

# Статистические методы обработки информации в научных исследованиях



К.м.н., доцент Рычкова Т.А.,  
К.м.н., доцент Таловерова Л.И.



**«В жизни, как правило,  
преуспевает тот, кто  
располагает лучшей  
информацией»**

**Б.Дизраэли**



# План:

- 1. Основные понятия статистического исследования.**
  - 2. Правила сбора и оформления данных для статобработки.**
  - 3. Выбор статистического метода исследования.**
  - 4. Анализ результатов статобработки данных и формулировка выводов.**
- 

**Статистика** — отрасль знаний, в которой излагаются общие вопросы:

- сбора,
- измерения,
- анализа данных.

Слово «статистика» происходит от латинского *status* — состояние дел.

В науку термин «статистика» ввёл немецкий ученый Готфрид Ахенваль в 1746 году.



**Статистика есть наука о том,  
как, не умея мыслить и понимать,  
заставить делать это цифры**

**Василий Осипович  
Ключевский**

# 4 ОСНОВНЫХ ЗАДАЧИ СТАТИСТИКИ

1. статистическое описание выборки
2. сравнение параметров двух и более выборок
3. определение связей между параметрами
4. динамический анализ (временной анализ изменения статистических параметров выборки)





# **Этапы статистического исследования**

**I. Составление плана  
статистического исследования,  
разработка программы**



**II. Регистрация и сбор  
статистического материала**

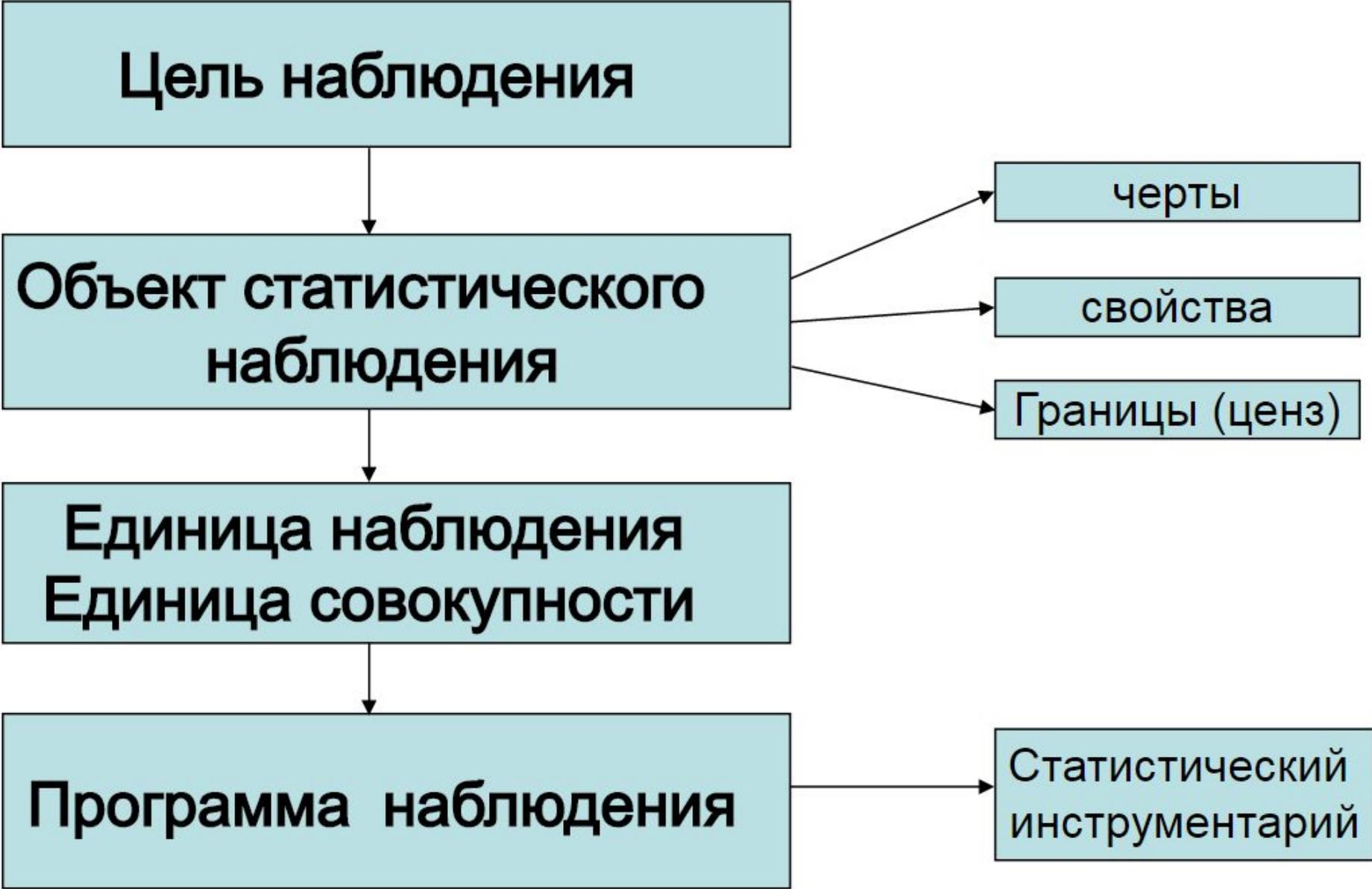


**III. Составление базы данных (БД)**



**IV. Статистический анализ**

**V. Выводы, внедрение  
результатов исследования в  
практику**



```
graph TD; A[Цель наблюдения] --> B[Объект статистического наблюдения]; B --> C[Единица наблюдения  
Единица совокупности]; C --> D[Программа наблюдения]; B --> E[черты]; B --> F[свойства]; B --> G[Границы (ценз)]; D --> H[Статистический  
инструментарий];
```

Цель наблюдения

Объект статистического наблюдения

черты

свойства

Границы (ценз)

Единица наблюдения  
Единица совокупности

Программа наблюдения

Статистический  
инструментарий



**Обработать статистическими методами  
можно лишь то, что поддается измерению**

**Измерение – присвоение объекту (явлению,  
событию, процессу и т.п.) **числа** (значения)  
по определенному правилу**

**Признак – измеренное свойство объекта  
наблюдения**

**Данные**

**Даты**

**Количественные  
данные**

**Качественные  
данные**

**Дискретные**

Целые - дети

**Непрерывные**

Рост, Вес, t, АД

**Интервальные,  
абсолютные**

**Относительные, %**

**Номинальные**

Коды групп

**Порядковые**

Шкалы, оценки,  
стадии

# Характеристика данных

- **Количественные дискретные признаки** измеряются целыми числовыми значениями (например, кол-во детей, беременностей)
- **Количественные непрерывные признаки** измеряются любыми числовыми значениями (например, возраст, рост, вес, давление).
- **Качественные признаки** – характеризуют некоторое состояние объекта, но не могут быть измерены количественно (например, пол, профессия, диагноз).
- **Порядковые признаки** – могут быть измерены в шкалах (например, школьные оценки, степень тяжести заболевания – легкая (1), средняя (2), тяжелая (3) и т.д.).

# Абсолютные величины

**Абсолютные величины** – отображают численную величину измеренного явления, уровень развития явления

## Абсолютные величины в статистике:

- являются именованными
- измеряются в конкретных единицах (г, см, кг)

# Относительные величины

Относительная величина (относительный показатель, коэффициент) - это обобщающий показатель, который дает **числовую меру соотношения** двух сопоставляемых абсолютных величин

**Данные**

**Даты**

**Количественные  
данные**

**Качественные  
данные**

**Дискретные**

Целые - дети

**Непрерывные**

Рост, Вес, t, АД

**Интервальные,  
абсолютные**

**Относительные, %**

**Номинальные**

Коды групп

**Порядковые**

Шкалы, оценки,  
стадии

# Фрагмент БД в MS EXCEL:

|    | A     | B     | C          | D        | E        | F         | G          | H    | I     | J | K      | L      | M      | N     | O      |
|----|-------|-------|------------|----------|----------|-----------|------------|------|-------|---|--------|--------|--------|-------|--------|
| 1  | VAR1A | VAR2A | VAR3A      | VAR4A    | VAR5A    | VAR6A     | VAR7A      | VAR8 | VAR9A |   | VAR10A | VAR11A | VAR12A | VAR13 | VAR14A |
| 2  | 1     | 1     | 26.01.2010 | 15.02.10 | 144      | Фаустова  | 04.12.2006 | 3,15 | 1     | 1 | 21     | 1      | 1      | 3350  | 2      |
| 3  | 2     | 1     | 21.05.2010 | 04.06.10 | 686      | Щербина   | 11.02.2007 | 3,27 | 1     | 1 | 21     | 2      | 2      | 2800  | 2      |
| 4  | 3     | 1     | 10.06.2011 | 24.06.11 | 670      | Шавлак Я  | 22.01.2009 | 2,38 | 1     | 1 | 15     | 6      | 2      | 4000  | 2      |
| 5  | 4     | 1     | 12.07.2012 | 02.08.12 | 797      | Устинова  | 16.05.2009 | 3,16 | 1     | 1 | 15     | 2      | 2      | 3450  | 1      |
| 6  | 5     | 1     | 12.07.2012 | 13.07.11 | 08-09-20 | Кулакеви  | 13.07.2009 | 2,00 | 1     | 1 | 1      | 1      | 1      | 4100  | 1      |
| 7  | 6     | 1     | 21.08.2012 | 20.08.12 | 08-10-23 | Игнатенко | 14.01.2010 | 2,60 | 1     | 1 | 1      |        |        |       |        |
| 8  | 7     | 1     | 03.07.2012 | 02.07.12 | 08-10-02 | Жукова    | 07.02.2010 | 2,40 | 1     | 1 | 1      |        |        |       |        |
| 9  | 8     | 1     | 21.08.2012 | 19.08.12 | 08-10-07 | Веренич   | 19.02.2010 | 2,50 | 1     | 1 | 1      |        |        |       |        |
| 10 | 9     | 1     | 21.08.2012 | 22.08.12 | 08-10-49 | Ярова     | 23.08.2010 | 2,00 | 1     | 1 | 1      |        |        |       |        |
| 11 | 10    | 1     | 11.09.2012 | 10.09.12 | 08-10-17 | Дзюба     | 11.09.2010 | 2,00 | 1     | 1 | 1      | 1      | 1      | 2900  | 3      |
| 12 | 11    | 1     | 22.05.2012 | 21.05.12 | 08-10-51 | Новікова  | 15.10.2010 | 1,60 | 1     | 1 | 1      |        |        |       |        |
| 13 | 12    | 1     | 02.10.2009 | 16.10.09 | 1087     | Макаров В | 30.12.2006 | 2,76 | 2     | 2 | 21     | 2      | 2      | 2800  | 1      |
| 14 | 13    | 1     | 17.02.2010 | 09.03.10 | 206      | Селютин   | 25.09.2006 | 3,40 | 2     | 2 | 3      | 1      | 1      | 3100  | 1      |
| 15 | 14    | 1     | 07.06.2010 | 27.06.10 | 712      | Изосимов  | 06.02.2007 | 3,33 | 2     | 2 | 18     | 3      | 1      | 3200  | 1      |

# Описание БД:

## Цель исследования:

разработать и экспериментально проверить эффективность системы коррекционных занятий по развитию наглядно-образного мышления детей дошкольного возраста с минимальной мозговой дисфункцией.

# Описание БД:

**Объект исследования –**

наглядно-образное мышление детей дошкольного возраста с минимальной мозговой дисфункцией.

**Предмет исследования –**

особенности развития наглядно-образного мышления детей дошкольного возраста с минимальной мозговой дисфункцией.

# Описание БД:

## Обследовано:

24 ребенка с ММД и  
20 здоровых детей 5-6 лет по 23  
признакам. Из них 15 качественных и 8  
количественных.

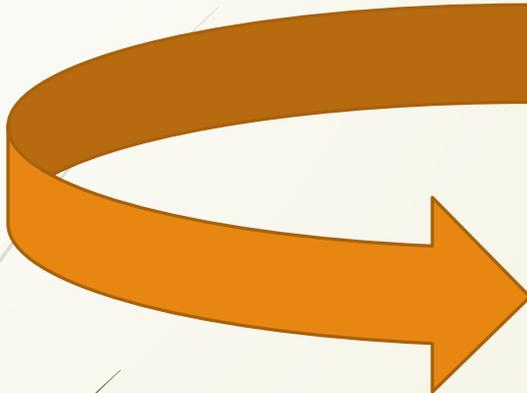
# Методы исследования:

- **тест рисования часов**
- **заучивание и воспроизведение 5 слов**
- **исследование памяти (серийный и обратный счет, вербальная и невербальная память, ассоциации)...**

## Задачи исследования:

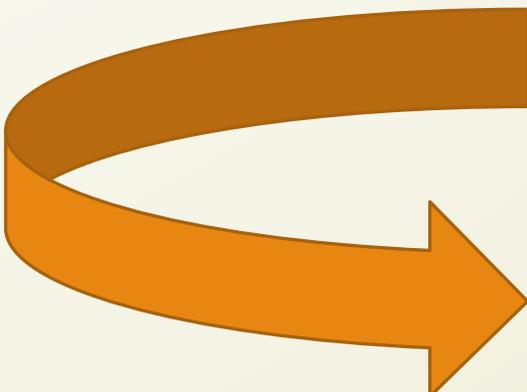
1. Установить частоту встречаемости, выраженность и характер нарушений \_\_\_\_\_ (VAR7-VAR14)...
2. Выявить частоту встречаемости, степень выраженности и факторы риска \_\_\_\_\_ VAR4A-VAR6A...
3. Выявить взаимосвязь между \_\_\_\_\_ VAR9A-VAR13A...

