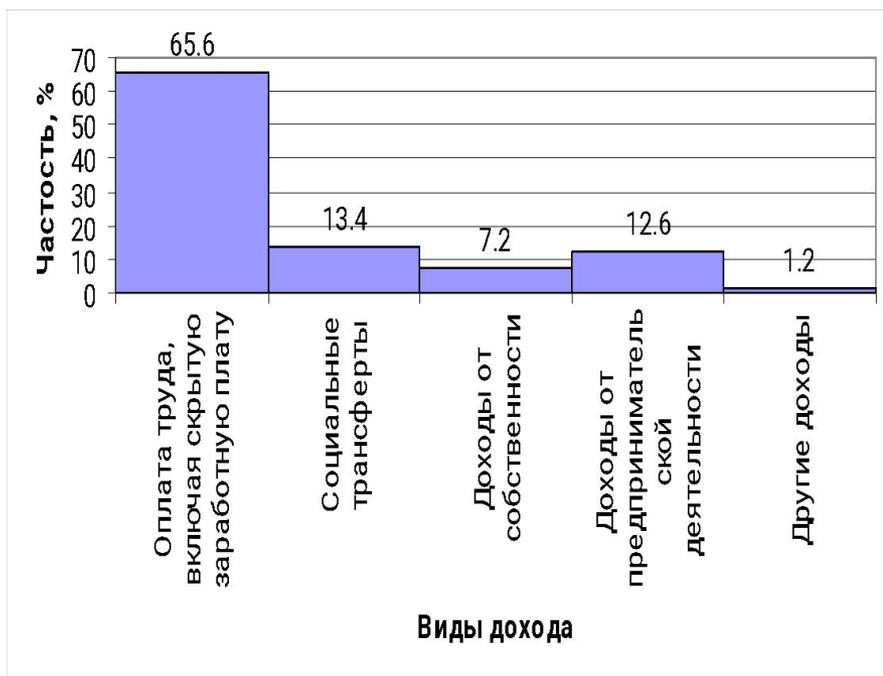
Элементы вариационного ряда:

1. **Варианты** – это отдельные значения (только для дискретного ряда) или отдельные интервалы значений варьирующего признака, которые он принимает в вариационном ряду.
2. **Частоты** – это численности отдельных вариантов. Также при построении вариационного ряда могут использоваться **частости** – частоты, выраженные в долях единицы или в процентах.
Важное место в анализе рядов распределения занимают **накопленные частоты** – суммарная частота всех групп от первой до текущей.

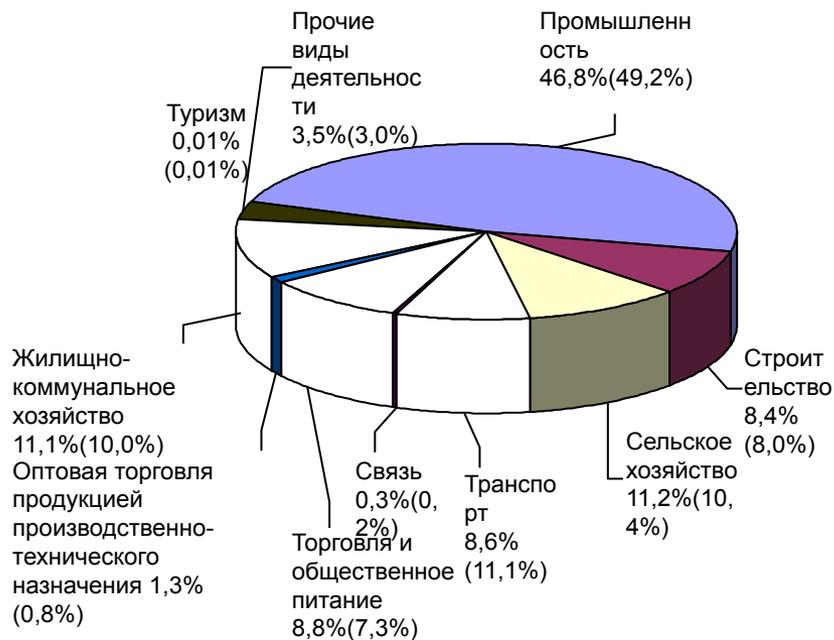
Графическое изображение рядов распределения

- Для изображения атрибутивного ряда могут использоваться гистограмма и круговая диаграмма.

Гистограмма:



Круговая диаграмма:

