

# Методы сравнения средних

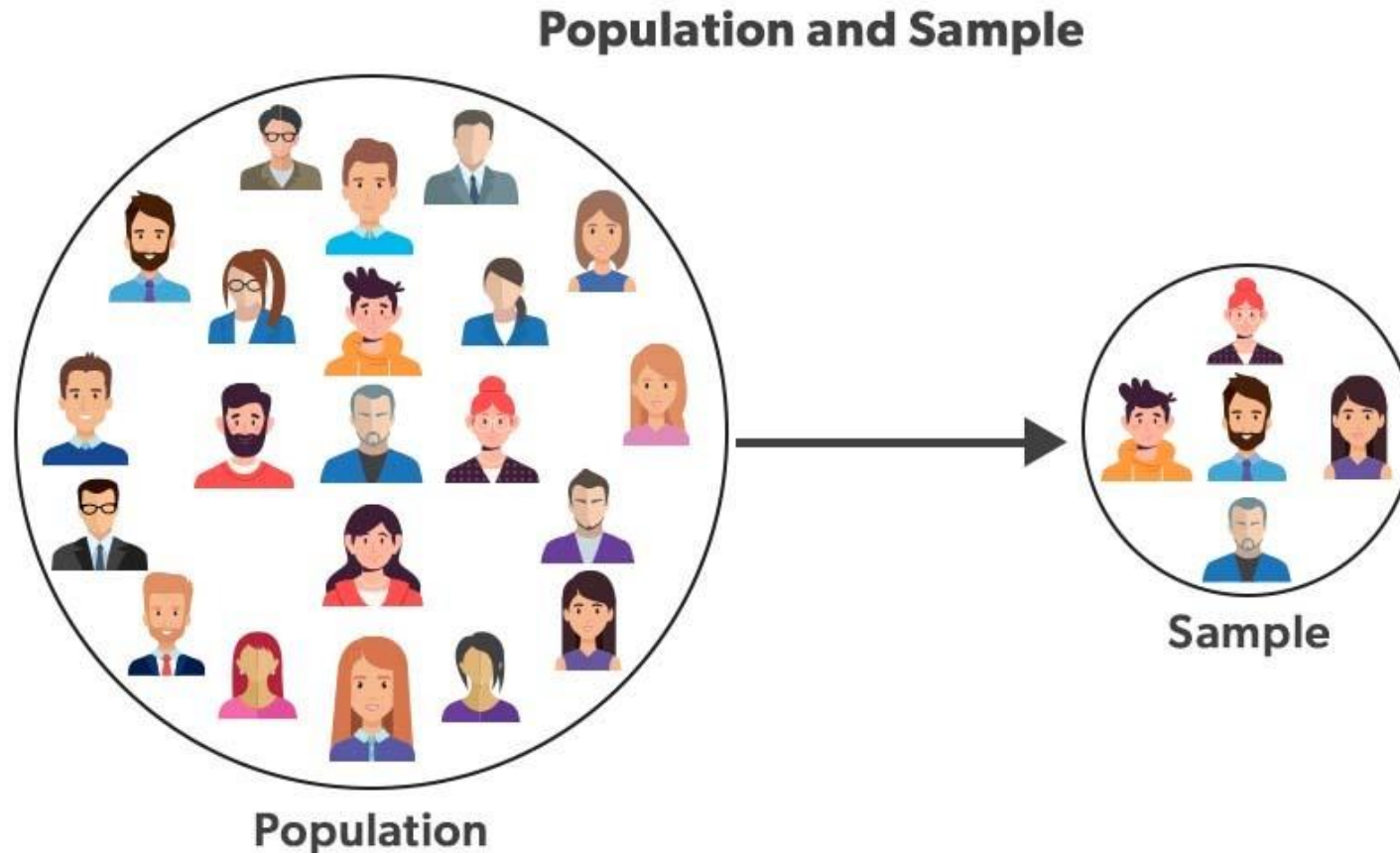
# Статистика

**Описательная статистика** – описание, обобщение и систематизация имеющихся данных в графиках и таблицах (без выводов о генеральной совокупности)

**Статистический вывод** – обобщение данных выборки, чтобы сделать выводы о генеральной совокупности, параметры которой неизвестны

# Основные понятия статистики

## Выборка и генеральная совокупность



# Гипотезы

**Гипотеза** – утверждение о мире, которое может быть принято или опровергнуто в ходе исследования.

**Экспериментальная гипотеза** – гипотеза, которая будет принята или отклонена на основе анализа доказательств, подтверждающих или опровергающих её.

# Экспериментальные гипотезы

- **Нулевая гипотеза ( $H_0$ ):** группы не различаются; переменные не связаны между собой; все обнаруженные различия случайны
- **Альтернативная гипотеза ( $H_1$ ):** группы различаются; связь между переменными существует; различия не случайны

В результате тестирования гипотез мы:

Отклоняем  $H_0$  в пользу  $H_1$ ; принимаем гипотезу  $H_1$

Отклоняем  $H_1$  и не отвергаем  $H_0$  (!)

# Асимметрия статистического вывода

Статистический вывод – это импликация:

- 1) Если значение статистики больше критерия, то  $H_0$  отклоняется
- 2) Если значение статистики НЕ больше критерия, то что?