# Особенности описания признаков:



Количественные признаки:

**VAR1** = «возраст, лет»



Качественные признаки:

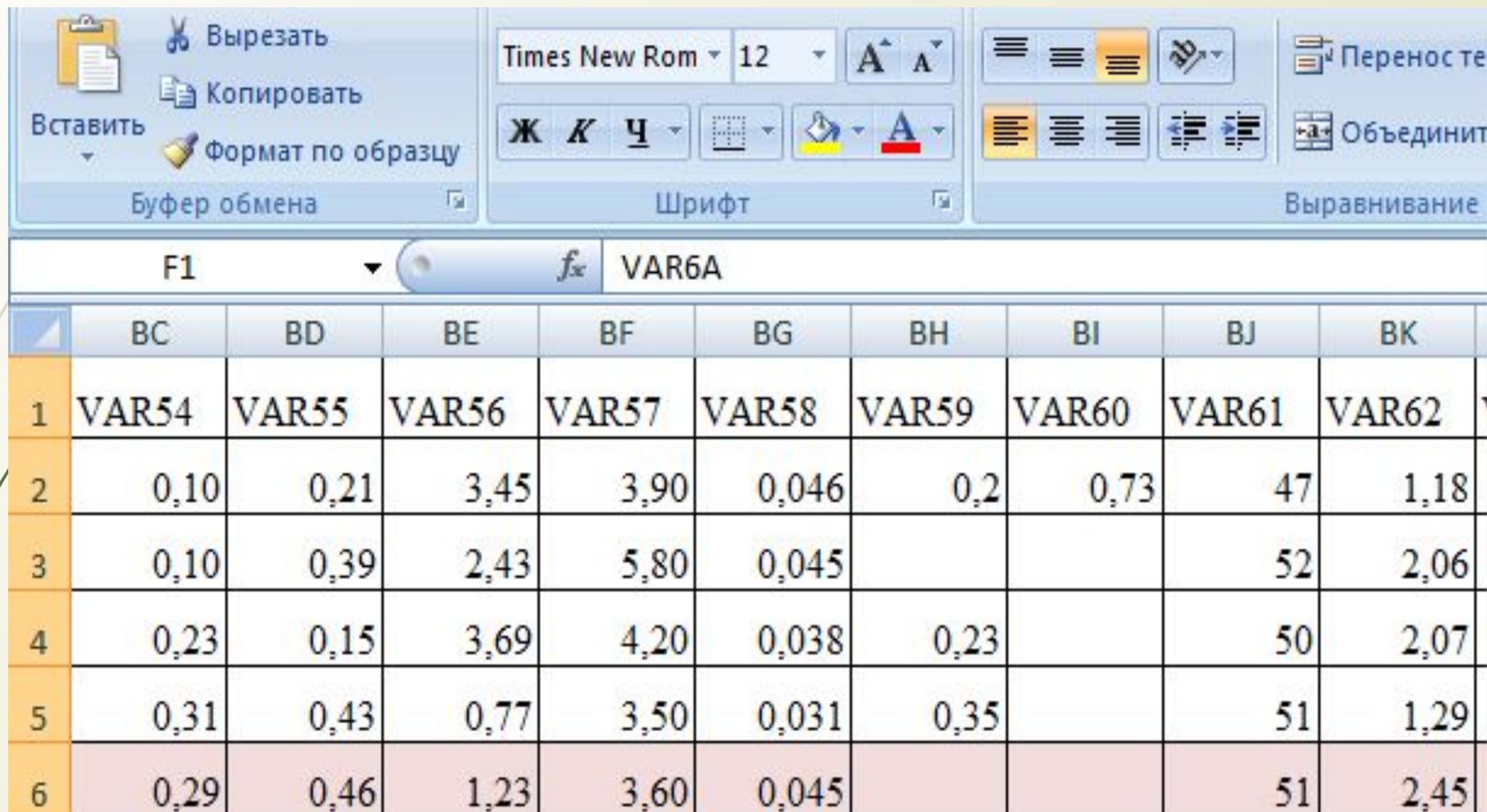
**VAR2A** = «пол»

1-девочка

2-мальчик

|   | A     | B     | C          |
|---|-------|-------|------------|
| 1 | VAR1A | VAR2A | VAR3A      |
| 2 | 1     | 1     | 26.01.2010 |
| 3 | 2     | 1     | 21.05.2010 |
| 4 | 3     | 1     | 10.06.2011 |
| 5 | 4     | 1     | 12.07.2012 |

# VAR61=«ОКРУЖНОСТЬ ГОЛОВЫ, СМ»



|   | BC    | BD    | BE    | BF    | BG    | BH    | BI    | BJ    | BK    |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | VAR54 | VAR55 | VAR56 | VAR57 | VAR58 | VAR59 | VAR60 | VAR61 | VAR62 |
| 2 | 0,10  | 0,21  | 3,45  | 3,90  | 0,046 | 0,2   | 0,73  | 47    | 1,18  |
| 3 | 0,10  | 0,39  | 2,43  | 5,80  | 0,045 |       |       | 52    | 2,06  |
| 4 | 0,23  | 0,15  | 3,69  | 4,20  | 0,038 | 0,23  |       | 50    | 2,07  |
| 5 | 0,31  | 0,43  | 0,77  | 3,50  | 0,031 | 0,35  |       | 51    | 1,29  |
| 6 | 0,29  | 0,46  | 1,23  | 3,60  | 0,045 |       |       | 51    | 2,45  |

# Описание признаков:

**VAR5A=«тревога, балл» НОРМА: 0-7**

**VAR6A=«тревога»**

**1-норма**

**2-не норма**

# Матрица результатов тестирования

| №  | ФИО<br>тестирующегося | Номер задания    |          |          |          |          |          |          |          |          |          |
|----|-----------------------|------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|    |                       | 1                | 2        | 3        | 4        | 5        | 6        | 7        | 8        | 9        | 10       |
|    |                       | Правильный ответ |          |          |          |          |          |          |          |          |          |
|    |                       | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>B</i> |
| 1  | Иванов В.А.           | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>A</i> | <i>A</i> | <i>C</i> | <i>D</i> |
| 2  | Петров П.П.           | <i>C</i>         | <i>D</i> | <i>C</i> | <i>D</i> | <i>A</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>C</i> | <i>C</i> |
| 3  | Дроздов П.У.          | <i>A</i>         | <i>C</i> | <i>B</i> | <i>A</i> | <i>A</i> | <i>A</i> | <i>A</i> | <i>A</i> | <i>A</i> | <i>D</i> |
| 4  | Вахненко М.К.         | <i>A</i>         | <i>C</i> | <i>D</i> | <i>A</i> | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>A</i> | <i>A</i> |
| 5  | Елизарова М.М.        | <i>C</i>         | <i>D</i> | <i>D</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>B</i> |
| 6  | Лысенко Т.О.          | <i>C</i>         | <i>B</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>B</i> |
| 7  | Голуб А.В.            | <i>C</i>         | <i>D</i> | <i>A</i> | <i>D</i> | <i>D</i> | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>C</i> |
| 8  | Файзулин М.Р.         | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>A</i> | <i>B</i> | <i>A</i> | <i>D</i> | <i>C</i> | <i>A</i> |
| 9  | Игнатович Р.Т.        | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>D</i> |
| 10 | Роднина М.В.          | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>D</i> | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> |
| 11 | Правдина М.Л.         | <i>C</i>         | <i>D</i> | <i>A</i> | <i>C</i> | <i>C</i> | <i>B</i> | <i>D</i> | <i>C</i> | <i>D</i> | <i>B</i> |

# Описание признаков:

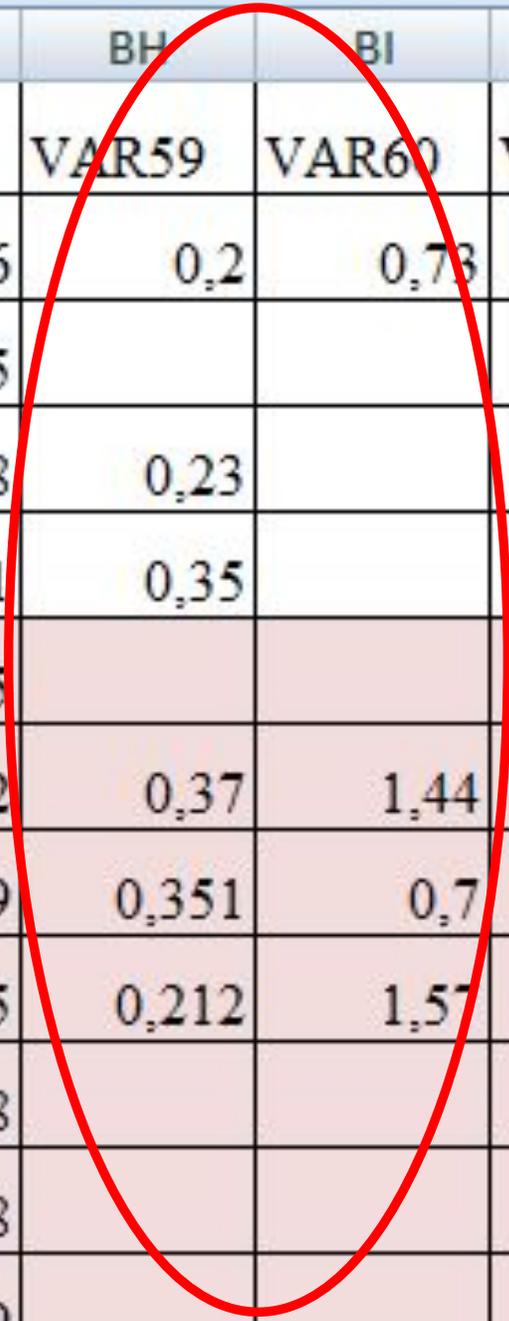
**VAR5A=«тревога, балл» НОРМА: 0-7**

**VAR6A=«тревога»**

**1-норма**

**2-не норма**





|    | BC    | BD    | BE    | BF    | BG    | BH    | BI    | BJ    | BK    | BL    |
|----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1  | VAR54 | VAR55 | VAR56 | VAR57 | VAR58 | VAR59 | VAR60 | VAR61 | VAR62 | VAR63 |
| 2  | 0,10  | 0,21  | 3,45  | 3,90  | 0,046 | 0,2   | 0,73  | 47    | 1,18  | 19    |
| 3  | 0,10  | 0,39  | 2,43  | 5,80  | 0,045 |       |       | 52    | 2,06  | 22    |
| 4  | 0,23  | 0,15  | 3,69  | 4,20  | 0,038 | 0,23  |       | 50    | 2,07  | 18    |
| 5  | 0,31  | 0,43  | 0,77  | 3,50  | 0,031 | 0,35  |       | 51    | 1,29  | 20    |
| 6  | 0,29  | 0,46  | 1,23  | 3,60  | 0,045 |       |       | 51    | 2,45  | 21    |
| 7  | 0,10  | 0,44  | 3,14  | 2,00  | 0,042 | 0,37  | 1,44  | 46    | 1,66  | 19    |
| 8  | 0,10  | 0,30  | 1,16  | 3,70  | 0,039 | 0,351 | 0,7   | 48    | 2,04  | 22    |
| 9  | 0,57  | 0,31  | 2,53  | 2,90  | 0,05  | 0,212 | 1,57  | 52    | 2,19  | 17    |
| 10 | 0,14  | 0,33  | 1,03  | 2,30  | 0,048 |       |       | 44    | 1,64  | 19    |
| 11 | 0,15  | 0,34  | 1,97  | 3,10  | 0,048 |       |       | 49    | 2,27  | 21    |
| 12 | 0,19  | 0,46  | 2,30  | 4,40  | 0,039 |       |       | 40    | 1,64  | 23    |

# Что такое переменная?

**Переменная (анг. variable) — это то, что можно измерять или контролировать. Иными словами, переменная — это то, что изменяется.**

**ПРИМЕРЫ: анкетные данные, АД, доля зрителей, скорость, температура, объем, оценка по шкале**

## **Характеристики переменных:**

**Переменные не постоянны, нужно научиться описывать их изменчивость.**

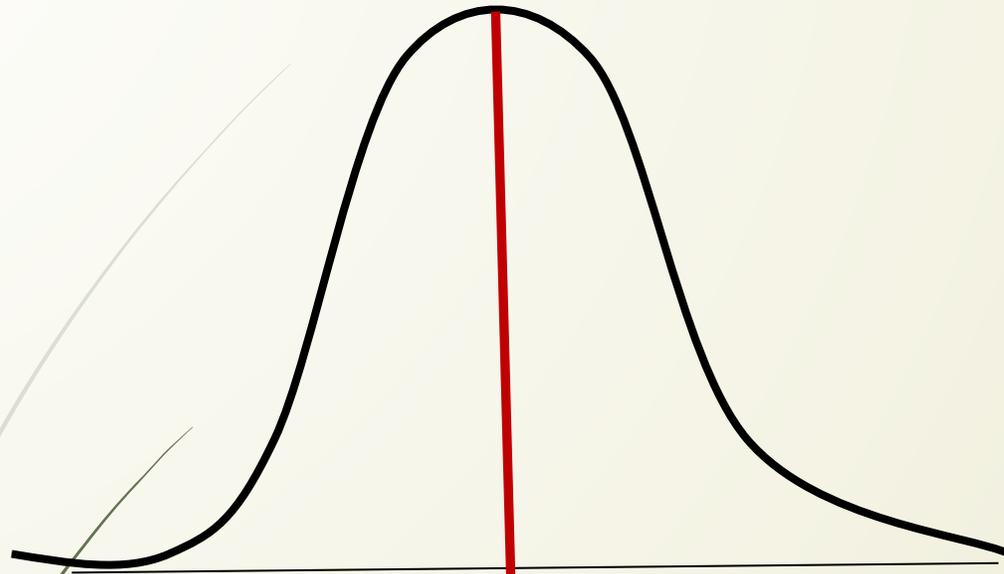
**Для этого придуманы описательные или дескриптивные статистики.**

## Характеристики переменных:

Медиана ( $Me$ ) разбивает выборку на две равные части. Половина значений переменной лежит ниже медианы, половина — выше.

$Me$  дает общее представление о том, где сосредоточены значения переменной, иными словами, где находится ее центр.