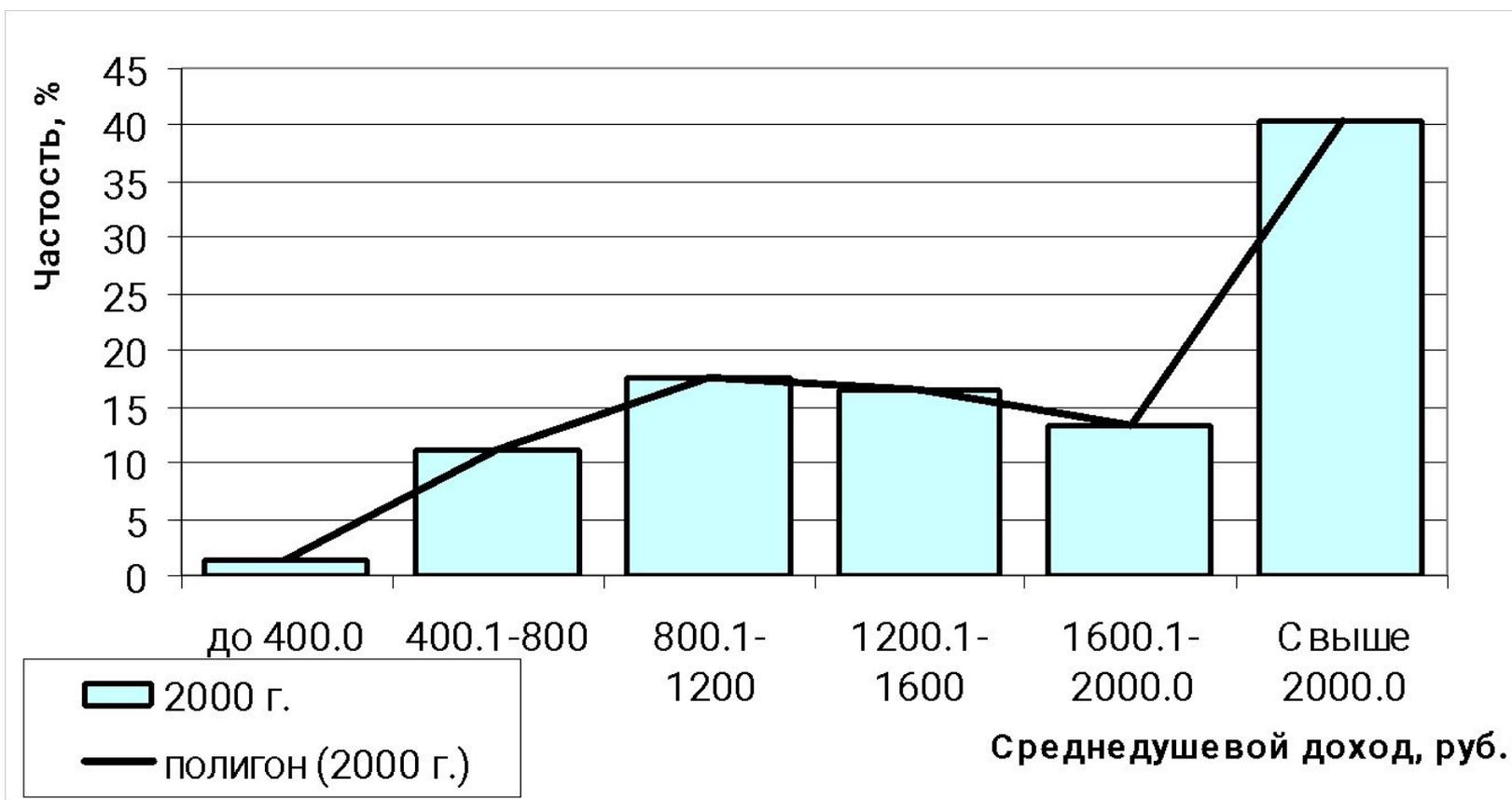


Графическое изображение рядов распределения

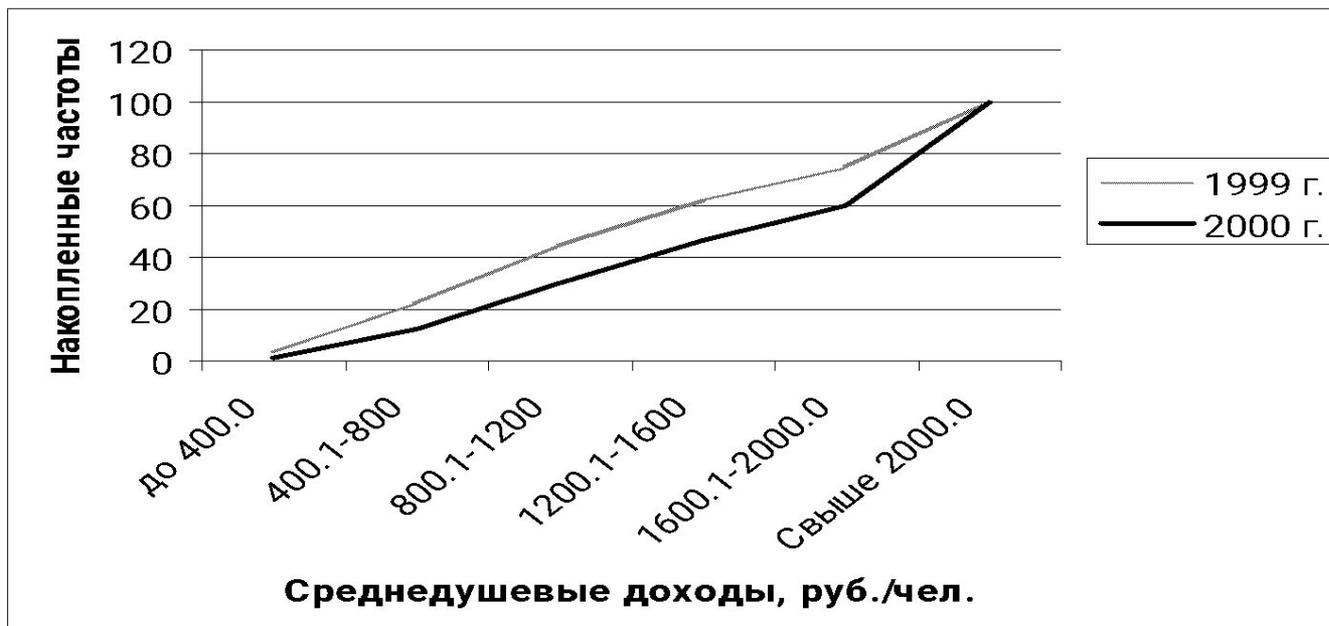
Для изображения вариационных рядов используются **полигон** и **гистограмма**.

- **Полигон** преимущественно используется при изображении дискретных вариационных рядов.
- **Гистограмма** применяется для изображения интервального вариационного ряда.

Гистограмма и полигон частот



Помимо полигона и гистограммы для графического изображения вариационных рядов может быть использована кумулятивная кривая (кумулята). **Кумулята** – это кривая, графически отображающая ряд накопленных частот.



Модуль 5. Статистические показатели

5.1. Сущность и виды статистических показателей

5.2. Абсолютные статистические показатели

5.3. Относительные статистические показатели

5.4. Средние величины

5.5. Структурные средние величины и другие
структурные показатели

5.6. Показатели вариации

5.7. Оценка дисперсии сгруппированных данных

5.8. Графическое представление сгруппированных
данных

5.9. Оценка статистических показателей

5.1. Сущность и виды статистических показателей

Статистический показатель – это количественная характеристика социально-экономических явлений и процессов в условиях качественной определенности.