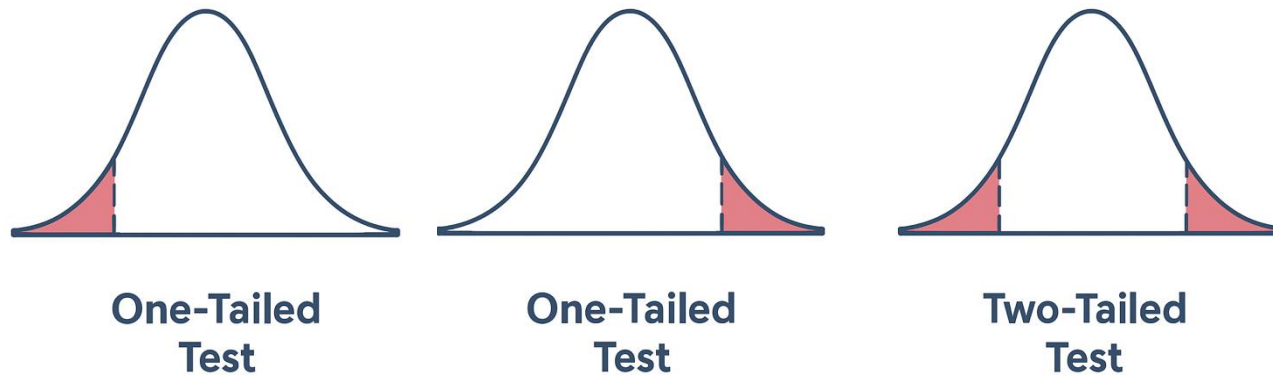
# Асимметрия статистического вывода

Статистический вывод – это импликация:

- 1) Если значение статистики больше критерия, то  $H_0$  отклоняется
- 2) Если значение статистики НЕ больше критерия, то нет оснований отклонить  $H_0$ , но оснований принять  $H_0$  тоже нет!

# Экспериментальные гипотезы

Гипотеза может быть односторонней (например,  $H_1: M_1 > M_2$ ) или двусторонней ( $H_1: M_1 \neq M_2$ ).



# Уровень значимости ( $\alpha$ ) и p-value

- **P-value** – показатель того, насколько мы уверены в полученном результате; «какова вероятность того, что полученный результат – случайность?»
- Чем меньше p-value, тем больше мы уверены, но p-value ничего не говорит о силе связи / величине различий / etc.

# Уровень значимости ( $\alpha$ ) и p-value

- **Уровень значимости** ( $\alpha$ , level of significance) – порог, относительно которого мы считаем результаты неслучайными.

Если **p-value**  $< \alpha$ , то мы считаем результаты значимыми

# Уровень значимости ( $\alpha$ ) и p-value

- **Уровень значимости ( $\alpha$ , level of significance)** – порог, относительно которого мы считаем результаты неслучайными.

Если **p-value <  $\alpha$** , то мы считаем результаты значимыми

В зависимости от того, насколько высока цена ошибки (насколько опасно ошибиться), мы можем повышать или понижать порог значимости. Конвенционально, используется порог:  $\alpha = 0.05$ . То есть значимыми считаются результаты при p-value < 0.05.

# Уровень значимости ( $\alpha$ ) и p-value

| $\alpha$ | p-value  | Значимо? |
|----------|----------|----------|
| 0.05     | 0.034    |          |
| 0.01     | 0.028    |          |
| 0.001    | 0.000577 |          |
| 0.1      | 0.002    |          |
| 0.05     | 0.048    |          |
| 0.05     | 0.05     |          |

# Уровень значимости ( $\alpha$ ) и p-value

| $\alpha$ | p-value  | Значимо? | На каком уровне значимо? |
|----------|----------|----------|--------------------------|
| 0.05     | 0.034    | Да       |                          |
| 0.01     | 0.028    | Нет      |                          |
| 0.001    | 0.000577 | Да       |                          |
| 0.1      | 0.002    | Да       |                          |
| 0.01     | 0.048    | Нет      |                          |
| 0.05     | 0.05     | Нет      |                          |

# Нормальное распределение

# Нормальное распределение

**Распределение** – это связь между значением величины и вероятностью того, что величина примет это значение.

**Нормальное распределение** – это особый тип распределения, при котором большинство значений сосредоточено около среднего