## Симметричное распределение

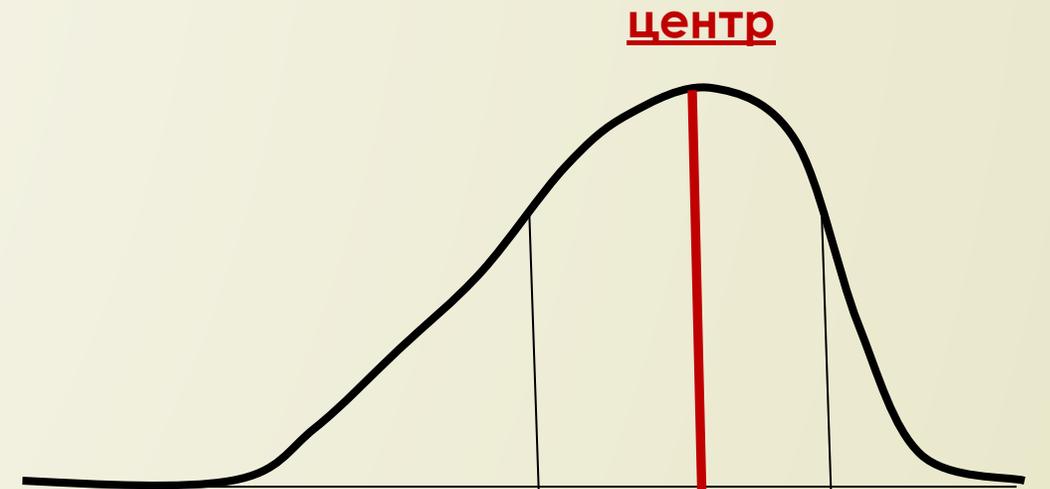


**Медиана**

**Среднее**

**Мода**

## Асимметричное распределение



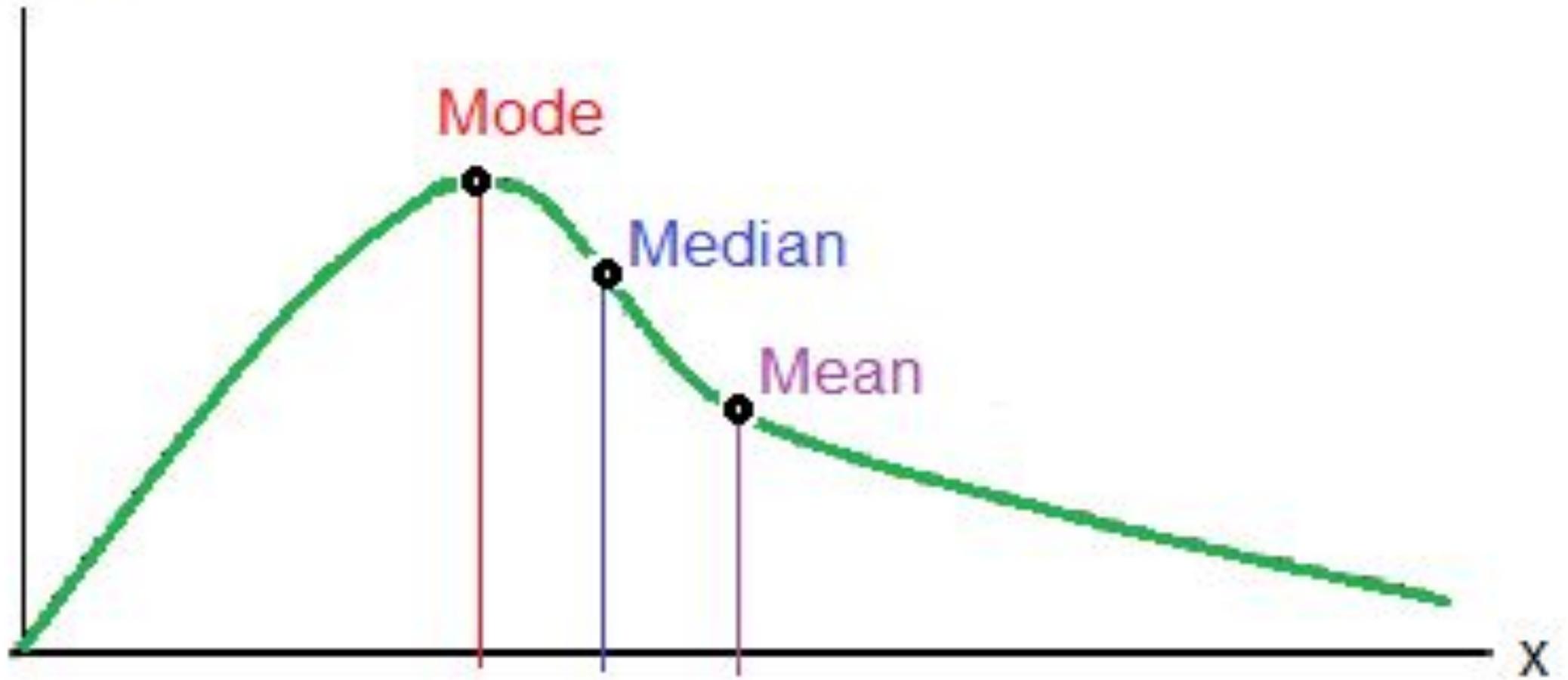
центр

**Среднее**

**Мода**

**Медиана**

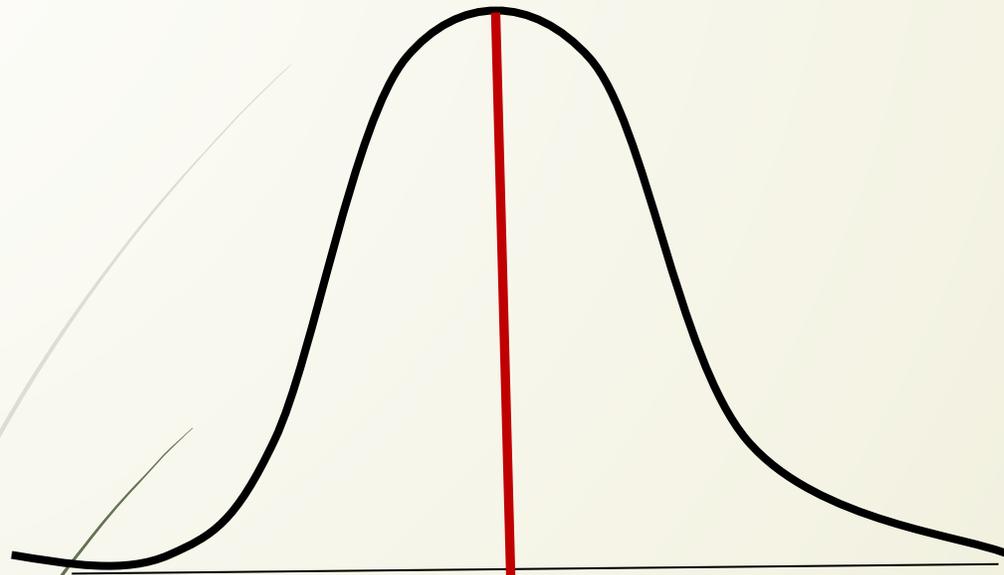
$p(x)$



# Характеристики переменных:

**Мода (Mo)** представляет собой максимально часто встречающееся значение переменной (иными словами, наиболее «модное» значение переменной).

## Симметричное распределение

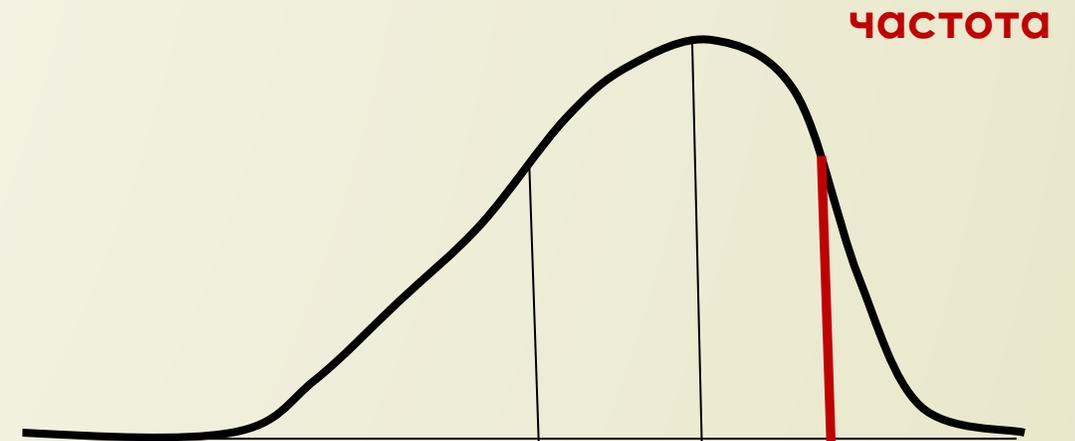


**Мода**

**Медиана**

**Среднее**

## Асимметричное распределение



**Частота**

**Среднее**

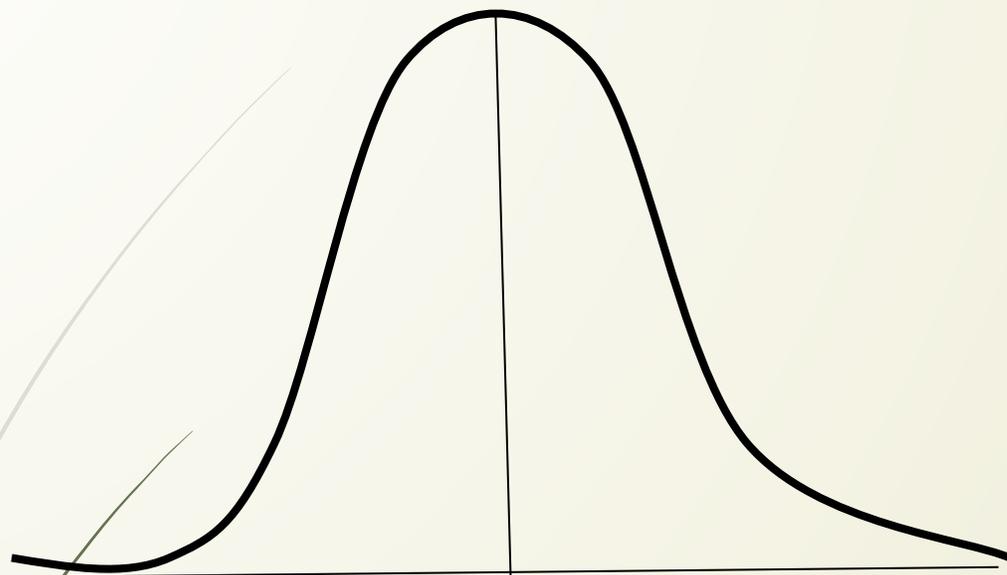
**Мода**

**Медиана**

# Характеристики переменных:

**Среднее - сумма значений  
переменной, деленная на  $n$  (число  
значений переменной).**

## Симметричное распределение

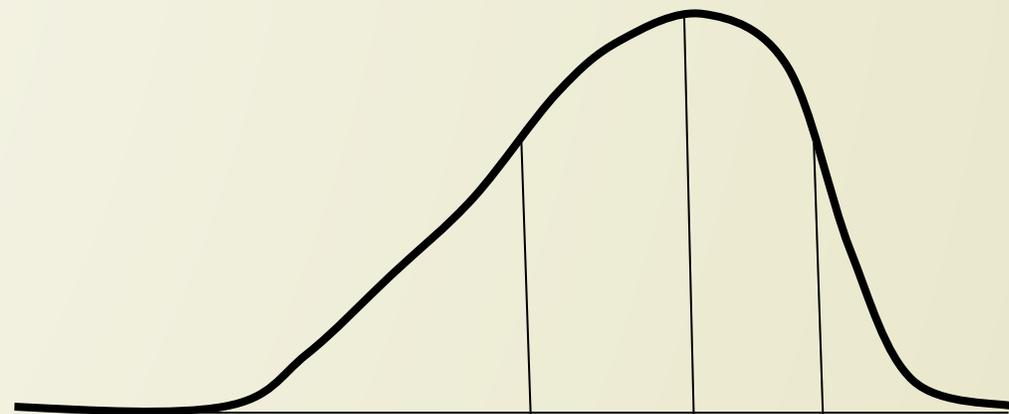


**Среднее**

**Медиана**

**Мода**

## Асимметричное распределение



**Среднее**

**Мода**

**Медиана**

# Характеристики переменных:

**Минимум и максимум — это минимальное и максимальное значения переменной.**



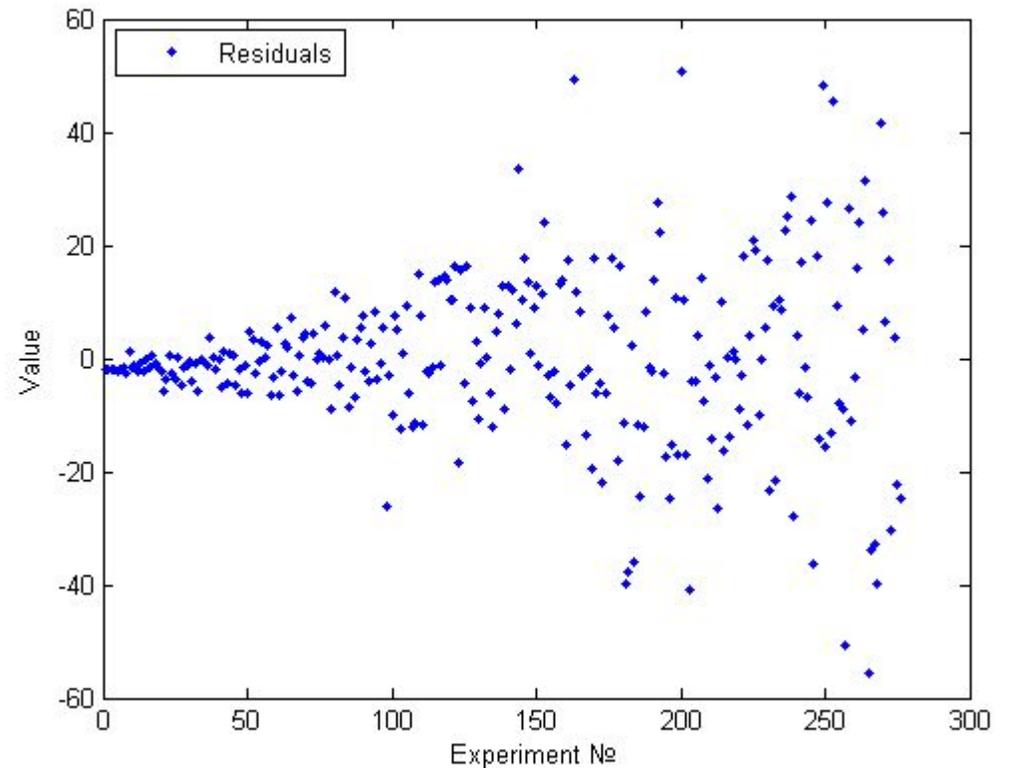
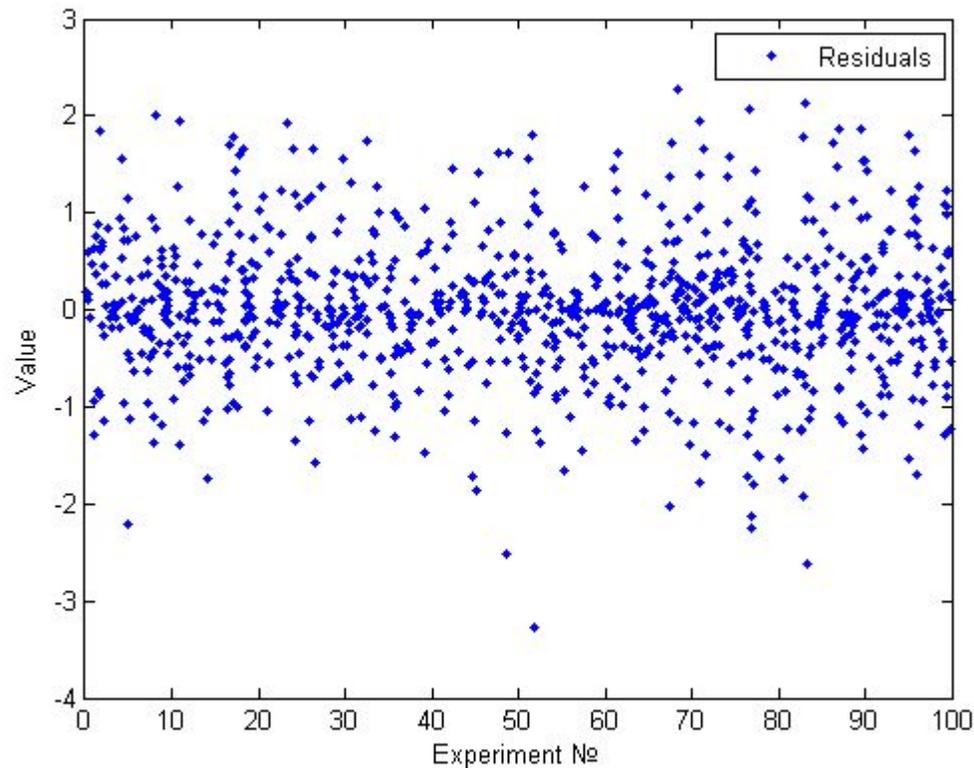
**Минимум**