5.1. Сущность и виды статистических показателей

Виды статистических показателей:

- по пространственной определенности:
 - общетерриториальные
 - региональные
 - локальные
- по форме выражения:
 - абсолютные
 - относительные
 - средние
- по временному фактору:
 - моментные
 - интервальные

5.1. Сущность и виды статистических показателей

- по принадлежности к объекту:
 - однообъектные
 - межобъектные
- по охвату единиц совокупности:
 - индивидуальные
 - сводные (расчетные и объемные).

5.2. Абсолютные статистические показатели

- Абсолютные величины представляют собой характеристику всего исследуемого явления по отдельно взятому признаку.
- Абсолютные величины являются результатом первичного учета, заключающегося в первоначальной регистрации предметов, событий хозяйственной деятельности, отражаемой в соответствующей документации.

5.2. Абсолютные статистические показатели

По степени охвата исследуемой совокупности выделяют несколько видов абсолютных величин:

- Индивидуальные (характеризуют отдельные единицы совокупности)
- Групповые (отражают размеры признака в отдельных частях совокупности)
- Общие (отражают размеры признака в совокупности в целом)

5.2. Абсолютные статистические показатели

Единицы измерения абсолютных показателей

- **натуральные единицы измерения** (тонны, метры, литры, штуки и т.п.)
- **стоимостные единицы измерения** (объем производства на предприятии, валовой национальный продукт)
- **Трудовые единицы измерения** (человеко-дни, человеко-часы)

5.3. Относительные статистические показатели

- **Относительные статистические показатели** – это показатели, отражающие соотношение между размерами изучаемого явления.
- Метод их расчета можно сформулировать следующим образом: соотношение сравниваемого показателя с другим показателем, принятым за **базу сравнения**.
- Показатель, с которым сравнивается изучаемый признак, называется **базисным**.

5.3. Относительные статистические показатели

Виды относительных статистических показателей:

- Динамики (характеризует отношение уровня исследуемого явления за данный период времени к уровню этого же явления в прошлом)
- Плана (рассчитывается как соотношение планируемого уровня явления с текущим уровнем этого же явления)
- Реализации плана (соотношение фактически достигнутого уровня явления в исследуемом периоде с планируемым уровнем этого явления; показывает, во сколько раз фактический уровень исследуемого явления в отчетном периоде отличается от запланированного уровня явления)

5.3. Относительные статистические показатели

Виды относительных статистических показателей:

- Структуры (показывают, какую долю составляет часть совокупности в общем ее объеме)
- Координации (показывают соотношение отдельных частей целого между собой (в качестве базы сравнения используется часть, имеющая наибольший удельный вес)
- Интенсивности (характеризуют степень распространения изучаемого процесса или явления в присущей ему среде)
- Сравнения (соотношение одноименных абсолютных показателей, характеризующих разные объекты)

5.4. Средние величины

- **Средняя величина** является основной обобщающей количественной характеристикой признака в статистической совокупности в конкретных условиях места и времени.

- **Определения средней величины:**

- 1) Наиболее типичное для совокупности значение признака;
- 2) Объем признака совокупности, распределенный поровну между единицами совокупности.

5.4. Средние величины

Виды средних:

- Средняя арифметическая

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \qquad \bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$$

- Средняя квадратическая лежит в основе вычислений ряда сводных расчетных показателей, в частности – показателей вариации

5.4. Средние величины

- **Средняя геометрическая** используется при анализе динамики для определения среднего темпа роста

$$\bar{x} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} = \sqrt[n]{\prod x_i}$$

- **Средняя гармоническая** используется для расчетов средних в рядах относительных показателей

$$\bar{x} = \frac{n}{\sum \frac{1}{x_i}}$$

5.4. Средние величины

Средние взвешенные величины

- средняя арифметическая взвешенная

$$\bar{x}_{\text{âç}} = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i} = \sum_{i=1}^n x_i d_i$$

Где $d_i = \frac{w_i}{\sum_{i=1}^n w_i}$ - относительный показатель структуры

Другие виды средних взвешенных рассчитываются аналогично.

5.5. Структурные средние величины и другие структурные показатели

Мода – это значение изучаемого признака, повторяющееся с наибольшей частотой.

Выделяют два способа расчета моды:

- Для дискретных рядов распределения мода определяется путем выбора варианта, имеющего наибольшую частоту;

5.5. Структурные средние величины и другие структурные показатели

- Для интервальных рядов распределения расчет моды производится по следующему алгоритму:
 - 1 шаг: Определение модального интервала – , такого, что его частота больше и чем у предыдущего, и чем у следующего;
 - 2 шаг: Расчет моды по формуле:

$$M_0 = x_0 + h \cdot \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})}$$

где x_0 – нижняя граница модального интервала;

h – величина модального интервала;

f_m – частота модального интервала;

f_{m-1} – частота интервала, предшествующего модальному;

f_{m+1} – частота интервала, следующего за модальным.

5.5. Структурные средние величины и другие структурные показатели

Медиана – это значение признака, приходящееся на середину ранжированной совокупности.

При расчете медианы можно выделить два подхода:

- Для несгруппированных данных медианой будет считаться значение признака, лежащего посередине упорядоченной совокупности;
- Для сгруппированных данных расчет осуществляется в три этапа:

1 шаг: Рассчитываются накопленные частоты

2 шаг: Определяется медианный интервал: такой, при котором сумма накопленных частот превысит половину общей численности совокупности;

5.5. Структурные средние величины и другие структурные показатели

3 шаг: Определение медианы по формуле:

$$M_e = x_0 + h \cdot \frac{\frac{1}{2} \sum f_i - S_{me-1}}{f_{me}}$$

где x_0 – нижняя граница медианного интервала;

h – величина медианного интервала;

S_{me-1} – сумма накопленных частот на интервале, предшествующем медианному;

f_{me} – частота медианного интервала.

| З/п, тыс.руб | Число работников | S |
|---------------------|-------------------------|----------|
| Менее 5 | 2 | 2 |
| 5-10 | 5 | 7 |
| 10-15 | 10 | 17 |
| 15-20 | 17 | 34 |
| 20-25 | 7 | 41 |
| Более 25 | 3 | 44 |

5.5. Структурные средние величины и другие структурные показатели

Квартили – это значения признака, делящие ранжированную совокупность на четыре равновеликие части.