@physicsfun

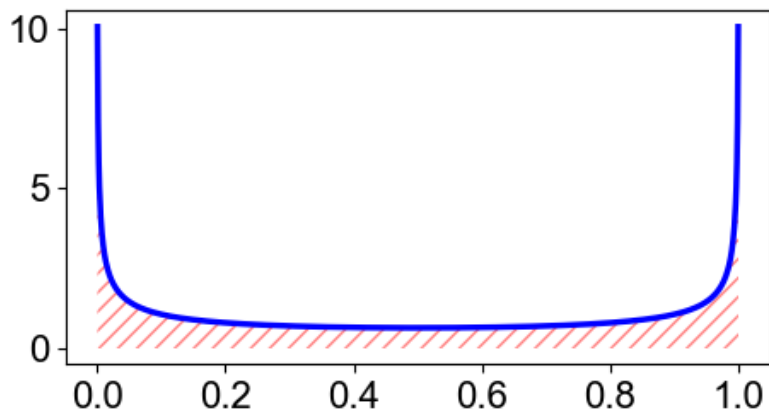
# Нормальное распределение

Параметры нормального распределения: среднее и стандартное отклонение

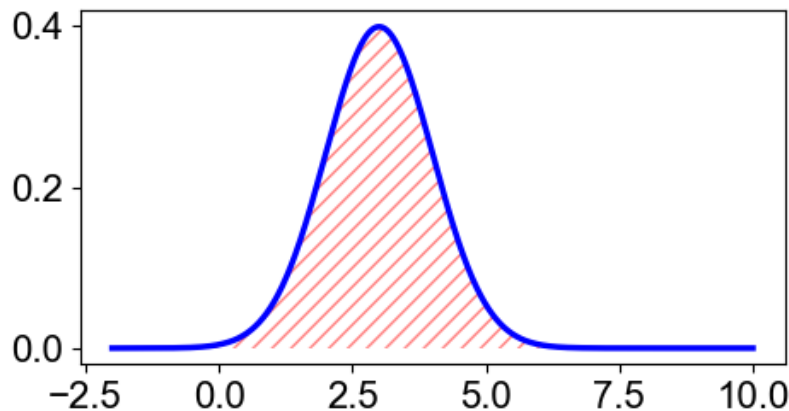
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2}$$

