

МОУ ДОД ДЮЦ «ЕДИНСТВО»

**СТАТИСТИЧЕСКИЕ
МЕТОДЫ
ОБРАБОТКИ ДАННЫХ**

22 слайда

Этапы статистического исследования

- 1. Сбор информации
- 2. Обработка информации
- 3. Анализ данных

Виды статистических признаков:

- *количественные*;
- *качественные* (атрибутивные, описательные)
- *альтернативные*, имеющие только два значения: да (1) или нет (0).

Виды ошибок

статистического наблюдения:

- I. Ошибки наблюдения *по источнику происхождения:*
 - - *преднамеренные* (тенденциозные);
 - - *непреднамеренные* (случайные);
 - - *ошибки запоминания;*
 - - *ошибки, связанные с фактором времени.*
- II. *Ошибки статистических расчетов.*

Виды статистического наблюдения

- I. Виды наблюдения *по характеру регистрации фактов*:
 - - *текущее (непрерывное)*;
 - - *периодическое (прерывное)*;
 - - *единовременное (разовое)*.
- II. Виды наблюдения *по степени охвата единиц изучаемой совокупности*:
 - - *сплошное*;
 - - *несплошное*.

Разновидности несплошного наблюдения

- ***выборочное***
- *способ основного массива*
- *анкетные обследования*
- *монографическое исследование*

Таблица оптимальных соотношений объема исследуемой совокупности (n) и числа групп (s)

n	15-24	25-44	45-89	90-179	180-359	360-719
s	5	6	7	8	9	10

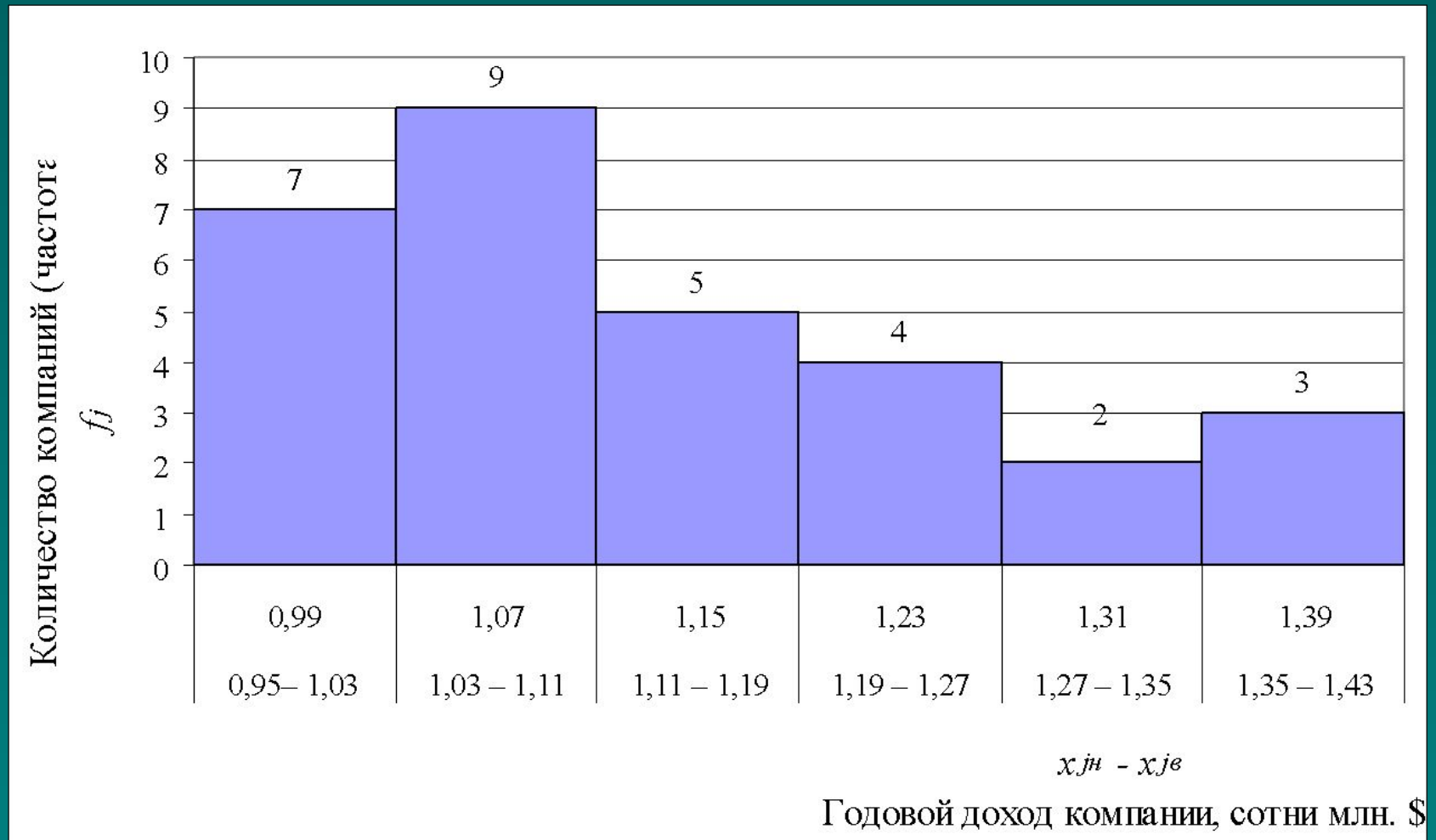
При группировке данных необходимо знать следующее:

- 1. Минимальное и максимальное значение признака в группе называется *нижней и верхней границей интервала*.
- 2. Интервалы, в которых указана лишь одна граница, называются *открытыми интервалами*.
- 3. Пограничные значения признака всегда включаются в следующий интервал.
- 4. Ширина открытого интервала равна ширине прилежащего интервала.
- 5. Все интервалы группировки должны быть заполнены единицами наблюдения.
- 6. Мелкие группы (менее 10% совокупности) при группировке могут быть объединены рубрикой «прочие».