- Выделяют нижний (Q1) и верхний (Q3) квартили. Нижний квартиль отделяет четвертую часть совокупности с наименьшими значениями признака, верхний – с наибольшими.

$$Q_1 = x_{Q_1} + h \frac{\frac{1}{4} \sum f - S_{Q_1-1}}{f_{Q_1}} \quad Q_3 = x_{Q_3} + h \frac{\frac{3}{4} \sum f - S_{Q_3-1}}{f_{Q_3}}$$

5.5. Структурные средние величины и другие структурные показатели

Квартили:

где x_{Q1} – нижняя граница интервала, содержащего нижний квартиль (интервал определяется по накопленной частоте, первой превышающей 25 %);

x_{Q3} – нижняя граница интервала, содержащего верхний квартиль (интервал определяется по накопленной частоте, первой превышающей 75 %);

h – величина интервала;

S_{Q_1-1} – накопленная частота интервала, предшествующего интервалу, содержащему нижний квартиль;

S_{Q_3-1} – накопленная частота интервала, предшествующего интервалу, содержащему верхний квартиль;

f_{Q1} – частота интервала, содержащего нижний квартиль;

f_{Q3} – частота интервала, содержащего верхний квартиль.

5.5. Структурные средние величины и другие структурные показатели

Децили – это варианты, которые делят ранжированный ряд на десять равных частей.

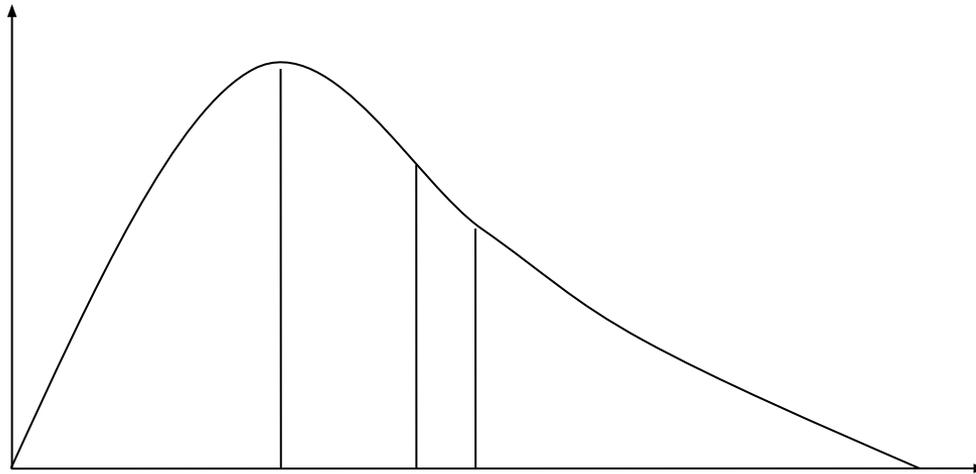
Процентили – это варианты, делящие ряд на сто частей. Они применяются лишь при необходимости подробного изучения структуры вариационного ряда.

5.5. Структурные средние величины и другие структурные показатели

Асимметрия

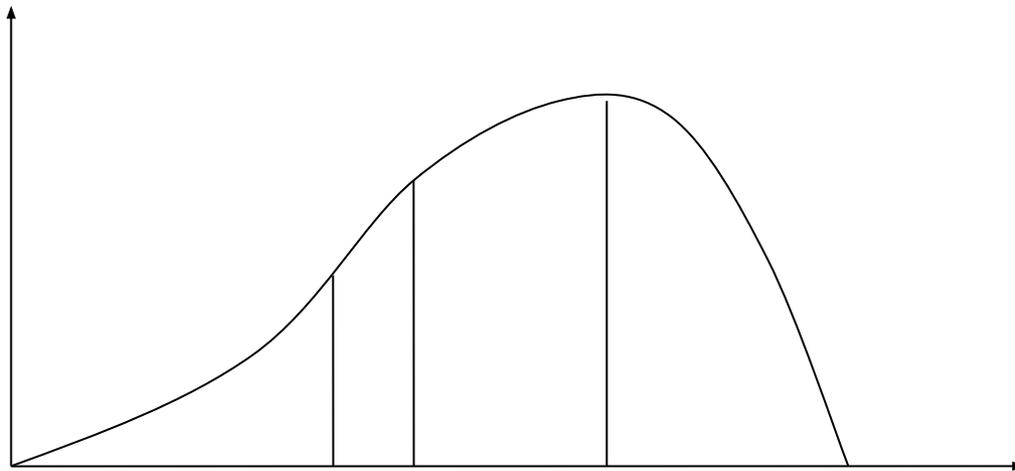
Асимметрия имеет место, когда наблюдается неравномерность распределения единиц вокруг средней величины.

В случае положительной асимметрии большая часть совокупности имеет значение ниже средней, однако значения выше средней имеют единицы с более значительным разбросом.



5.5. Структурные средние величины и другие структурные показатели

В случае **отрицательной асимметрии** ситуация обратная – большинство единиц совокупности располагаются выше средней, в то время как отдельные единицы обладают значительно меньшими, по сравнению со средней, значениями признака.



5.6. Показатели вариации

- **Размах вариации** – это разность между единицами совокупности с наибольшим и наименьшим значениями варьирующего признака.

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

где x_{\max} и x_{\min} – наибольшее и наименьшее значения варьирующего признака.