**Максимум**

# Характеристики переменных:

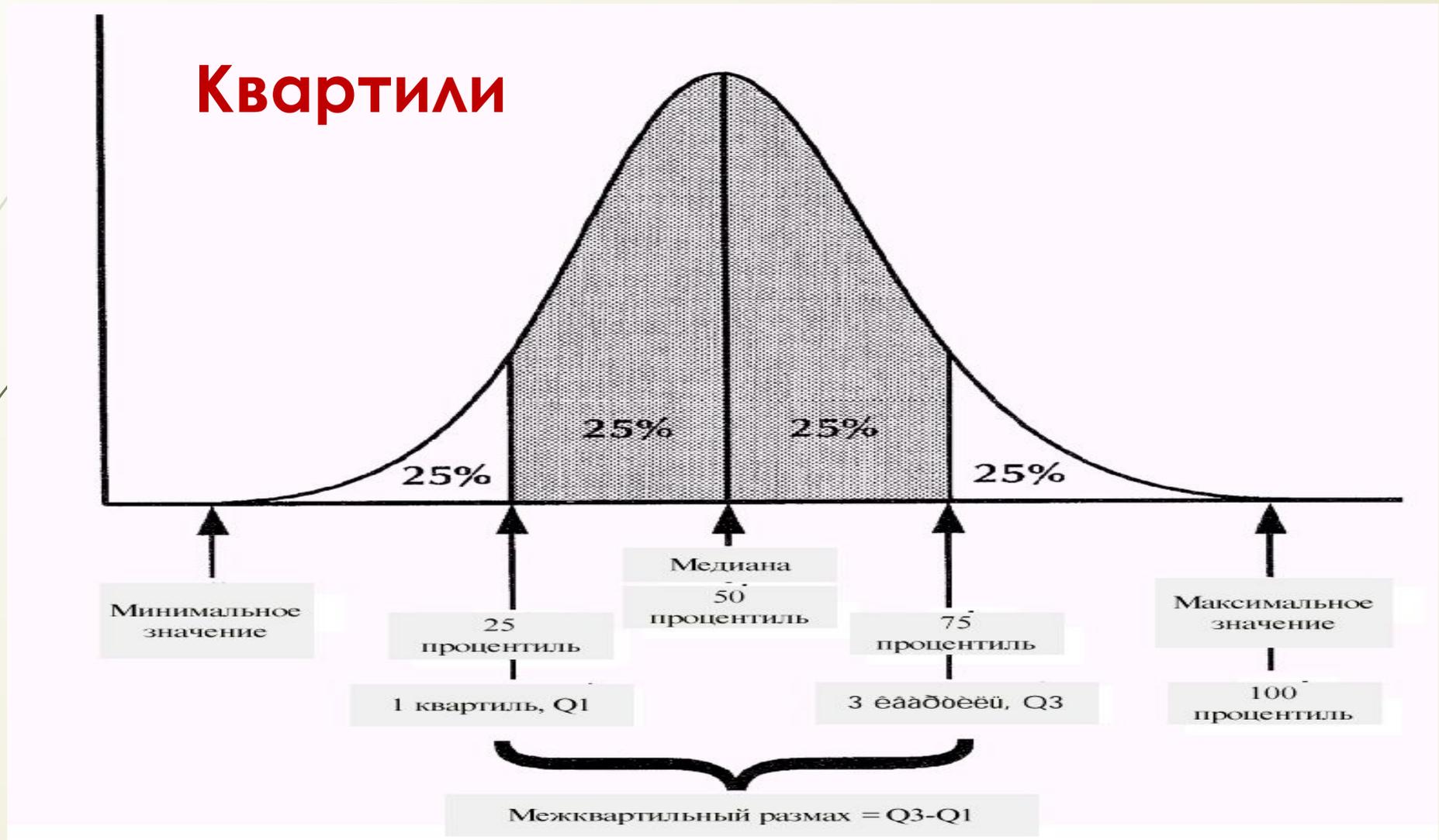
**Дисперсия - мера разброса данной переменной.**



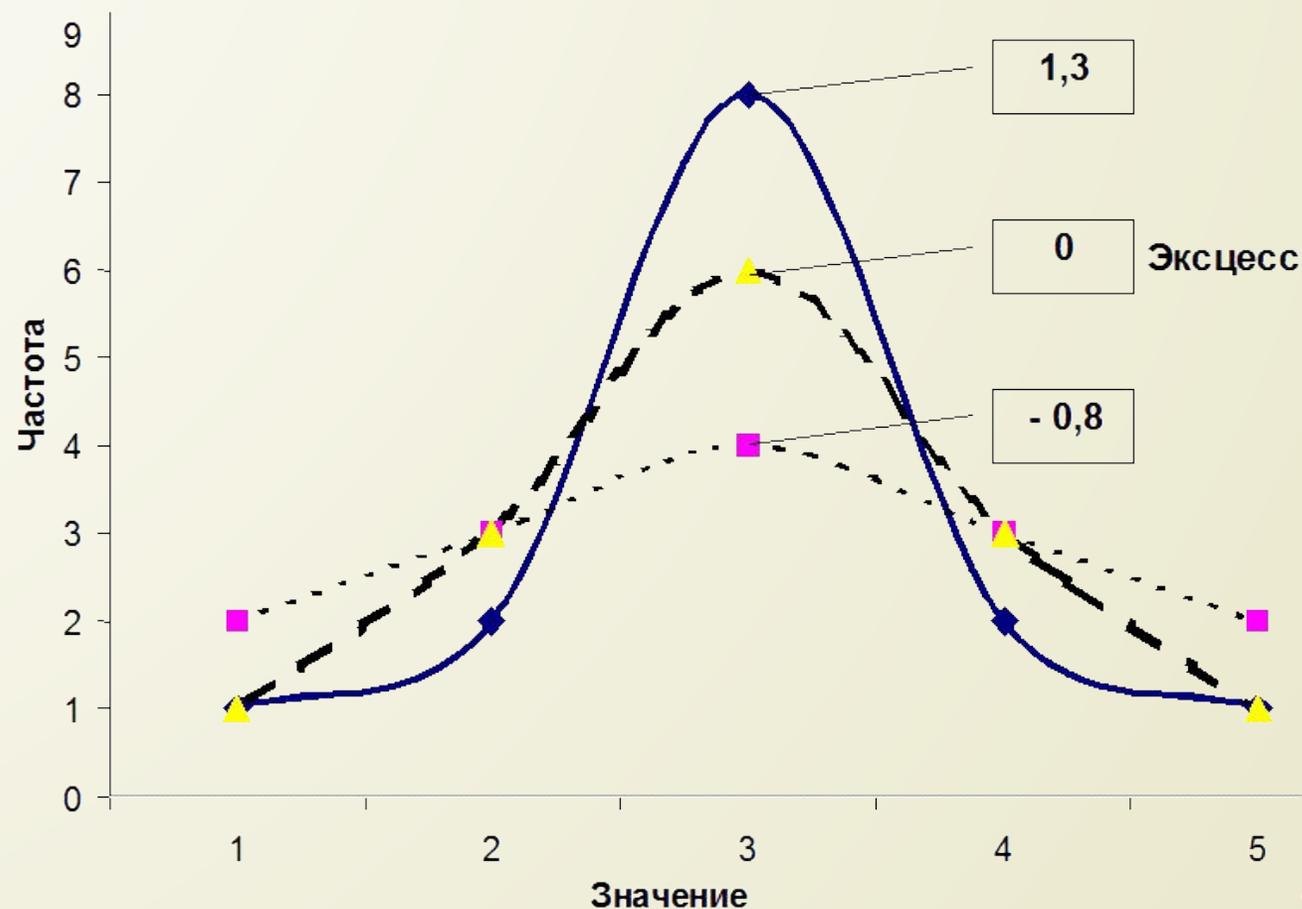
# Характеристики переменных:

- **квартили,**
- **коэффициент асимметрии,**
- **эксцесс,**
- **коэффициент корреляции и др.**

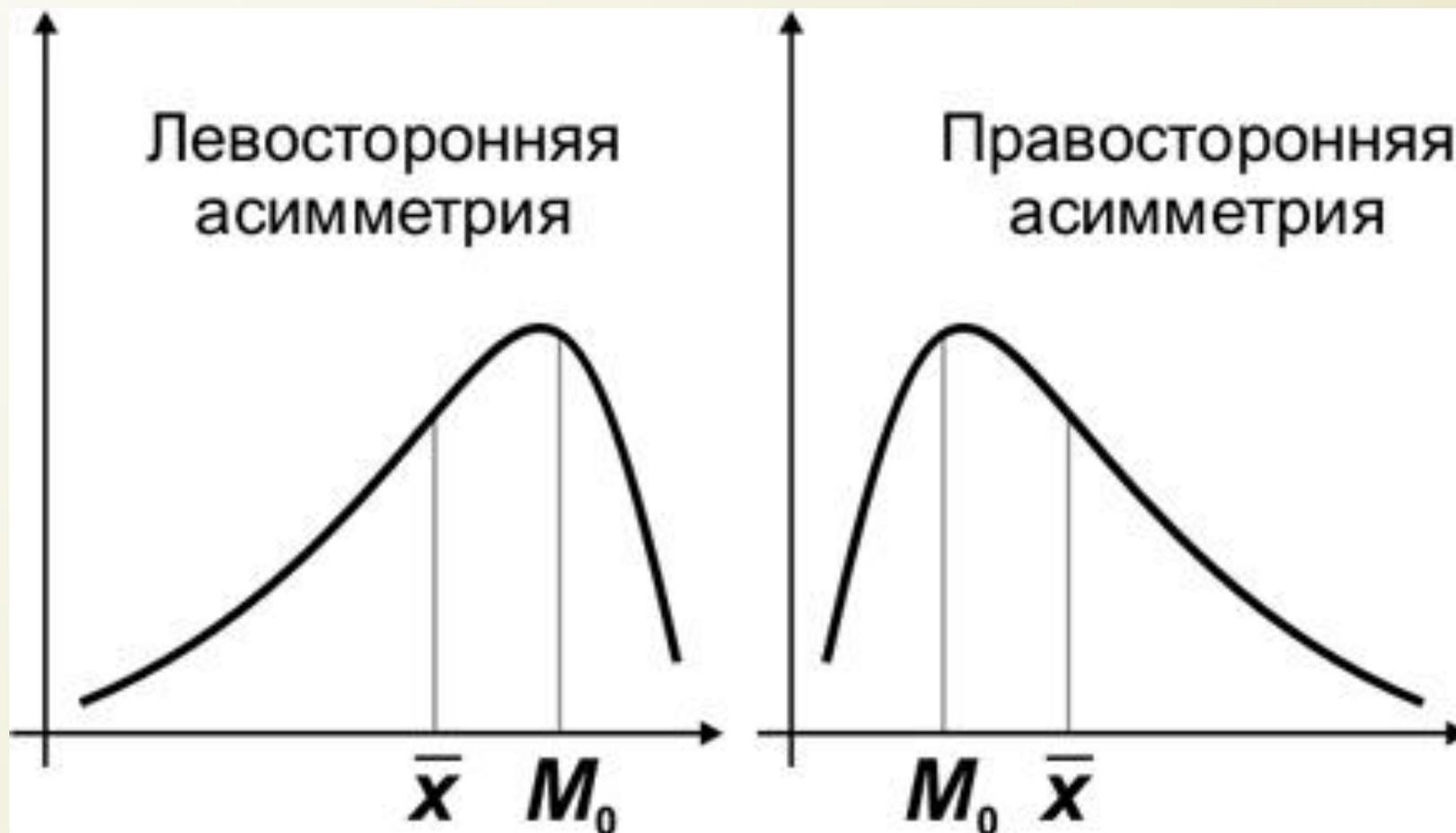
# Характеристики переменных:



**Эксцесс** – это показатель относительной крутости кривой вариационного ряда по сравнению с нормальным распределением. Эксцесс нормально распределенной случайной величины равен 0.



**Коэффициент асимметрии** — величина, характеризующая асимметрию распределения данной случайной величины



ФАЙЛ ГЛАВНАЯ ВСТАВКА РАЗМЕТКА СТРАНИЦЫ ФОРМУЛЫ **ДАННЫЕ** РЕЦЕНЗИРОВАНИЕ

Из Access Из Интернета Из других источников Из текста Существующие подключения Обновить все Подключения Свойства Изменить связи Сортировка Фильтр

Сергей Багузин **Анализ данных** Поиск решения

Получение внешних данных Подключения Сортировка и с Анализ

|    | A    | B                      | C    | D     |
|----|------|------------------------|------|-------|
| 1  | -6,1 | Столбец1               |      |       |
| 2  | -2,8 |                        |      |       |
| 3  | -1,2 | Среднее                |      | 6,08  |
| 4  | -0,7 | Стандартная ошибка     |      | 1,71  |
| 5  | 4,3  | Медиана                |      | 6,50  |
| 6  | 5,5  | Мода                   | #Н/Д |       |
| 7  | 5,9  | Стандартное отклонение |      | 6,62  |
| 8  | 6,5  | Дисперсия выборки      |      | 43,81 |
| 9  | 7,6  | Экссесс                |      | -0,22 |
| 10 | 8,3  | Асимметричность        |      | -0,15 |
| 11 | 9,6  | Интервал               |      | 24,60 |
| 12 | 9,8  | Минимум                |      | -6,10 |
| 13 | 12,9 | Максимум               |      | 18,50 |
| 14 | 13,1 | Сумма                  |      | 91,20 |
| 15 | 18,5 | Счет                   |      | 15,00 |

### Описательная статистика

**Входные данные**  
 Входной интервал: **SAS1:SAS15**

Группирование:
 

- по столбцам
- по строкам

 Метки в первой строке

**Параметры вывода**

- Выходной интервал: **SCS1**
- Новый рабочий лист:
- Новая рабочая книга
- Итоговая статистика
  - Уровень надежности: 95 %
  - К-ый наименьший: 2
  - К-ый наибольший: 2

ОК Отмена Справка

# Понятие «Нулевая гипотеза»:

**«Нулевая гипотеза» — это предположение о том, что в сравниваемых группах отсутствует различие в распределении частот.**

# Уровень значимости

**- это такая вероятность, которую принимают за основу при статистической оценке гипотезы.**

## Уровень значимости 5%.

В качестве максимального уровня значимости, при котором нулевая гипотеза еще отклоняется, **принимается 5%**.

При уровне значимости **более 5%** «нулевая гипотеза» принимается, различия между сравниваемыми совокупностями принимаются статистически незначимыми.



# Как выбрать статистический метод?

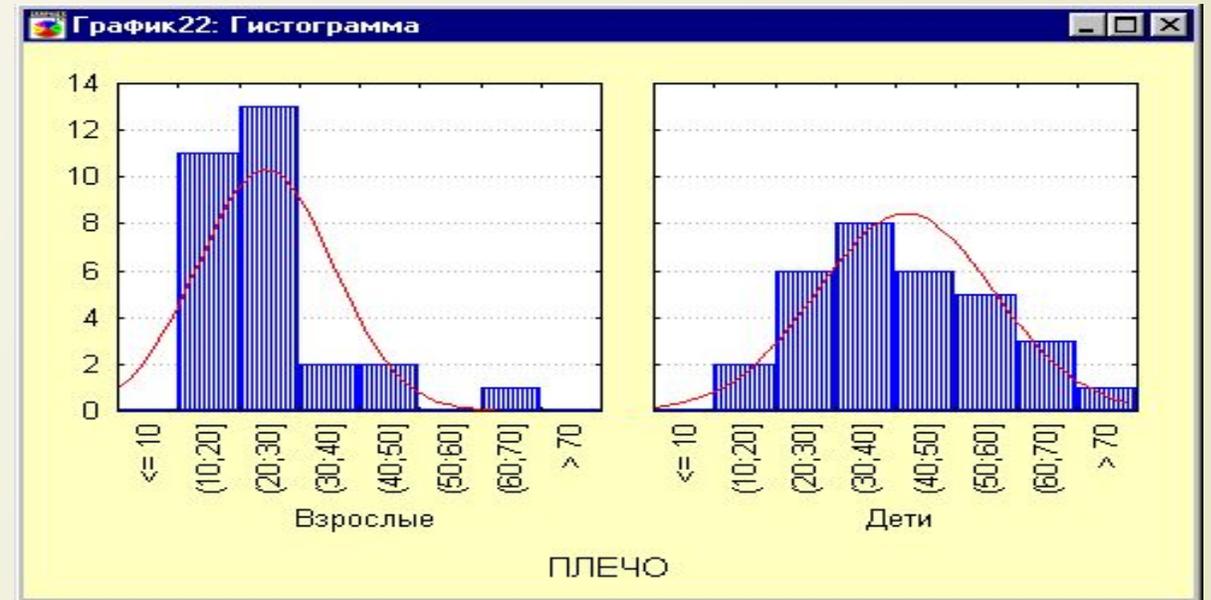
# Как выбрать статистический метод?