# Другие распределения

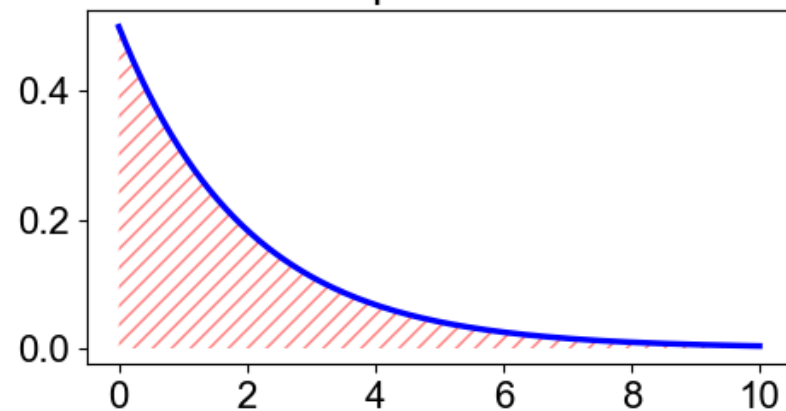
Beta



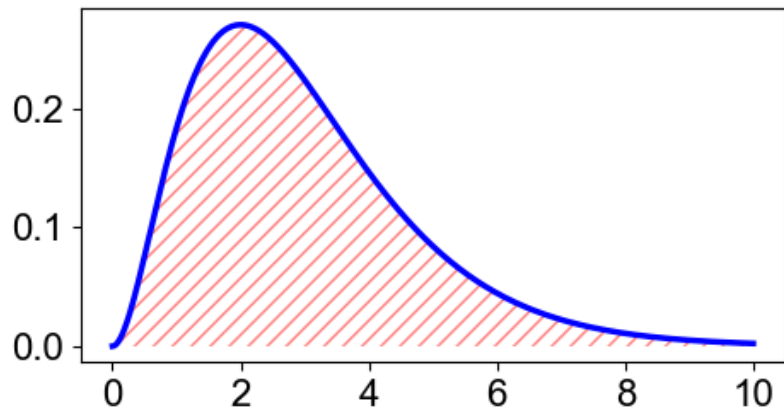
Normal



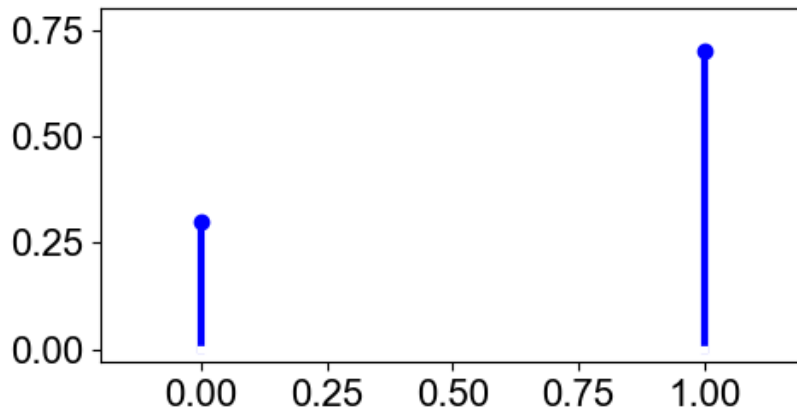
Exponential



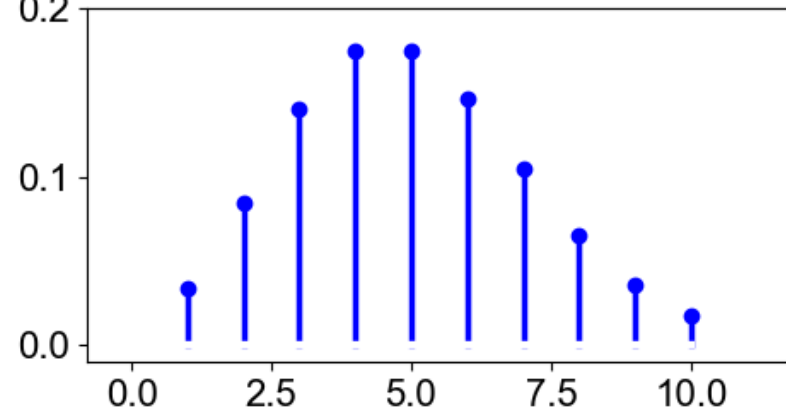
Gamma



Bernoulli



Poisson

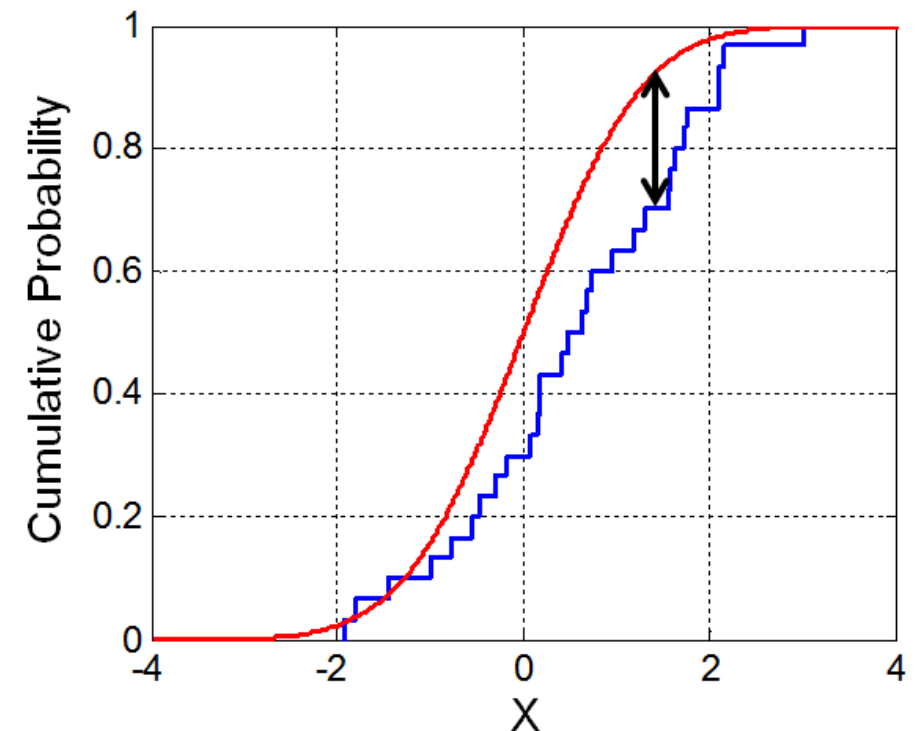


# Проверка распределения на нормальность

Тесты Колмогорова-Смирнова и Шапиро-Уилка позволяют оценить то, насколько эмпирическое распределение отличается от нормального:

**H<sub>0</sub>**: эмпирическое распределение не отличается от нормального

**H<sub>1</sub>**: распределение переменной отличается от нормального



# Проверка распределения на нормальность

**Тест Колмогорова-Смирнова** используется, когда выборка относительно большая ( $> 50$  наблюдений).

**Тест Шапиро-Уилка** используется, когда выборка относительно небольшая ( $< 50$  наблюдений).

```
> shapiro.test(data)
```

```
Shapiro-Wilk normality test
```

```
data: data  
W = 0.9889, p-value = 0.5768
```