Группировки

- **1. Структурные группировки** – имеют своей целью выявление структуры совокупности.
- **2. Типологические группировки** – строятся по качественным признакам, имея своей целью выявление групп и типов явлений.
- **3. Аналитические группировки** – направлены на выявление взаимосвязи между отдельными признаками. В основании аналитических группировок лежит **факторный** признак x . Признак-следствие y называется результативным, т.е. **результативный признак** – это признак, который изменяется под влиянием факторного признака (-ов).

Построения гистограммы на основе равноинтервальной группировки



Средние величины

Средняя степенная

$$\bar{x} = \sqrt[k]{\frac{\sum x_i^k}{n}}$$

Средняя
арифметическая

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

$$\bar{x} = \frac{\sum x_j \cdot f_j}{\sum f_j}$$

Средняя
геометрическая

$$\bar{x}_g = \sqrt[n]{k_{p_1} \cdot k_{p_2} \cdot \dots \cdot k_{p_n}} = \sqrt[n]{\prod k_{p_j}}$$

Средняя
квадратическая

$$\bar{x}_q = \sqrt{\frac{\sum x_i^2}{n}} \quad \bar{x}_q = \sqrt{\frac{\sum x_j^2 \cdot f_j}{\sum f_j}}$$

Мода и медиана

- **Мода** (M_o) – это значение признака, наиболее часто встречающееся в данной совокупности.
- **Медиана** (M_e) – значение признака у единицы, стоящей в середине ранжированного массива.

Виды показателей вариации

1. Абсолютные показатели вариации признака:

• *Размах вариации (колебаний)* $R = x_{max} - x_{min}$

• *Дисперсия признака* $\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n}$

• *Среднее квадратическое отклонение* $\sigma_x = \sqrt{\sigma_x^2}$

$$\sigma_w = \sqrt{\sigma_w^2} = \sqrt{w(1-w)}$$

2. Относительные показатели вариации:

• *Коэффициент вариации* $V_\sigma = \frac{\sigma_x}{x} \cdot 100\%$

Ошибка репрезентативности состоит из трех компонент:

- - *случайная компонента* – возникает в том случае, когда в выборочную совокупность случайно попадают единицы, существенно искажающие представление о генеральной совокупности;
- - *систематическая компонента* – возникает при нарушении принципа случайности отбора;
- - *сущностная компонента* – связана с различиями в численности единиц и степени вариации признака в выборочной и генеральной совокупности (т.е. с сущностью выборочного наблюдения).

Средняя ошибка репрезентативности

Отбор	генерального среднего	генеральной доли
Повторный	$\mu_x = \sqrt{\frac{\sigma_x^2}{n}}$	$\mu_w = \sqrt{\frac{\sigma_w^2}{n}} = \sqrt{\frac{w(1-w)}{n}}$
Бесповторный	$\mu_x = \sqrt{\frac{\sigma_x^2}{n} \left(1 - \frac{n}{N}\right)}$	$\mu_w = \sqrt{\frac{\sigma_w^2}{n} \left(1 - \frac{n}{N}\right)}$

Способы отбора в выборочную совокупность

- **1. Собственно случайный отбор**
(метод жеребьевки или метод лото)
- **2. Механическая выборка**
- **3. Типический отбор с механической
выборкой**
- **4. Многоступенчатая выборка.**
- **5. Многофазная выборка**
- **6. Серийная (гнездовая) выборка**

Ошибка *многоступенчатой* выборки

$$\mu_x = \sqrt{\mu_1^2 + \frac{\mu_2^2}{n_1} + \frac{\mu_3^2}{n_1 \cdot n_2} + \dots + \frac{\mu_m^2}{n_1 \cdot n_2 \cdot \dots \cdot n_{m-1}}}$$

Доверительный интервал генерального среднего или генеральной доли

- - расчет границ генерального среднего

$$\bar{X} = \bar{x} \pm \Delta_x = \bar{x} \pm t\mu_x$$

- - расчета границ генеральной доли

$$W = w \pm t\mu_w = \pm w$$

Значения коэффициента кратности при различных уровнях доверительной вероятности

Уровень доверительной вероятности $p(t)$	0,683	0,950	0,954	0,990	0,997	0,999
Критерий кратности t	1,00	1,96	2,00	2,58	3,00	4,00

Формулы расчета необходимой численности выборки

Отбор	Для средней	Для доли
Повторный	$n = \frac{t^2 \sigma_x^2}{\Delta_x^2}$	$n = \frac{t^2 \sigma_w^2}{\Delta_w^2} = \frac{t^2 w(1-w)}{\Delta_w^2}$
Бесповторный	$n = \frac{t^2 \sigma_x^2 N}{\Delta_x^2 N + t^2 \sigma_x^2}$	$n = \frac{t^2 w(1-w)N}{\Delta_w^2 N + t^2 w(1-w)}$

Связи между явлениями и признаками

- **функциональная связь** (полная) – определенному значению факторного признака соответствует одно или несколько строго определенных значений результативного признака;
- **корреляционная связь** (статистическая, неполная) – одному и тому же значению факторного признака могут соответствовать разные значения результативного признака.
 - - *слабая,*
 - - *умеренная*
 - - *сильная (тесная)*

Аналитические методы корреляционного анализа

- **1. Метод приведения параллельных данных**
- **2. Метод построения корреляционных таблиц**
- **3. Графический метод**
- **4. Дисперсионный анализ**