- Если признак числовой, стоит подумать, нормально ли его распределение.
- Если Вы имеете дело с порядковыми и качественными признаками, то подходят только непараметрические методы.
- Если данных мало (или Вы не хотите думать о типе распределения) - воспользуйтесь непараметрическими методами.

# Основные цели первичного (разведочного) анализа

Определение характера распределений переменных, визуальный анализ зависимостей и идентификация возможных выбросов.

Нормальное  
Можно применять стандартные методы:  $t$ -критерии и дисперсионный анализ.



Отличное от нормального (или малая выборка)  
Необходимо использовать непараметрические критерии.

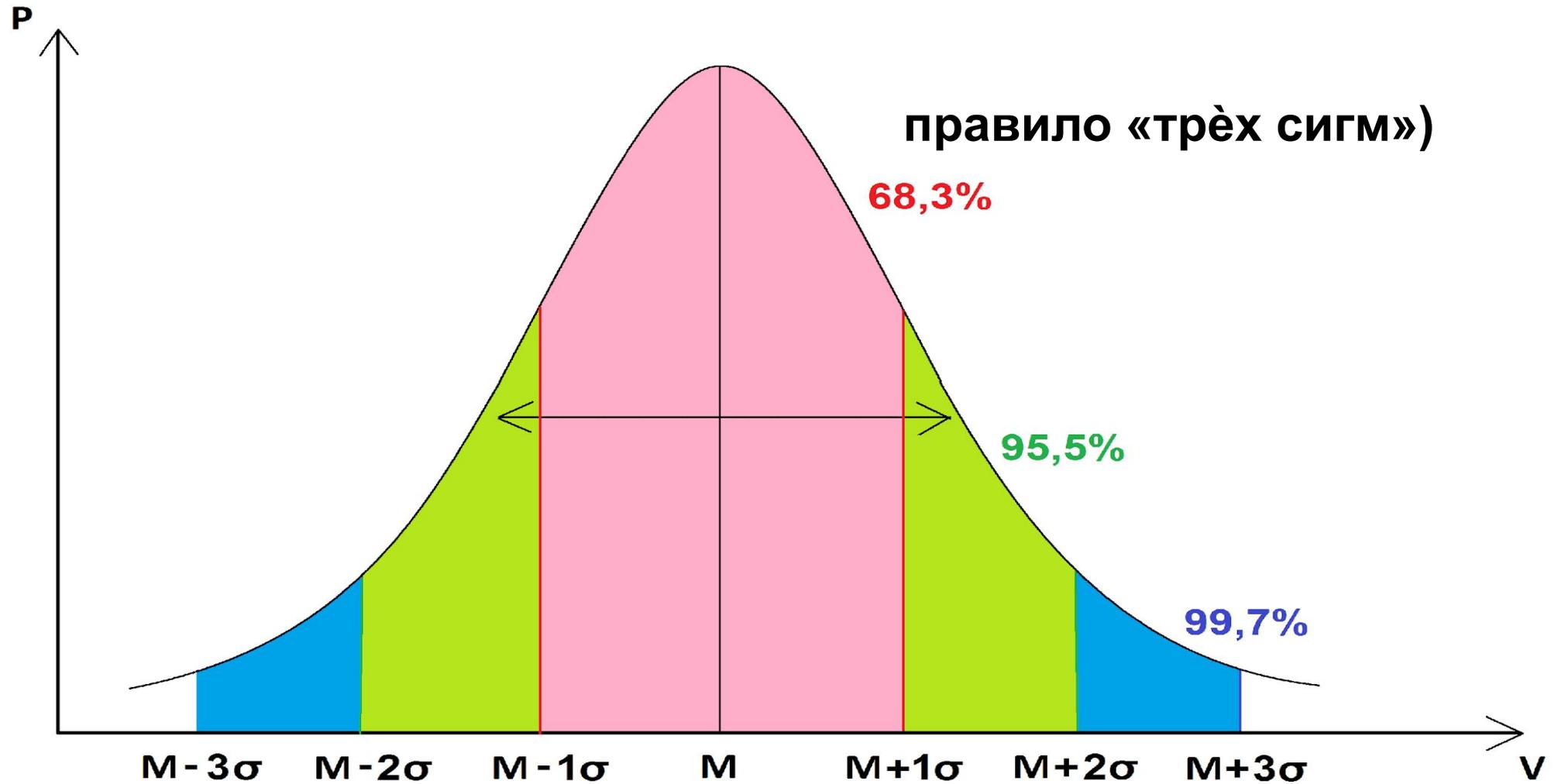
# Проверка на нормальность:

- Визуальный метод
- График вероятности
- Критерий Колмогорова-Смирнова (K-S).
- Если  $K-S d > 0,0895$ , распределение не соответствует нормальному на уровне значимости 0,05.
- Критерий Шапиро-Уилка
- Совпадение значений моды и медианы
- Равенство нулю показателей асимметрии и эксцесса (таблицы критических значений для заданной выборки)

# Нормальное распределение - фундаментальный закон природы

Мировая практика: **20-25% !!!**

# Закон нормального распределения вариационного ряда



# Проверка на нормальность с помощью графика вероятности



# Распределения отличное от нормального

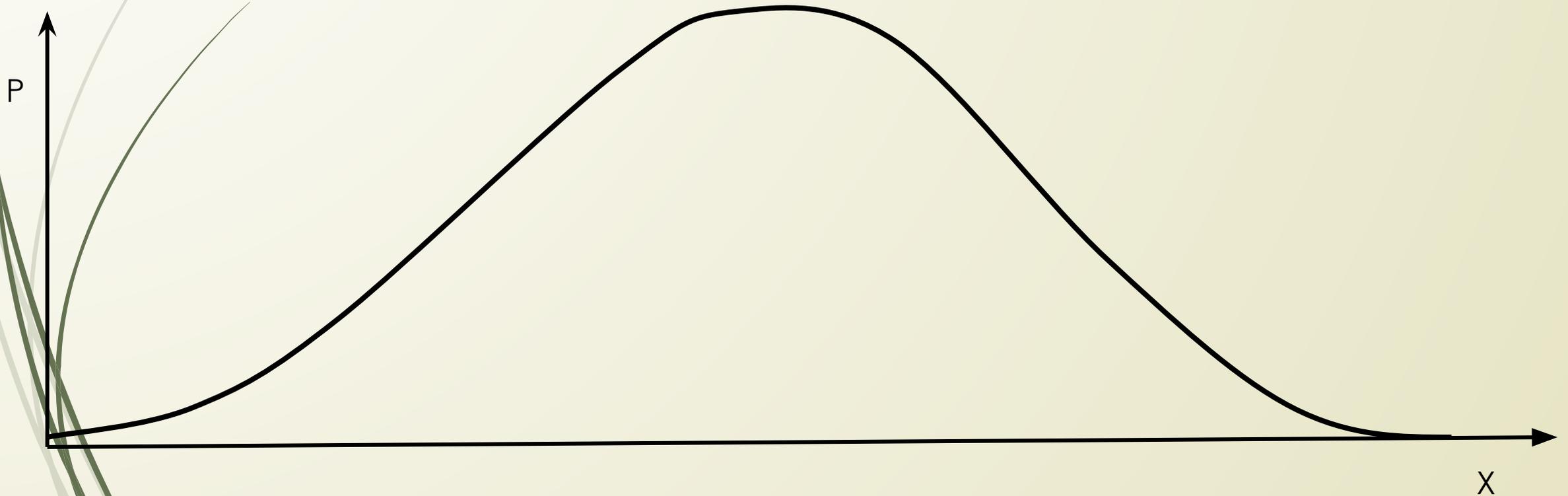


# Распределения отличное от нормального



Гистограмма медицинского показателя **до** и **после** лечения (пример отсутствия нормального распределения)

**Нормальное (гауссово, симметричное, колоколообразное) распределение – используем параметрические методы**



**В противном случае должны  
применяться  
непараметрические методы.**

# Непараметрические методы

- это методы статистической обработки данных, применение которых не требует знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров.

| <b>ПРИЗНАК</b>  | <b>ИССЛЕДОВАНИЕ</b>                                |                                     |                                  |   |   |
|---|--|-------------------------------------|----------------------------------|---|---|
|   | <b>Две группы</b>                                  | <b>Более двух групп</b>             | <b>Группа до и после лечения</b> | <b>Одна группа несколько видов лечения</b>      | <b>Связь признаков</b>  |
| <b>Количественный<br/>(нормальное распределение*)</b> | <b>Критерий Стьюдента</b>                          | <b>Дисперсионный анализ</b>         | <b>Парный критерий Стьюдента</b> | <b>Дисперсионный анализ повторных измерений</b> | <b>Линейная регрессия, корреляция, или метод Блэнда-Алтмана</b> |
| <b>Качественный</b>                                   | <b>Критерий <math>\chi^2</math><br/>Z-критерий</b> | <b>Критерий <math>\chi^2</math></b> | <b>Критерий Мак-Нимара</b>       | <b>Критерий Кокрена</b>                         | <b>Коэффициент сопряженности</b>                                |
| <b>Порядковый</b>                                     | <b>Критерий Манна Уитни</b>                        | <b>Критерий Крускала Уоллиса</b>    | <b>Критерий Уилкоксона</b>       | <b>Критерий Фридмана</b>                        | <b>Коэффициент ранговой корреляции Спирмена</b>                 |

## Принцип действия критериев:

- Сравниваются нужные признаки или группы.
- Проверяется нулевая гипотеза. Находится фактическая вероятность ошибки **ОТКЛОНИТЬ верную нулевую гипотезу ( $p$ )**.

Говоря упрощенно,  $p$  – это вероятность справедливости нулевой гипотезы.

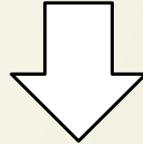
## Принцип действия критериев:

- Максимальную приемлемую вероятность отвергнуть нулевую гипотезу называют уровнем значимости и обозначают  $\alpha$ .
- Обычно в медико-биологических исследованиях принимают  $\alpha = 0.05$ .
- Если  $p < 0,05$  – нулевая гипотеза отвергается, следовательно найдено статистически значимое различие в сравниваемых группах.

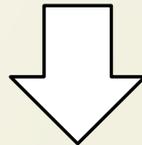
# Сравнение различных совокупностей

«Нулевая гипотеза»

Предположение о том, что в сравниваемых группах отсутствуют различия в распределении частот



расчет «ожидаемых» величин (результатов)



сравнение «ожидаемых» величин и фактических

разницы нет

разница есть

Статистической значимости различий нет

Статистическая значимость различий есть

$p > 0,05$

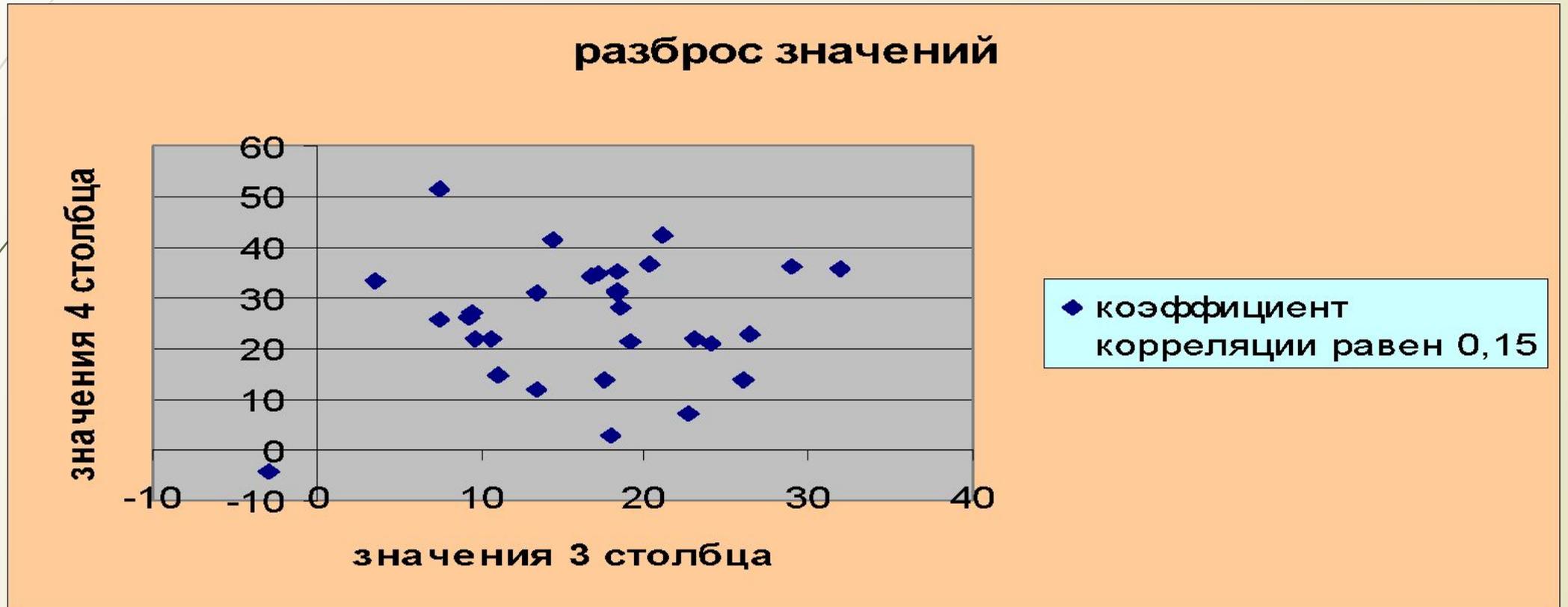
$p < 0,05$

# Корреляция и причинная связь

Когда проверка гипотезы показывает, что существует значимая линейная связь между переменными, исследователь должен рассмотреть возможные виды связи между переменными и выбрать ту, которая диктуется логикой исследования.