- **Размах квартилей** – разность между значениями третьего и первого квартилей:

$$Q = Q_3 - Q_1$$

5.6. Показатели вариации

- **Квартильное отклонение.** (приблизненно показывает, насколько отклоняются от среднего значения 50% ближайших к нему значений признака)

$$q = \frac{Q}{2} = \frac{(Q_3 - Q_1)}{2}$$

- **Среднее линейное отклонение.** (определяется как средняя арифметическая из абсолютных значений отклонений вариант признака от их средней)

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

5.6. Показатели вариации

- **Дисперсия**

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

- **Среднее квадратическое отклонение** (является обобщающей характеристикой размеров вариации признака в совокупности)

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

5.6. Показатели вариации

Относительные показатели вариации

- Коэффициент осцилляции

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

- Линейный коэффициент вариации

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

- Коэффициент вариации (наиболее часто используется)

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

5.7. Оценка дисперсии сгруппированных данных

Межгрупповая дисперсия характеризует вариацию признака в совокупности, обусловленную влиянием на анализируемый признак группировочного

$$\delta^2 = \frac{\sum_{j=1}^m (\bar{x}_j - \bar{x})^2 \cdot n_j}{n}$$

где \bar{x}_j – среднее значение признака в j -й группе,

\bar{x} – среднее значение признака во всей совокупности

n_j – размер j -й группы,

m – число групп

5.7. Оценка дисперсии сгруппированных данных

Средняя из внутригрупповых дисперсий показывает вариацию анализируемого признака, обусловленную влиянием признаков, не вошедших в основание группировки

$$\overline{\sigma^2} = \frac{\sum_{j=1}^m \sigma_j^2 \cdot n_j}{n}$$

Общая вариация признака в совокупности описывается **общей дисперсией**, которая может быть рассчитана как сумма межгрупповой и средней из внутригрупповых дисперсий

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

5.9. Оценка статистических показателей

- Большинство величин, с которыми имеет дело исследователь в статистике, являются не жестко детерминированными, а вероятностными величинами. Мы можем лишь оценить их значения, предположить с определенной вероятностью границы, в которых они могут находиться на самом деле.
- Существует два типа оценок – точечные и интервальные.

5.9. Оценка статистических показателей

- **Точечная оценка** представляет собой отдельное число (точку), которое используется в качестве оценки параметра генеральной совокупности.
- **Интервальная оценка** представляет собой интервал, в котором с известной вероятностью находится истинное значение исследуемого признака. Такой интервал называется доверительным, а соответствующая ему вероятность – доверительной вероятностью. Доверительные интервалы определяются на основе площади под кривой нормального распределения

$$\mu = \bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

где $Z_{\alpha/2}$ – значения, определяемые из таблицы площади под кривой нормального распределения

Выборочное наблюдение

Это несплошное наблюдение, при котором признаки регистрируются у отдельных единиц совокупности, отобранных с использованием специальных методов, а полученные в процессе обследования результаты с определенным уровнем вероятности распространяются на всю исходную совокупность

Генеральная совокупность – основа выборки, исходная совокупность

- Повторный отбор – попавшая в выборку единица подвергается обследованию и возвращается в генеральную совокупность.
- Число единиц, участвующих в отборе, остается постоянным, обычно используется, когда объем генеральной совокупности неизвестен
- Бесповторные отбор – попавшая в выборку единица в дальнейшем отборе

Ошибки репрезентативности

- Обусловлены тем, что выборочная совокупность не точно воспроизводит генеральную
- Систематические ошибки связаны с нарушением принципов формирования выборочной совокупности
- Случайные ошибки обусловлены действием случайных факторов

Этапы проведения выборочного наблюдения

- Определение цели обследования
- Установление границ генеральной совокупности
- Составление программы наблюдения
- Определение вида выборки, процента и метода отбора
- Отбор и регистрация признаков
- Расчет выборочных характеристик и их ошибок
- Распространение полученных результатов на ген

Модуль 7. Статистический анализ взаимосвязей

7.1. Сущность и виды взаимосвязей в статистике

7.2. Методы выявления наличия связи

7.3. Корреляционный анализ взаимосвязей количественных признаков

7.4. Корреляционный анализ признаков, выраженных в шкале порядка (ранговая корреляция)

7.5. Основы регрессионного анализа статистической связи

7.1. Сущность и виды взаимосвязей в статистике

Основные виды связей, изучаемых в статистике:

- **Балансовая.** Такой вид связи характерен для взаимообусловленных явлений и процессов (н-р, совокупные доходы всех субъектов национальной экономики равны их совокупным расходам)
- **Компонентная.** Такая связь характерна для показателей, один из которых можно определить как произведение остальных (н-р, взаимосвязь между составляющими основного уравнения количественной теории денег)
- **Факторная.** Этот вид связи подразумевает согласованную вариацию исследуемых статистических признаков. При изучении факторной связи выделяют **результативный** (тот, на который оказывают влияние) и **факторные** (оказывающие влияние) признаки.

7.1. Сущность и виды взаимосвязей в статистике

Факторная связь может быть:

- **функциональной (детерминированной)** (вариация результативного признака всецело обусловлена вариацией факторных)
- **статистической (стохастической, вероятностной)** (факторные признаки могут только частично объяснить вариацию результативного).

Наиболее часто стохастическая связь можно обнаружить в форме **корреляционной связи**, под которой понимают зависимость между определенными значениями факторных признаков и средним значением результативного.