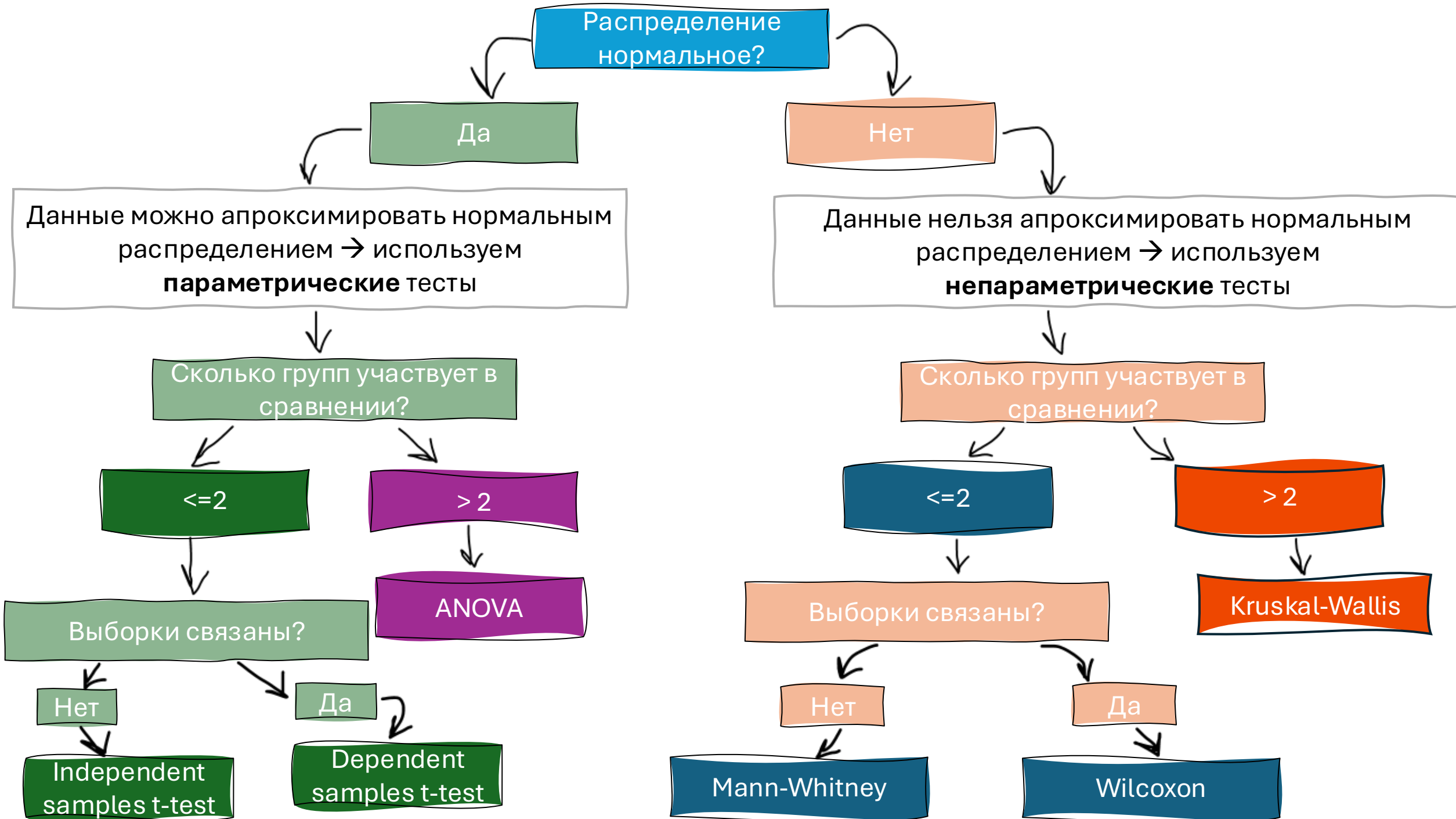
```
> |
```

```
> x = rnorm(50)  
> ks.test(x, "pnorm")
```

```
One-sample Kolmogorov-Smirnov test
```

```
data: x  
D = 0.1698, p-value = 0.0994  
alternative hypothesis: two-sided
```

Сравнение средних



Распределение нормальное?

Да

Данные можно аппроксимировать нормальным распределением → используем **параметрические** тесты

Сколько групп участвует в сравнении?

$\leq 2$

$> 2$

Выборки связаны?

Нет

Independent samples t-test

Да

Dependent samples t-test

ANOVA

Нет

Данные нельзя аппроксимировать нормальным распределением → используем **непараметрические** тесты

Сколько групп участвует в сравнении?

$\leq 2$

$> 2$

Выборки связаны?

Нет

Mann-Whitney

Да

Wilcoxon

Kruskal-Wallis

# Тест Стьюдента (t-тест)

Позволяет проверить гипотезу о равенстве средних значений в двух выборках.

## **Требования к данным:**

- 1) Числовая переменная
- 2) Распределение не отличается от нормального

$H_0$ : Распределение средних в двух группах не отличается

$H_1$ : Средние в группах различны

\*Почему важна нормальность?\*

The diagram illustrates the formula for a two-sample t-test. The formula is 
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
. Annotations with green arrows point to specific parts of the formula: 'Mean of the sample 1' points to  $\bar{X}_1$ , 'Mean of the sample 2' points to  $\bar{X}_2$ , 'Standard deviation sample 1 and 2' points to the  $s_1^2$  and  $s_2^2$  terms, and 'Number of cases sample 1 and 2' points to the  $n_1$  and  $n_2$  terms.