# Пример разброса значений двух величин



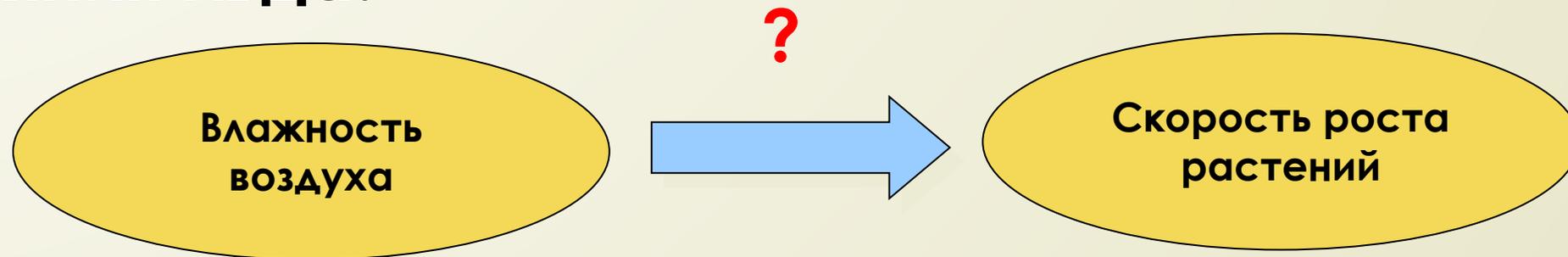
# **Пять видов связи между переменными**

- 1. Прямая причинно-следственная связь**
- 2. Обратная причинно-следственная связь**
- 3. Связь вызвана третьей (скрытой) переменной**
- 4. Взаимосвязь вызвана несколькими скрытыми переменными**
- 5. Связи нет, наблюдаемая зависимость случайна**

# 1. Прямая связь

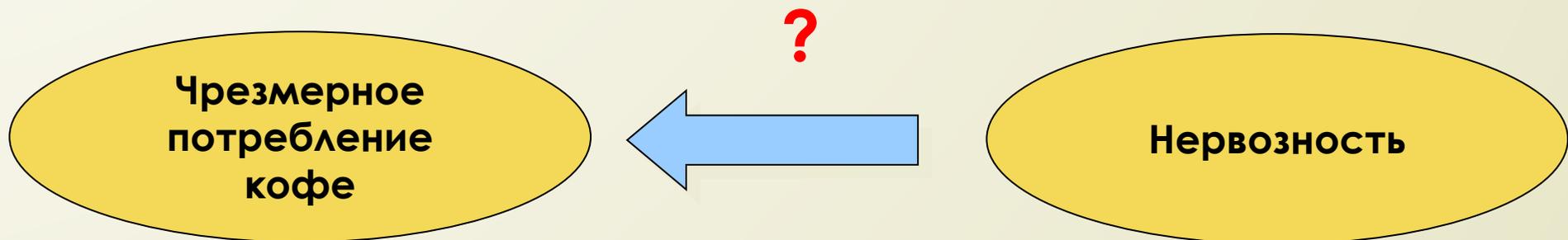
**Прямая причинно-следственная связь между переменными (переменная  $x$  определяет значение переменной  $y$ ).**

- Наличие воды ускоряет рост растений.**
- Яд вызывает смерть.**
- Температура воздуха прямо влияет на скорость таяния льда.**



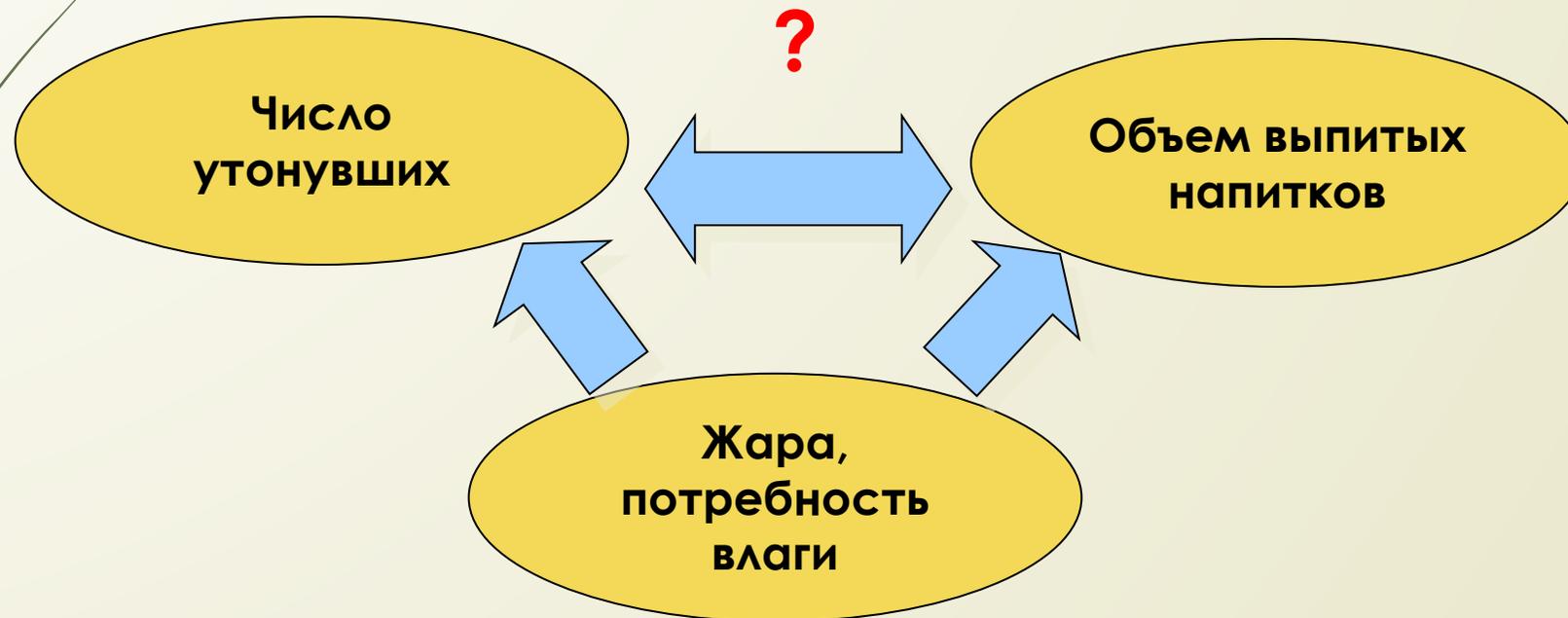
## 2. Обратная связь

- Обратная причинно-следственная связь между переменными (переменная  $y$  определяет значение переменной  $x$ ).
- Исследователь может думать, что чрезмерное потребление кофе вызывает нервозность. Но, может быть, очень нервный человек выпивает кофе, чтобы успокоить свои нервы?



### 3. Связь определена третьей переменной

- Исследователь установил, что существует некая зависимость между числом утонувших людей и числом выпитых безалкогольных напитков в летнее время. А может быть, обе переменные связаны с жарой и потребностью людей во влаге?



## 4. Несколько переменных

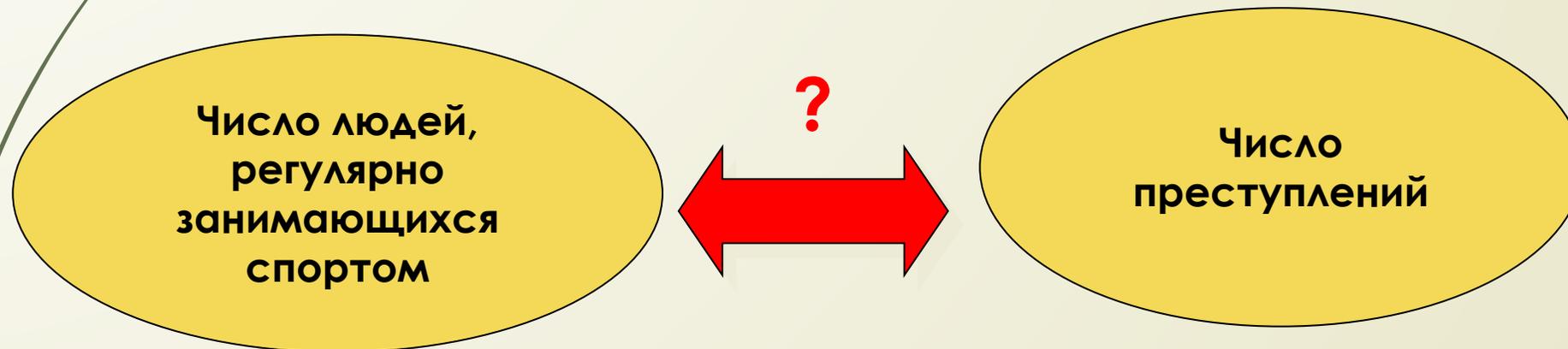
Исследователь может обнаружить значимую связь между оценками студентов в университете и оценками в школе.

Но, возможно, действуют и другие переменные: IQ, количество часов занятий, влияние родителей, мотивация, возраст, авторитет преподавателей.



## 5. Зависимость случайна

- Исследователь может найти значимую зависимость между увеличением количества людей, которые занимаются спортом и увеличением количества людей, которые совершают преступления.
- Но здравый смысл говорит, что любая связь между этими двумя переменными должна быть случайной.



# Диаграммы

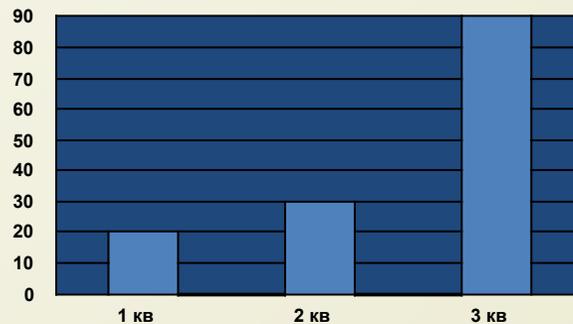
Диаграмма – графическое изображение статистических величин с помощью различных геометрических фигур, знаков

## I. Классификация диаграмм по форме:

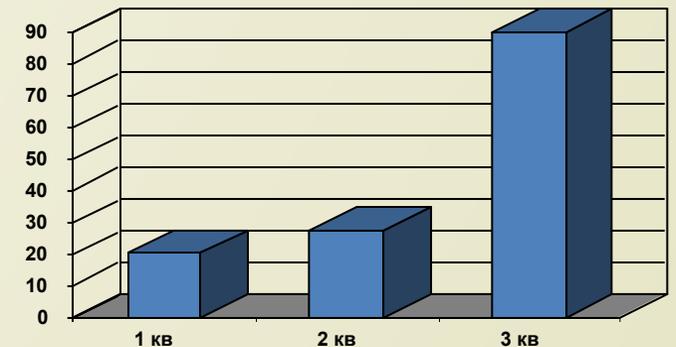
ЛИНЕЙНЫЕ



ПЛОСКОСТНЫЕ



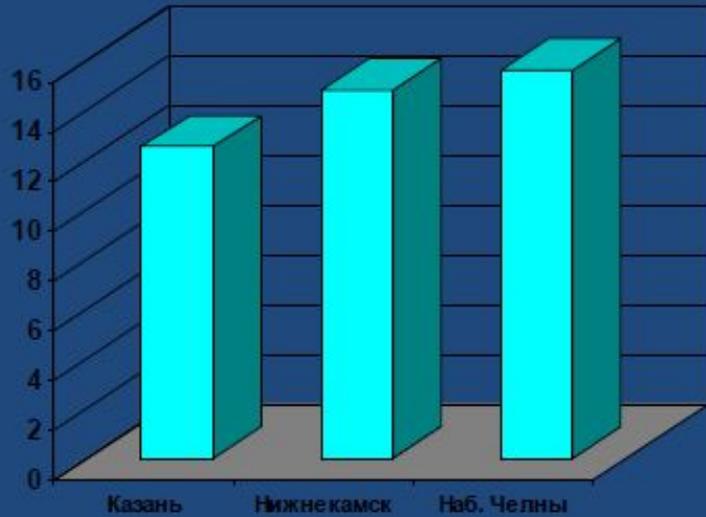
ОБЪЕМНЫЕ



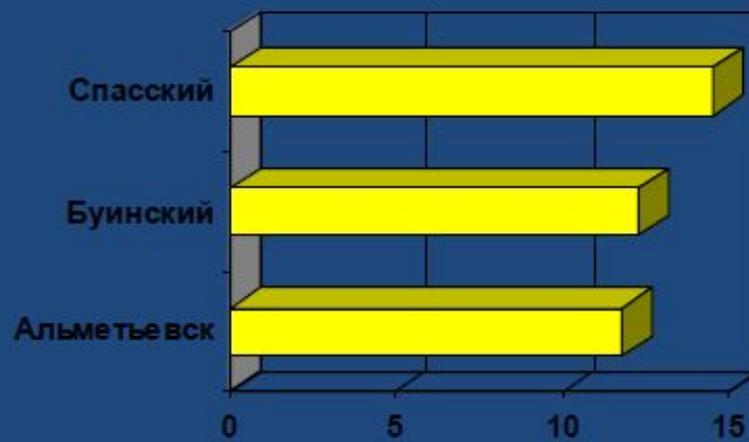
# II. Классификация диаграмм по назначению:

## 1. Диаграммы сравнения

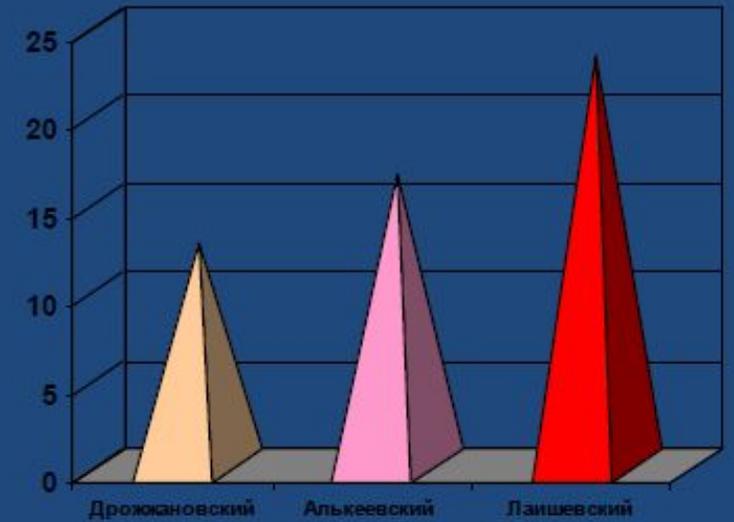
**СТОЛБИКОВЫЕ**



**ЛЕНТОЧНЫЕ**



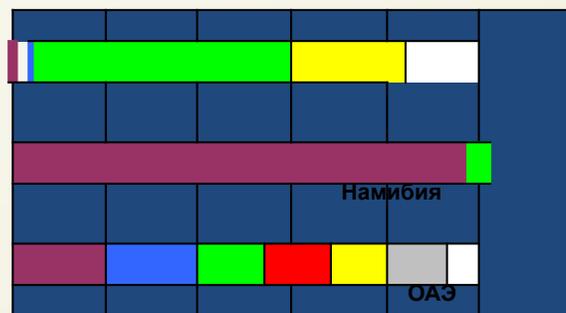
**ФИГУРНЫЕ**



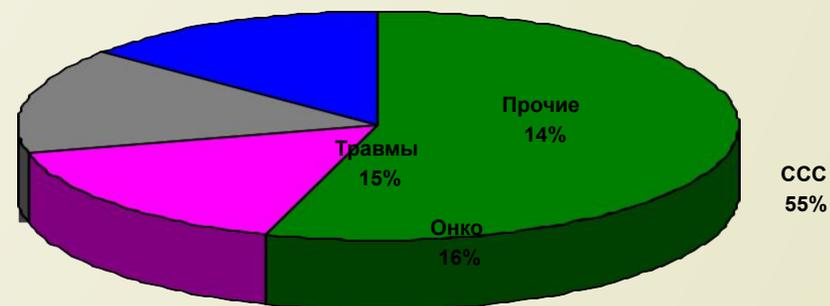
## II. Классификация диаграмм по назначению:

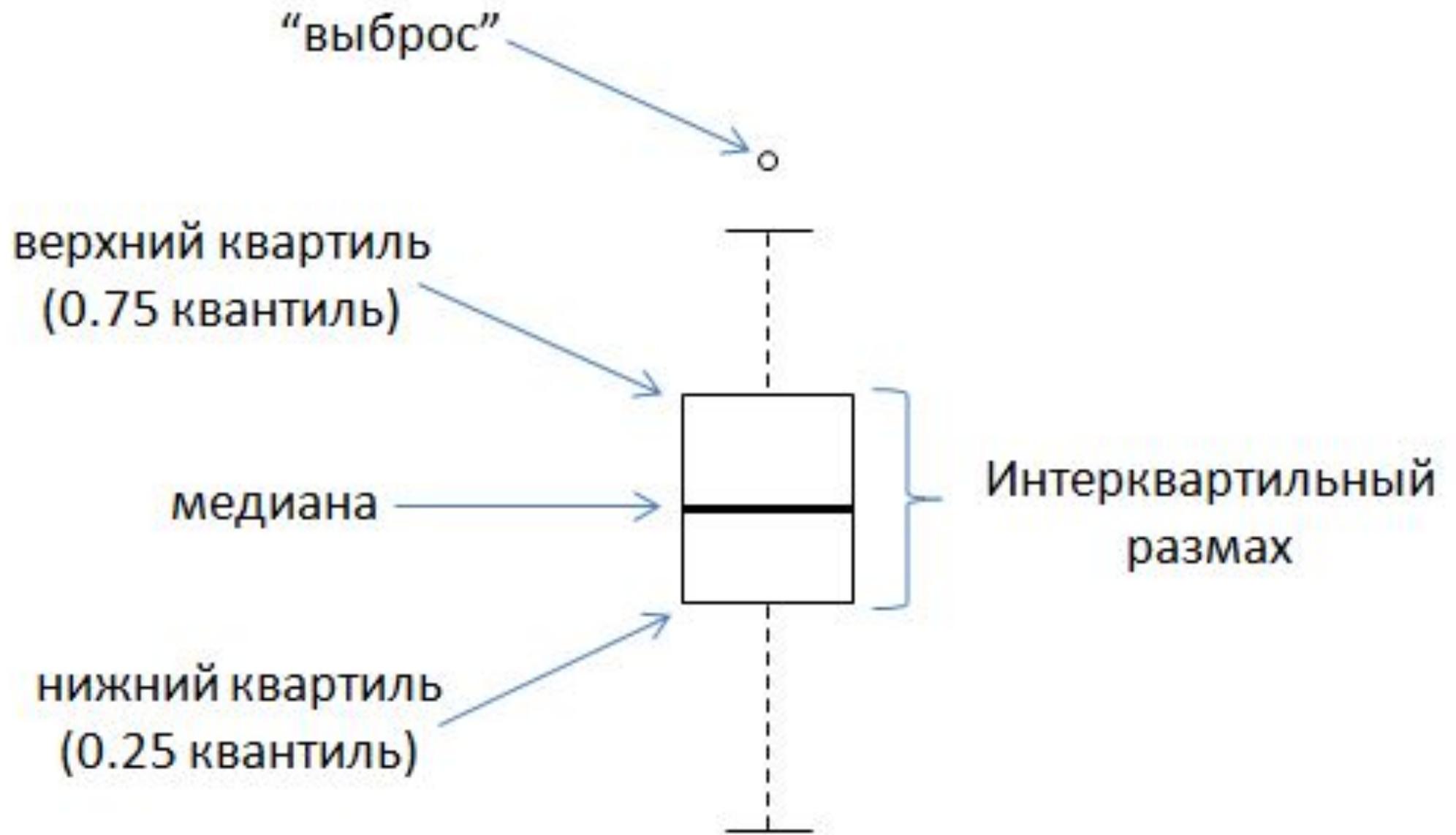
### 2. Диаграммы структуры

ВНУТРИ-ЛЕНТОЧНЫЕ

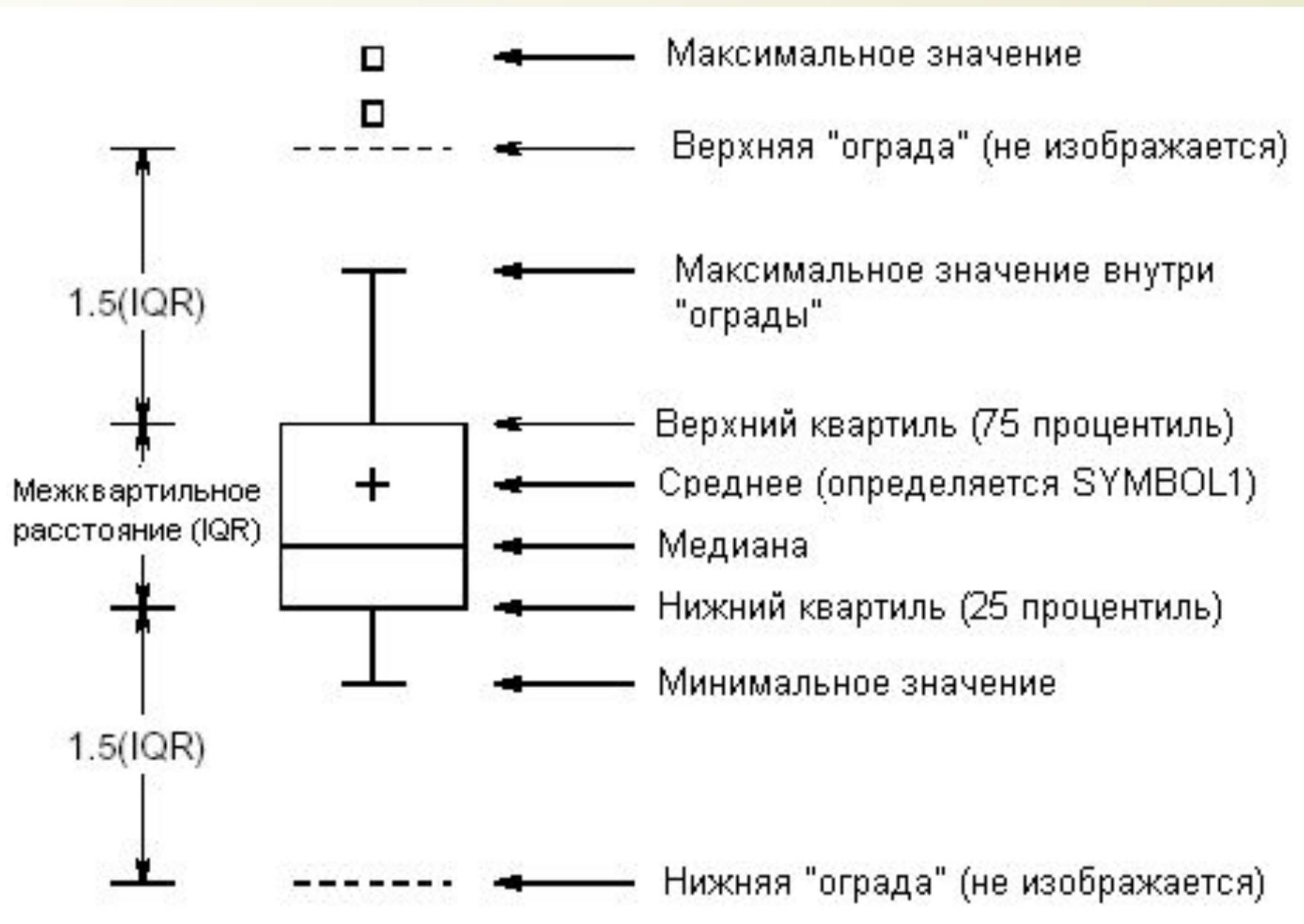


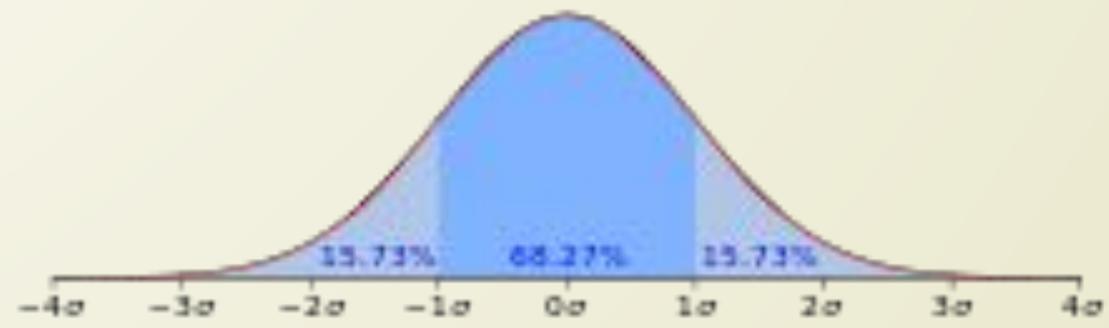
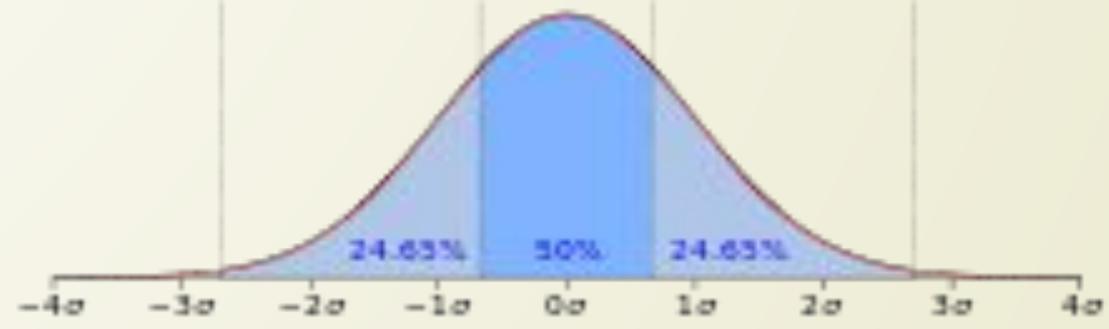
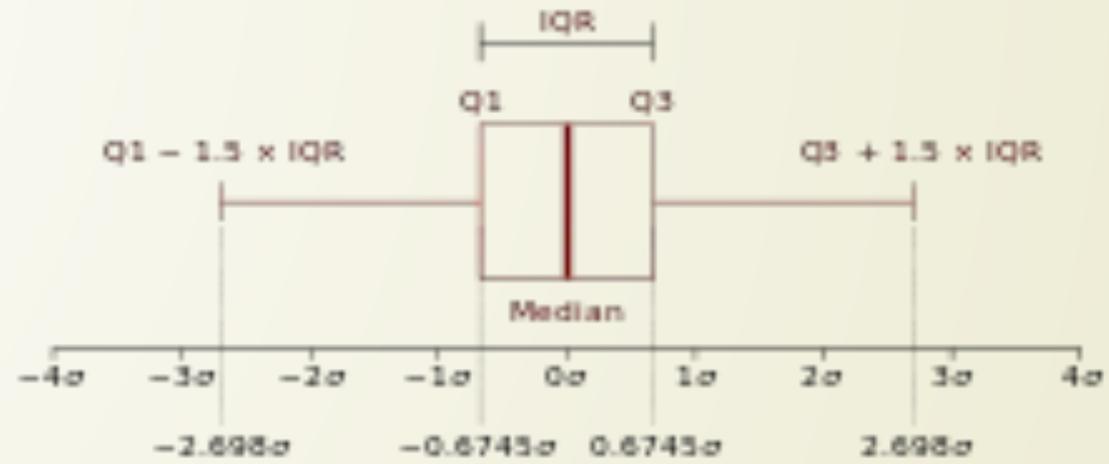
СЕКТОРНЫЕ