7.1. Сущность и виды взаимосвязей в статистике

Виды корреляционной связи:

- в зависимости от количества факторных признаков:
 - парная
 - множественная
- по направлению:
 - прямая
 - обратная
- по силе:
 - слабая
 - умеренная
 - средней силы
 - сильная
- по характеру аналитической зависимости:
 - линейная
 - нелинейная

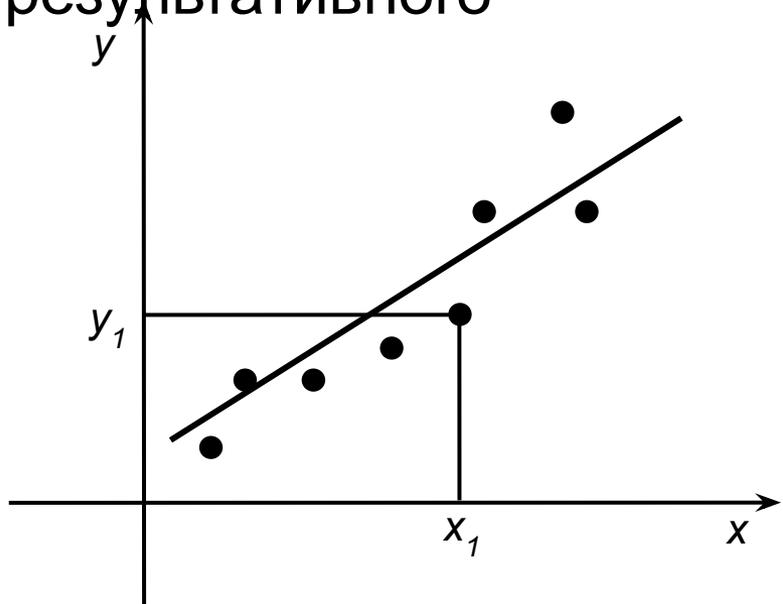
7.1. Сущность и виды взаимосвязей в статистике

Этапы статистического анализа корреляционной связи:

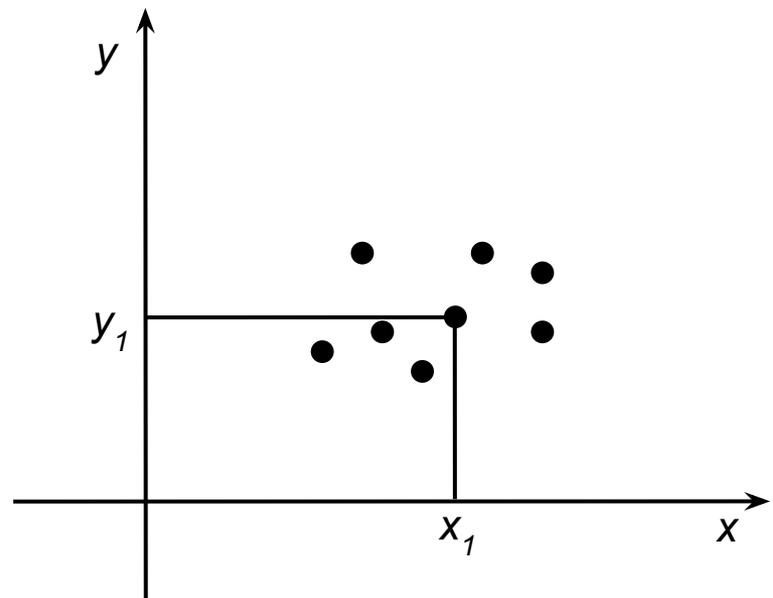


7.2. Методы выявления наличия связи

- **Графический метод.** Метод основан на построении и анализе *корреляционного поля*. Корреляционное поле строится в прямоугольной системе координат, в которой по оси абсцисс откладываются значения факторного признака, а по оси ординат – значения результативного



наличие связи (линейной)



отсутствие связи

7.2. Методы выявления наличия связи

- **Метод приведения параллельных данных** Этот метод основан на построении упорядоченного по факторному признаку списка единиц в совокупности. Если в этом списке наблюдается устойчивая тенденция изменения и результативного признака, то делается вывод о существовании связи.
- Расчет **коэффициента Фехнера** (коэффициента корреляции знаков). Этот метод основан на анализе поведения отклонений индивидуальных значений признака от среднего по факторному и результативному признакам.
- **Метод построения и анализа таблиц взаимной сопряженности.** Этот метод основан на построении комбинационных таблиц, в которых приводится распределение совокупности по двум признакам. Если таблица взаимной сопряженности строится по количественным признакам, то она называется **корреляционной таблицей**.
- **Дисперсионный анализ.** Дисперсионный анализ основан на проведении аналитической группировки, расчете и сопоставлении общей и межгрупповой дисперсии для результативного признака.

7.3. Корреляционный анализ взаимосвязи количественных признаков

- **Корреляционный анализ парной связи.** Для корреляционного анализа связи между двумя признаками используется *линейный коэффициент корреляции*:

$$r = \frac{\sum_{i=1} (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

После расчета линейного коэффициента корреляции требуется оценить его статистическую значимость при помощи t -статистики Стьюдента

$$t = \begin{cases} \frac{r}{\sqrt{1-r^2}} \sqrt{n}, & n > 30 \\ \frac{r}{\sqrt{1-r^2}} \sqrt{n-2}, & n \leq 30 \end{cases}$$

7.3. Корреляционный анализ взаимосвязи количественных признаков

- **Корреляционный анализ множественной связи**

- Анализ *частных коэффициентов корреляции*. Эти коэффициенты показывают силу и направление связи между несколькими признаками при условии независимости остальных

$$r_{y, x_1 / x_2, x_3, \dots, x_k} = \frac{r_{y, x_1 / x_2, x_3, \dots, x_{k-1}} - r_{y, x_k / x_2, x_3, \dots, x_{k-1}} \cdot r_{x_1, x_k / x_2, x_3, \dots, x_{k-1}}}{\sqrt{(1 - r_{y, x_k / x_2, x_3, \dots, x_{k-1}}^2) \cdot (1 - r_{x_1, x_k / x_2, x_3, \dots, x_{k-1}}^2)}}$$

- **Множественный коэффициент корреляции.** Позволяет установить связь между всеми признаками Для случая 2х факторных и 1 результативного:

$$R_{x_1, x_2, y} = \sqrt{\frac{r_{x_1, y}^2 - 2r_{x_1, y}r_{x_1, x_2}r_{y, x_2} + r_{y, x_2}^2}{1 - r_{x_1, y}^2}}$$

7.4. Корреляционный анализ признаков, выраженных в шкале порядка (ранговая корреляция)

Под **рангом** понимают порядковое место единицы в упорядоченной совокупности.