Mean of the sample 1      Mean of the sample 2

$\bar{X}_1 - \bar{X}_2$

$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

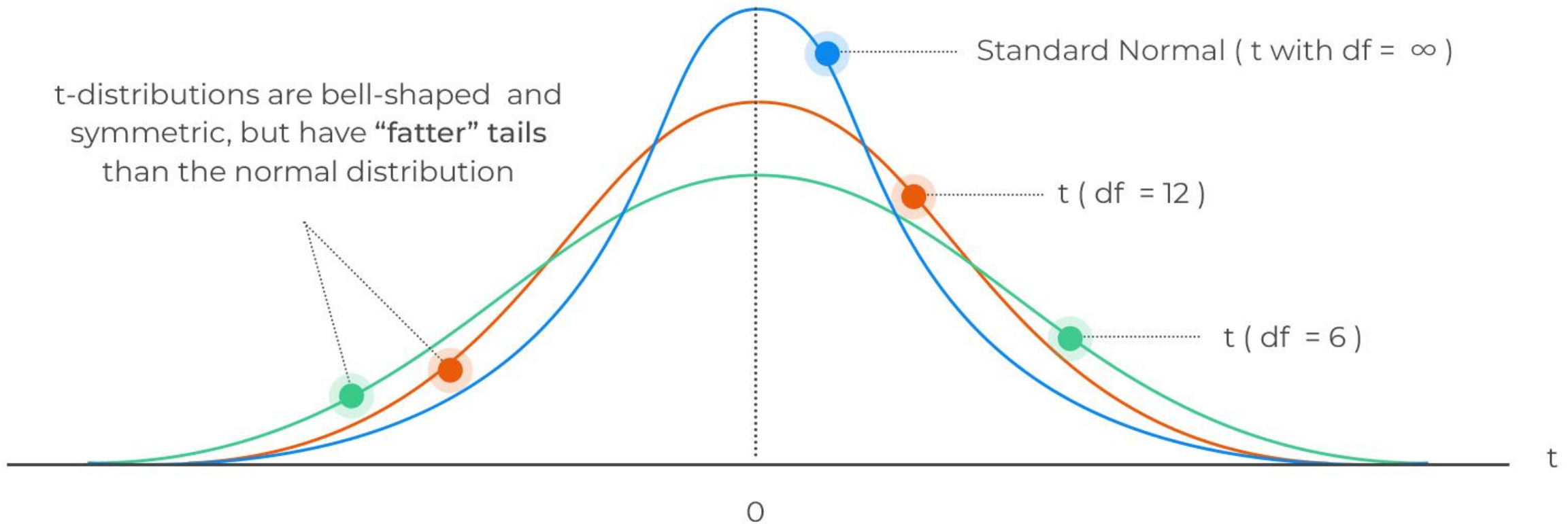
Standard deviation sample 1 and 2

Number of cases sample 1 and 2



# Student's t Distribution vs Normal Distribution

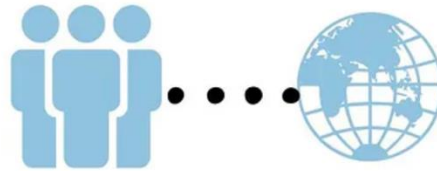
t-distributions are bell-shaped and symmetric, but have “fatter” tails than the normal distribution



# Тест Стьюдента (t-тест)

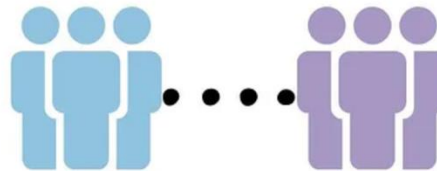
## Types of T-test

One sample  
t-test



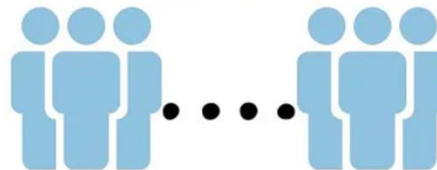
Is there a difference  
between a group and  
the population?

Independent samples  
t-test



Is there a difference  
between two groups?

Paired sample  
t-test



Is there a difference in  
a group between two  
points in time?

# t-test

**t-тест для независимых выборок** используется для сравнения средних в двух независимых(!) группах.

**H0:** Среднее по переменной не отличается между группами

**H1:** Среднее по переменной отличается между группами

*Пример:* средний рост парней и девушек различается.

# t-test

**t-тест для зависимых выборок** используется для сравнения средних, посчитанных на одной группе в разные моменты времени.

**H0:** Среднее по переменной не отличается между двумя моментами времени

**H1:** Среднее по переменной отличается между двумя моментами времени

*Пример:* результаты выполнения заданий по анализу данных до и после подготовки.

# t-test

**Одновыборочный t-тест** используется для сравнения среднего в выборке с некоторым априори заданным средним (например, средним в генеральной совокупности)

**H0:** Среднее по выборке не отличается от заданного значения

**H1:** Среднее по выборке отличается от заданного значения

*Пример:* средний рост студентов ФиКЛ отличается от 175 см.

# Однофакторный дисперсионный анализ (1-way ANOVA)

ANOVA используется для сравнения средних, если групп более двух.