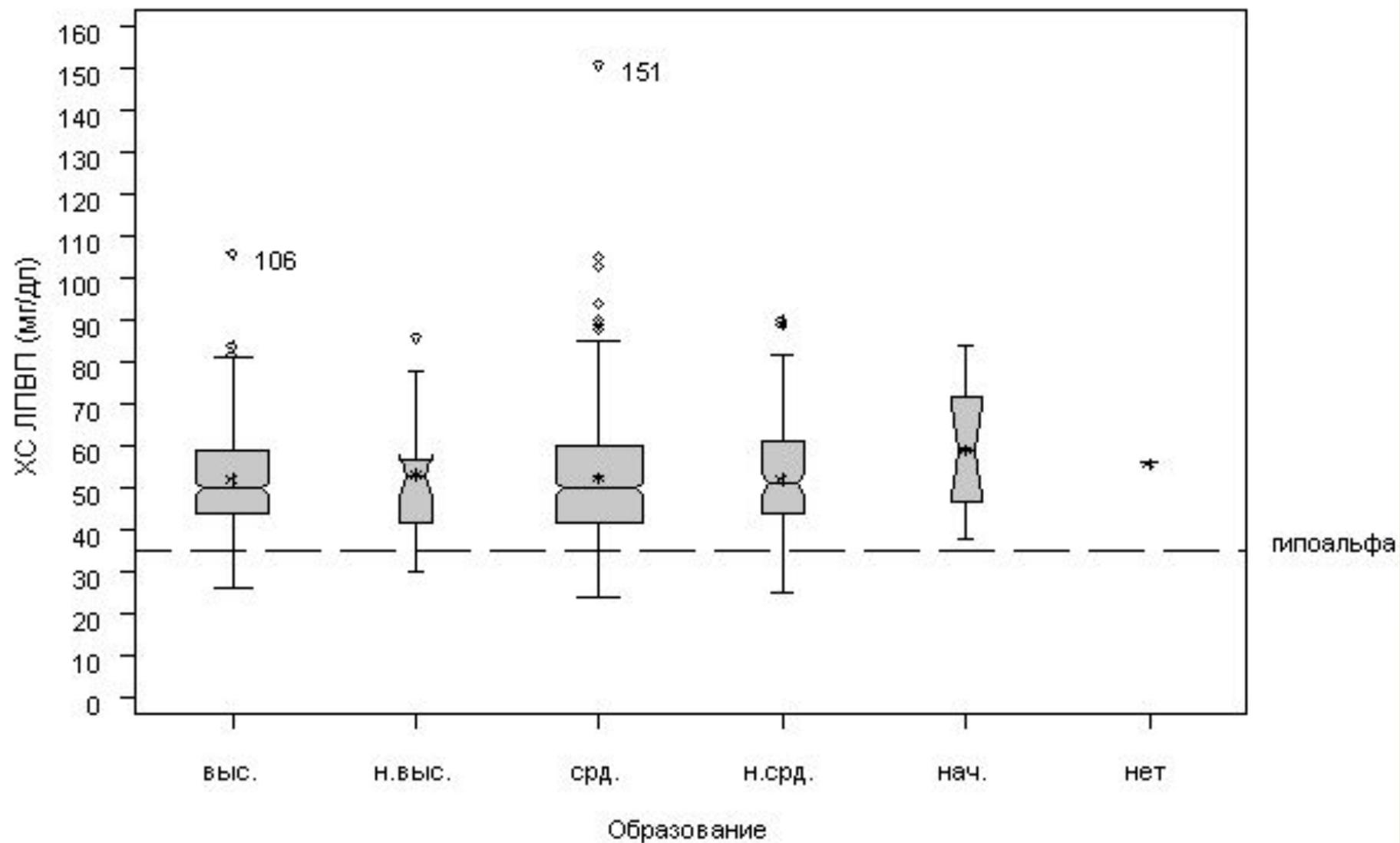




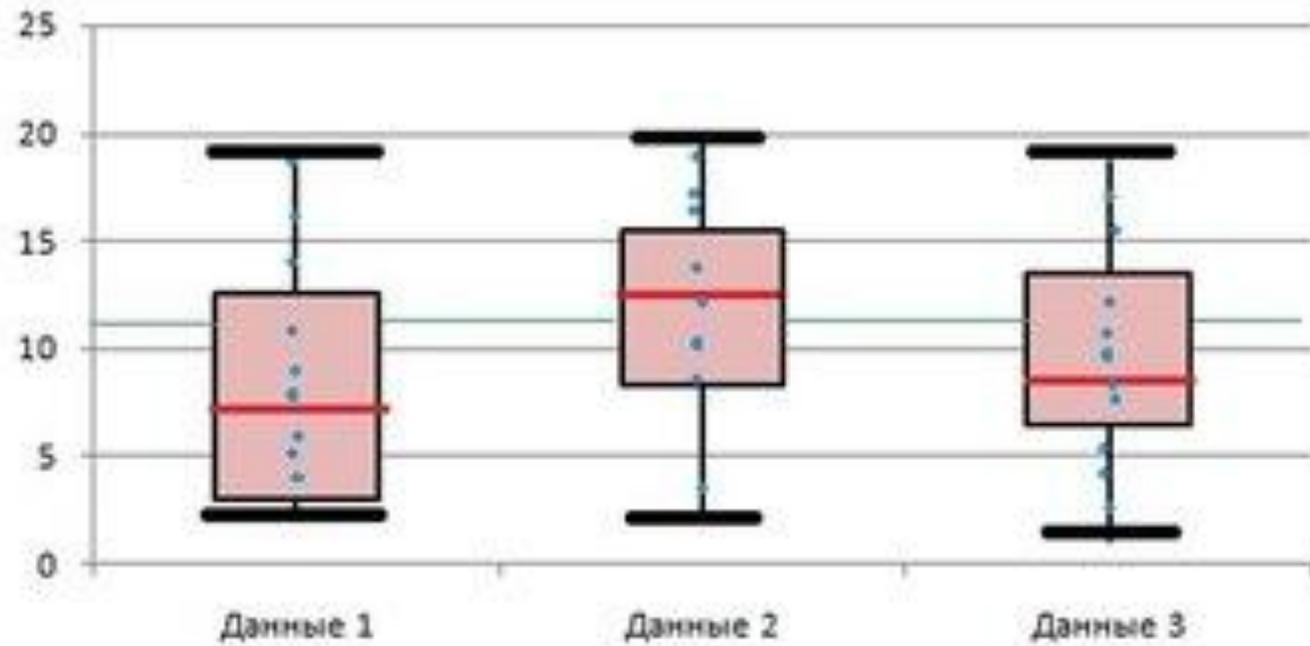
Группа А

Группа В

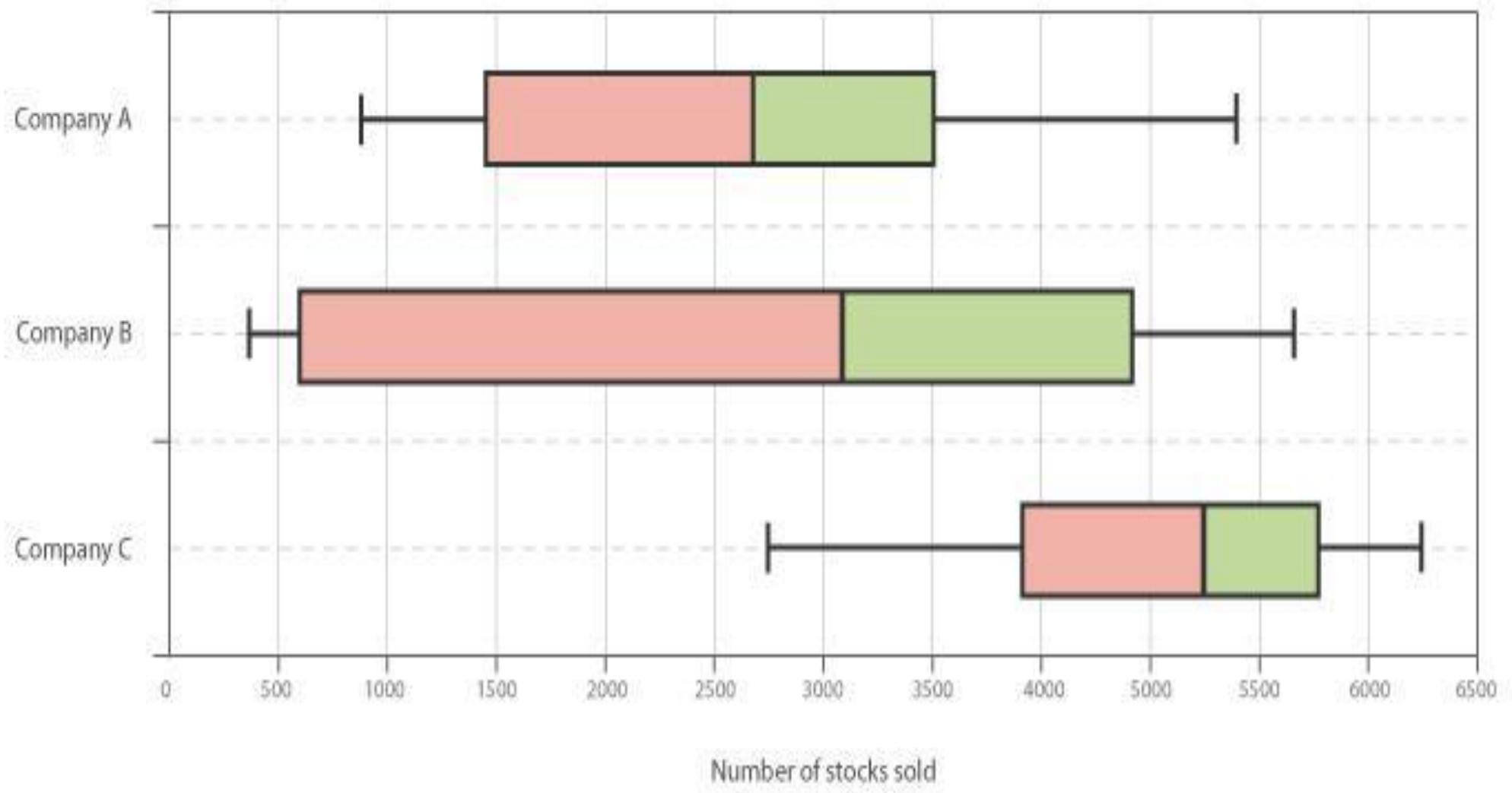
## Уровень ХС ЛПВП в разных образовательных группах



## Ящик с усами



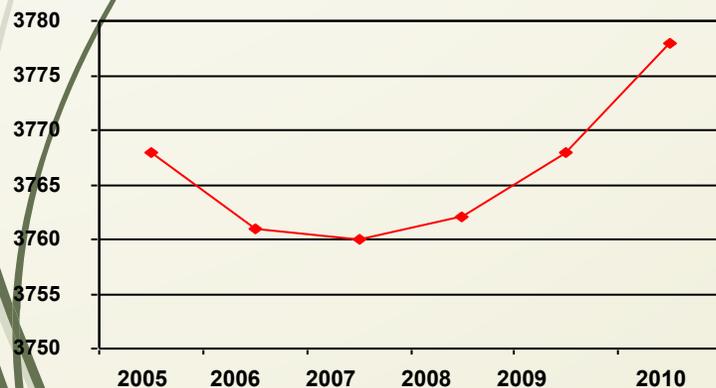
1. Добавить все значения для каждого измерения
2. Добавить "усы"
3. Добавить медиану
4. Добавить среднее значение



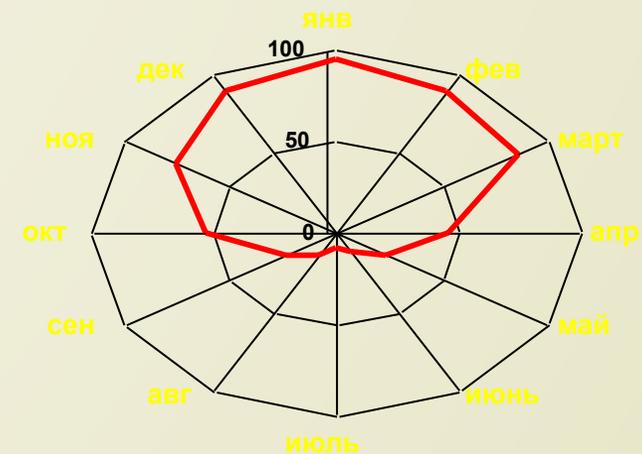
## II. Классификация диаграмм по назначению:

### 3. Диаграммы динамики

#### ЛИНЕЙНЫЕ



#### РАДИАЛЬНАЯ



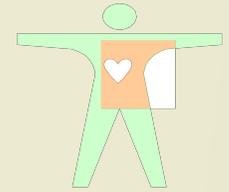
Финансирование медицинской помощи  
по Программе государственных гарантий  
в 1999-2005 гг. (млн.руб.)



Число пролеченных в круглосуточных стационарах  
и стационарах дневного пребывания  
в Красноярском крае в 2002-2005 гг.

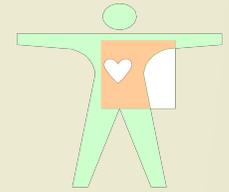


# Структура первичной заболеваемости взрослых в Красноярском крае в 2005 г. (%)



Среди впервые выявленных заболеваний у взрослых в 2005 г. лидируют болезни органов дыхания, а так же травмы и отравления.

# Структура первичной заболеваемости взрослых в Красноярском крае в 2005 г. (%)



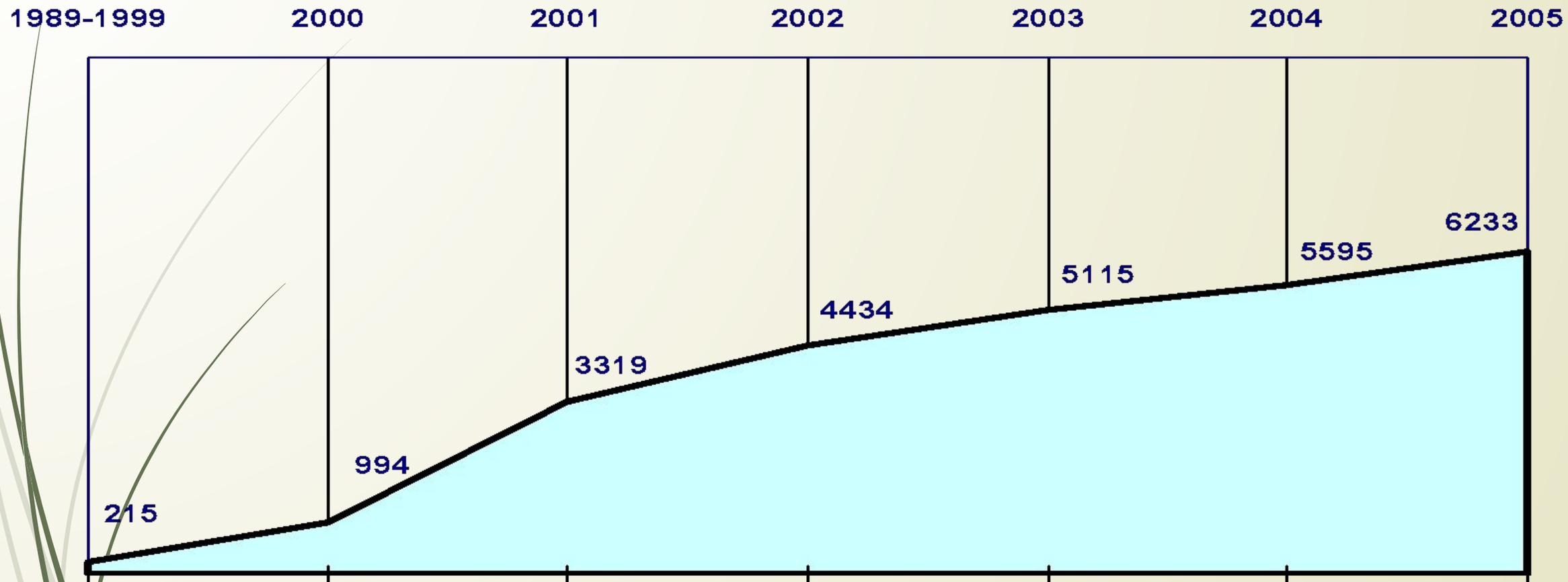
Среди впервые выявленных заболеваний у взрослых в 2005 г. лидируют болезни органов дыхания, а так же травмы и отравления.

**Исполнение программы государственных гарантий обеспечения граждан Красноярского края бесплатной медицинской помощью в 2005 году (%).**



Значительное снижение подушевого норматива допущено в г. Боготоле, Норильске, Ачинске, Лесосибирске, Красноярске, районах – Абанском, Дзержинском, Иланском, Нижне-Ингашском, Северо-Енисейском, Краснотуранском.

## Кумулятивное число зарегистрированных случаев ВИЧ-инфекции в Красноярском крае в 1989-2005 г.г.



По уровню пораженности Красноярский край относится к субъектам Российской Федерации с высоким уровнем развития эпидемии (от 151 до 300 ВИЧ-инфицированных на 100 тыс. населения) и находится на четвертом месте среди регионов Сибирского федерального округа после Иркутской (805,8), Кемеровской (269,8) областей и Республики Бурятия (244,60).

Рождаемость и смертность населения Красноярского края  
в 1990-2005 гг. (‰)



Соотношение числа родившихся к числу умерших в 2005 году составило 1 : 1,45 (по РФ в 2004 г. – 1 : 1,37).

# Правила построения диаграмм

- 1. Должно быть четкое, ясное, краткое название, порядковый номер**
- 2. Все элементы диаграммы (фигуры, знаки, окраска, штриховка) должны быть объяснены на самой диаграмме или в ее легенде**

# Правила построения диаграмм

**3. Изображаемые графические величины должны иметь цифровые обозначения на самой диаграмме или в прилагаемой таблице**

**4. Данные на диаграмме должны размещаться от большего к меньшему, слева направо, снизу вверх и по часовой стрелке,**

**5. Элемент «прочие» всегда располагается**

**«... Знание статистики подобно знанию иностранных языков или алгебры: оно может пригодиться в любое время и при любых обстоятельствах»**

**А.Боули, английский экономист-статистик**

**Благодарю за внимание!**