Для оценки парной ранговой корреляции используется два коэффициента:

- **Ранговый коэффициент Спирмена** (коэффициент линейной корреляции между рангами)

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

где d_i – ранговая разность, $d_i = R_{xi} - R_{yi}$

7.4. Корреляционный анализ признаков, выраженных в шкале порядка (ранговая корреляция)

- **Ранговый коэффициент Кенделла.** В основе его расчета лежит сопоставление числа «правильных» и «неправильных» рангов. Для этого используется следующая формула:

$$\tau = \frac{2S}{n(n-1)}$$

$$S = \sum_{i=1}^n P_i - \sum_{i=1}^n Q_i$$

- P_i – число единиц в совокупности, для которых *большему*, чем у i -й единицы рангу по x *соответствует больший*, чем у i -й единицы ранг по y .
- Q_i – число единиц в совокупности, для которых *меньшему*, чем у i -й единицы рангу по x *соответствует больший*, чем у i -й единицы ранг по y .

7.5. Основы регрессионного анализа статистической связи

- Регрессионный анализ состоит в отыскании аналитической формы зависимости между статистическими показателями. Наиболее простой является парная зависимость, описываемая парной регрессионной моделью

$$y = \alpha_0 + \alpha_1 x + \varepsilon \quad \text{- линейная парная регрессия}$$

(спецификация).

- Уравнение линейной парной регрессии: $\hat{y} = f(x) = a_0 + a_1 \cdot x$

где a_0 и a_1 – оценки теоретических коэффициентов регрессии α_0 и α_1 .

- Для определения коэффициентов чаще всего используют **метод наименьших квадратов**

7.5. Основы регрессионного анализа статистической связи

- Суть метода наименьших квадратов:

наилучшими считаются оценки параметров уравнения, обеспечивающие минимальное значение суммы квадратов случайных отклонений.

- Математическое условие минимизации квадратов:

$$F = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a_0 - a_1 \cdot x_i)^2 \rightarrow \min$$

- **Стандартная форма нормальных уравнений:**

решив это уравнение относительно

a_0 и a_1 ,

получим оценки параметров

$$\begin{cases} a_0 \cdot n + a_1 \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a_0 \cdot \sum_{i=1}^n x_i + a_1 \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i \cdot x_i \end{cases}$$

теоретического уравнения α_0 и α_1 .

- Индекс — это показатель сравнения двух состояний одного и того же явления (простого или сложного, состоящего из соизмеримых или несоизмеримых элементов).
- При помощи индексов можно решить следующие основные задачи:
 - определить средние изменения сложных, непосредственно несоизмеримых совокупностей во времени (индексы выступают в роли показателей динамики);
 - оценить среднюю степень выполнения плана по совокупности в целом или ее части (выступают в роли показателей выполнения плана);
 - установить средние отношения сложных явлений в пространстве (выступают в роли показателей сравнения);
 - определить роль отдельных факторов в общем изменении сложных явлений во времени или в пространстве и, в частности, изучается влияние структурных сдвигов (выступают в роли аналитического средства).

Федеральной таможенной службы от 18 декабря 2006 г. N 1329 «Об утверждении методологии исчисления системы индексов внешней торговли на основе данных таможенной статистики» и предназначена для описания и анализа краткосрочной и долгосрочной динамики показателей внешней торговли: стоимости, цен и физического объема.

- Система индексов внешней торговли Российской Федерации включает:
- а) индивидуальные индексы:
 - - стоимости $i_{pq} = q_1 p_1 / q_0 p_0$;
 - - цен $i_p = p_1 / p_0$;
 - - физического объема для сопоставимых на низшем иерархическом уровне ТН ВЭД России товаров (10-значная товарная подсубпозиция) $i_q = q_1 / q_0$.
- б) сводные индексы:
 - - стоимости $i_{pq} = \Sigma p_1 q_1 / \Sigma p_0 q_0$;
 - - физического объема по формулам Ласпейреса и Пааше;
 - - средних цен (удельной стоимости) по формуле Пааше.
- в) индекс условий торговли как отношение сводных индексов средних цен экспорта и импорта (рассчитывается только для внешней торговли Российской Федерации в целом).

- Изменение географического распределения экспорта товаров по странам может привести к значительным изменениям цен.
- Так же стоит отметить влияние ассортиментных сдвигов на стоимостные показатели внешней торговли, для чего применяется индекс ассортиментных сдвигов.
- В рамках факторного анализа существенно упомянуть так же и об изменениях физического объема, однако, данный фактор учитывается при формировании данных официальной таможенной статистики.
- Таким образом, система взаимосвязанных индексов, характеризующих влияние отдельных факторов на экспорт или импорт, будет иметь вид:
- $$I_{\text{стоимости}} = I_{\text{ср. цен без учета структ. сдвигов}} \times I_{\text{геогр. стр.}} \times I_{\text{ассорт. сдв.}} \times I_{\text{физического объема}}$$

Модуль 6. Анализ динамики

- Процесс развития, движения социально-экономических явлений во времени в статистике принято называть динамикой. Для отображения динамики строят **ряды динамики** (хронологические ряды, временные ряды), которые представляют собой ряды изменяющихся во времени значений статистического показателя, расположенных в хронологическом порядке.
- Составными элементами ряда динамики являются **показатели уровней ряда (уровни ряда)** и **периоды времени** (годы, кварталы, месяцы, сутки) или **моменты (даты) времени**.
- Уровни ряда обычно обозначаются через « u », моменты или периоды времени, к которым относятся уровни, - через « t ».