**H<sub>0</sub>**: Средние в группах не различаются

**H<sub>1</sub>**: Средние в группах различаются

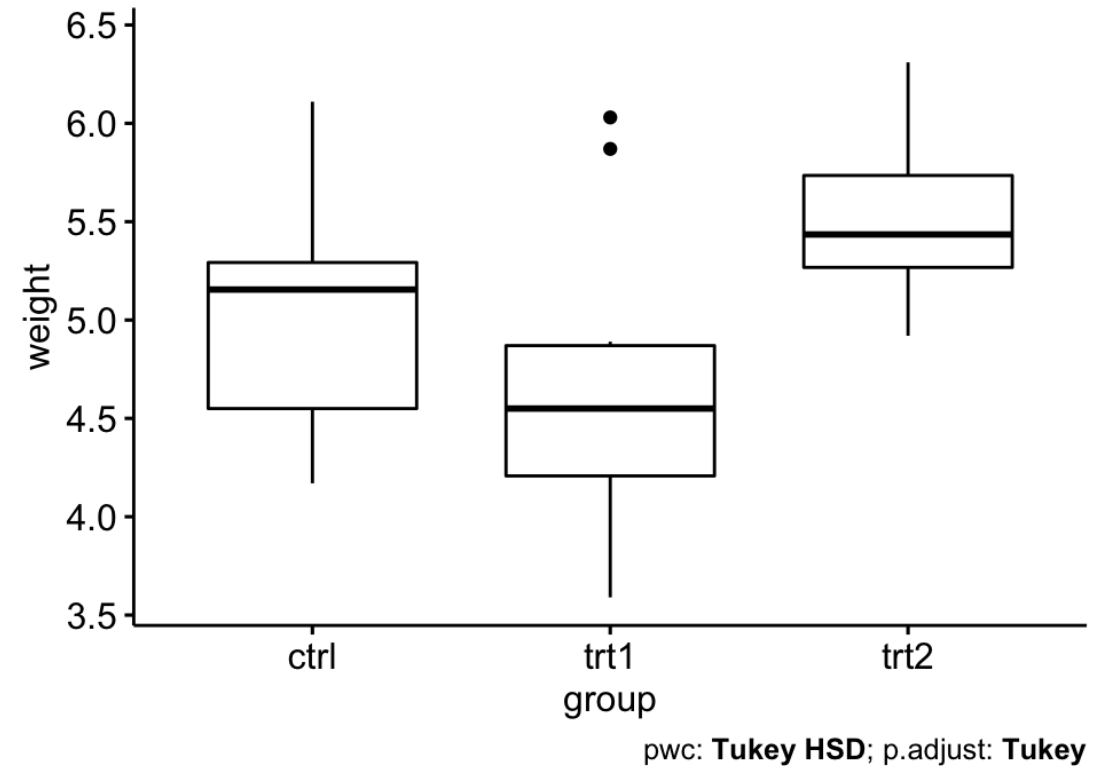
Если ANOVA показывает, что есть различия между группами, мы проводим post-hoc тесты, где попарно сравниваем все группы, чтобы найти те группы, между которыми существует значимое различие

# Допущения ANOVA

1. Нормальность распределения остатков
2. Гомогенность дисперсий (Levene's test)
3. ~одинаковое количество наблюдений в каждой группе

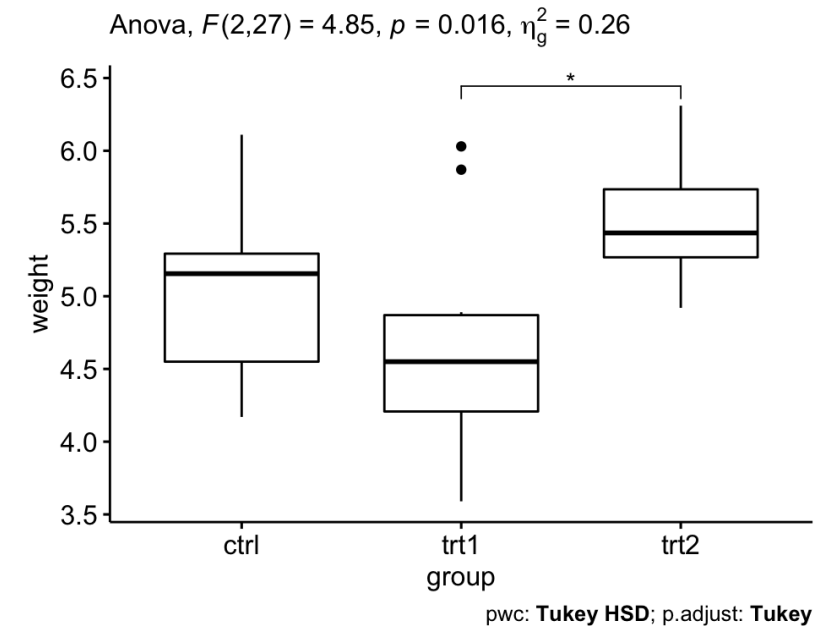
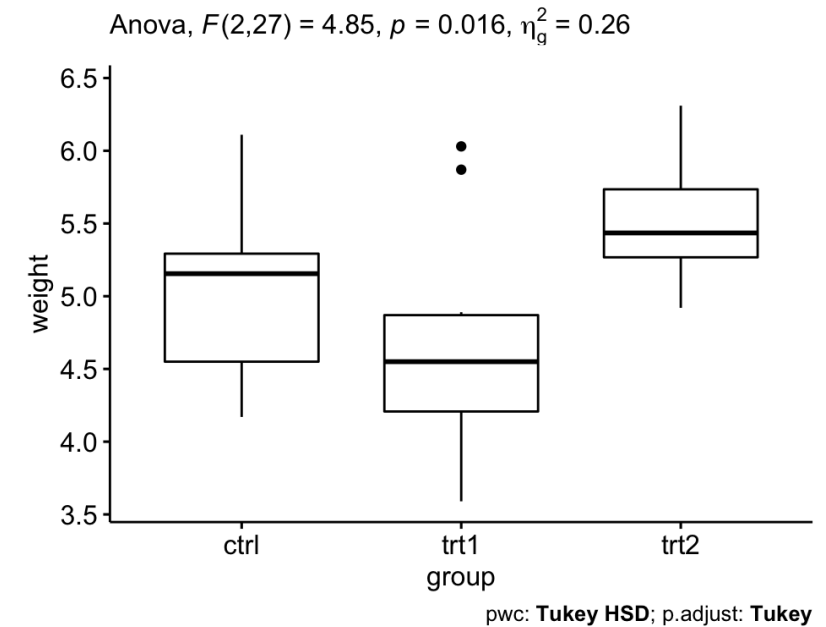
# ANOVA pipeline