Виды рядов динамики

- 1. По **отношению ко времени** уровней ряда – интервальные и моментные ряды. **Интервальный ряд динамики** – последовательность, в которой уровень явления относится к результату, накопленному или вновь произведенному за определенный интервал времени. Например, ряды показателей объема продукции по месяцам года, количества отработанных человеко-дней по отдельным периодам т. д. В **моментном ряду** уровень ряда показывает фактическое наличие изучаемого явления в конкретный момент времени. Примеры моментных рядов: последовательности показателей численности населения на начало года, величины запаса какого-либо материала на начало периода и т. д. Важное аналитическое отличие моментных рядов от интервальных состоит том, что сумма уровней интервального ряда дает конкретный реальный показатель, обобщающий весь ряд динамики. Сумма уровней моментного ряда реального содержания, как правило, не имеет.
- 2. По **способу выражения** уровней ряды динамики подразделяются на ряды **абсолютных, относительных и средних величин**.
- 3. По **расстоянию между датами или интервалами времени** различают полные и неполные хронологические ряды. **Полные** ряды имеют место, когда даты регистрации или окончания периодов следуют друг за другом с равными интервалами. Это равноотстоящие ряды динамики. **Неполные** – когда принцип равных интервалов не соблюдается.

- 4. По **числу показателей** можно выделить изолированные и комплексные (многомерные) ряды динамики. Если ведется анализ во времени одного показателя, имеем **изолированный** ряд динамики. **Комплексный** ряд динамики получаем в том случае, когда в хронологической последовательности дается система показателей, связанных между собой единством процесса или явления.
- 5. По **наличию основной тенденции** изучаемого процесса ряды динамики подразделяются на **стационарные и нестационарные**. Если математическое ожидание значения признака и дисперсия (основные характеристики случайного процесса) постоянны, не зависят от времени, то процесс считается стационарным и ряды динамики также называются стационарными. Экономические процессы во времени обычно не являются стационарными, так как содержат основную тенденцию развития, но их можно преобразовать в стационарные путем исключения основной тенденции.

- Ряд динамики может быть подвержен влиянию факторов эволюционного и осциллятивного характера, которое можно представить в виде комбинации циклических, сезонных и случайных колебаний.
- Влияния эволюционного характера – это изменения, определяющие общее направление развития. Такие изменения динамического ряда называются **тенденцией развития, или трендом.**

- Следовательно, первоначальные значения ряда динамики подвергаются самым разнообразным воздействиям. Выделим его четыре основные компоненты: основную тенденцию (тренд) (Т), циклическую или конъюнктурную (К), сезонную (S), случайные колебания (Е). Если ряд динамики разбить на различные компоненты, то он представляется в следующем виде:
- $y=f(T,K,S,E)$ (8.1)
- В зависимости от взаимосвязи их между собой может быть построена **аддитивная** или **мультипликативная модель** ряда динамики.
- **Аддитивная** модель ряда динамики
- $y=T+K+S+E$ (8.2)
- характеризуется главным образом тем, что характер циклических и сезонных флюктуации (колебаний) остается постоянным.
- **Мультипликативная** модель ряда динамики
- $y = T \cdot K \cdot S \cdot E.$ (8.

- Ряды динамики охватывают значительные периоды времени, за которые могли произойти изменения, приводящие к несопоставимости статистических данных. Основными причинами несопоставимости уровней ряда динамики являются:
 - 1) изменение **единиц измерения или единиц счета**.
 - 2) изменение **методологии учета или расчета показателей**. Например, расчет средней урожайности с засеянной площади или с убранной.
 - 3) **экономический смысл интервалов или моментов**, по которым определены уровни.
 - 4) **неравные интервалы**, по которым даны уровни. Нельзя сравнивать квартальную продукцию с годовой.
 - 5) **изменение круга охватываемых объектов** вследствие перехода ряда объектов из одного подчинения в другое.
 - 6) **изменение территориальных границ** областей, районов и т. д.

Показатели динамики

- **Уровень ряда** является исходным показателем, характеризующим динамику социально-экономического явления. Его значения получают из практики, а не в результате расчета.
- **Абсолютный прирост** – это показатель, характеризующий абсолютные изменения уровня ряда, произошедшие за определенный период времени. Если рассматривается один, последний период, то говорят о цепном абсолютном приросте, если за период от базы до текущего момента времени – то о базисном
- $\Delta = Y_1 - Y_0$ $\Delta = Y_1 - Y_{i-1}$
- Между цепными и базисными абсолютными приростами существует взаимосвязь: базисный прирост равен сумме цепных приростов
- Положительное значение абсолютного прироста показывает, что наблюдается увеличение уровней с течением времени, отрицательное говорит о снижении уровней.

- **Абсолютное ускорение** – это разница между соседними абсолютными приростами. Оно характеризует скорость изменения абсолютных приростов. Если абсолютные ускорения колеблются вокруг 0, то это свидетельствует о равенстве абсолютных приростов, то есть равномерном росте (снижении) уровней ряда. Обычно абсолютное ускорение рассчитывается по цепным темпам роста:
- **Абсолютное значение одного процента прироста** – показатель, который используется для анализа других показателей (в частности, темпа прироста) и самостоятельного значения не имеет. Цепной показатель рассчитывается по формуле:
 - $A = Y_0 / 100$
 -

Относительные показатели

- Следует отметить, что все относительные показатели анализа динамики являются безразмерными, могут выражаться либо в долях единицы, либо в процентах. Обычно при расчетах используют запись в долях единицы, а при интерпретации результатов – в процентах. Само слово «процент» означает «сотая часть», и содержательно запись в процентах не отличается от записи в долях единицы
- Иногда, чтобы разделить запись с процентами и без них, используют различия в терминологии: если показатель записывается без процентов, то говорят о коэффициенте (роста, прироста), а если с процентами, то о темпе (роста, прироста). Такое разделение несколько искусственно, и мы не будем его использовать. В дальнейшем относительные показатели, вне зависимости от формы записи, мы будем называть темпами (роста, прироста).
- Рассмотрим расчетные формулы относительных показателей анализа динамики.
- **Темп роста** – основной относительный показатель анализа динамики. Он показывает, во сколько раз текущий уровень ряда больше либо предшествующего (если рассчитывается цепной показатель), либо базисного (если рассчитывается базисный показатель). Расчет темпа роста выполняется по формулам – цепной показатель и – базисный показатель.

- $T_p = y_1 / y_0$
- $T_p = y_1 / y_{i-1}$
- $TR_{\text{баз}} = \prod T_p \text{ цеп}$