- 1) Проверяем допущения
- 2) Проводим тест ANOVA
- 3) Если статистика теста значима, проводим попарные сравнения, чтобы понять, между какими группами есть различия



# ANOVA pipeline

- 1) Проверяем допущения
- 2) Проводим тест ANOVA
- 3) Если статистика теста значима, проводим попарные сравнения, чтобы понять, между какими группами есть различия
- 4) Интерпретируем результаты



```
# Run ANOVA
anova_model <- aov(Score ~ Group, data = data)

# Check ANOVA Results
summary(anova_model)
```

## Output:

|           | Df | Sum Sq | Mean Sq | F value | Pr(>F)       |
|-----------|----|--------|---------|---------|--------------|
| Group     | 2  | 968.9  | 484.4   | 19.1    | 6.77e-06 *** |
| Residuals | 27 | 684.7  | 25.4    |         |              |

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
# Run Tukey's HSD Test
tukey_test <- TukeyHSD(anova_model)

# View Results
print(tukey_test)
```

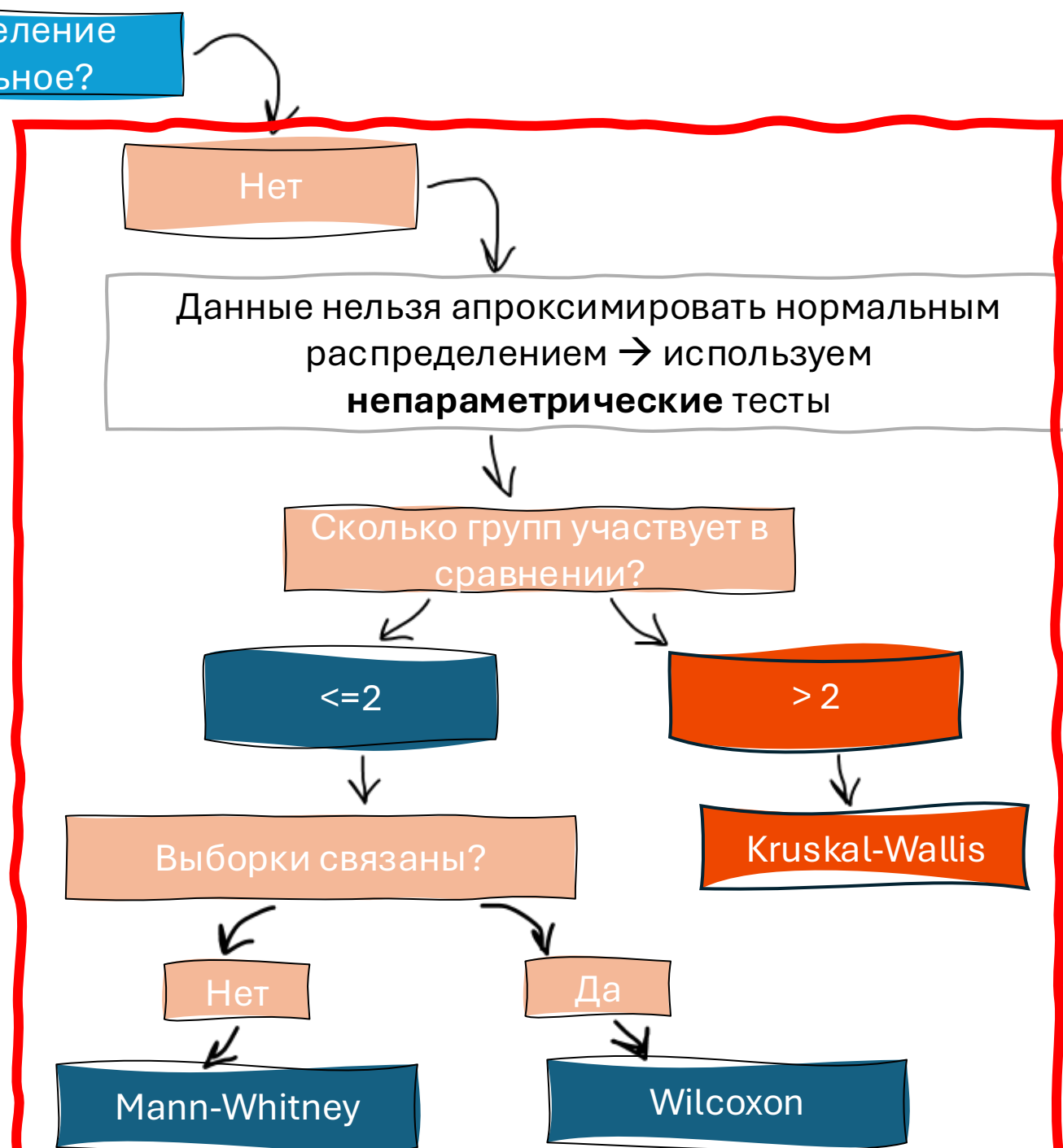
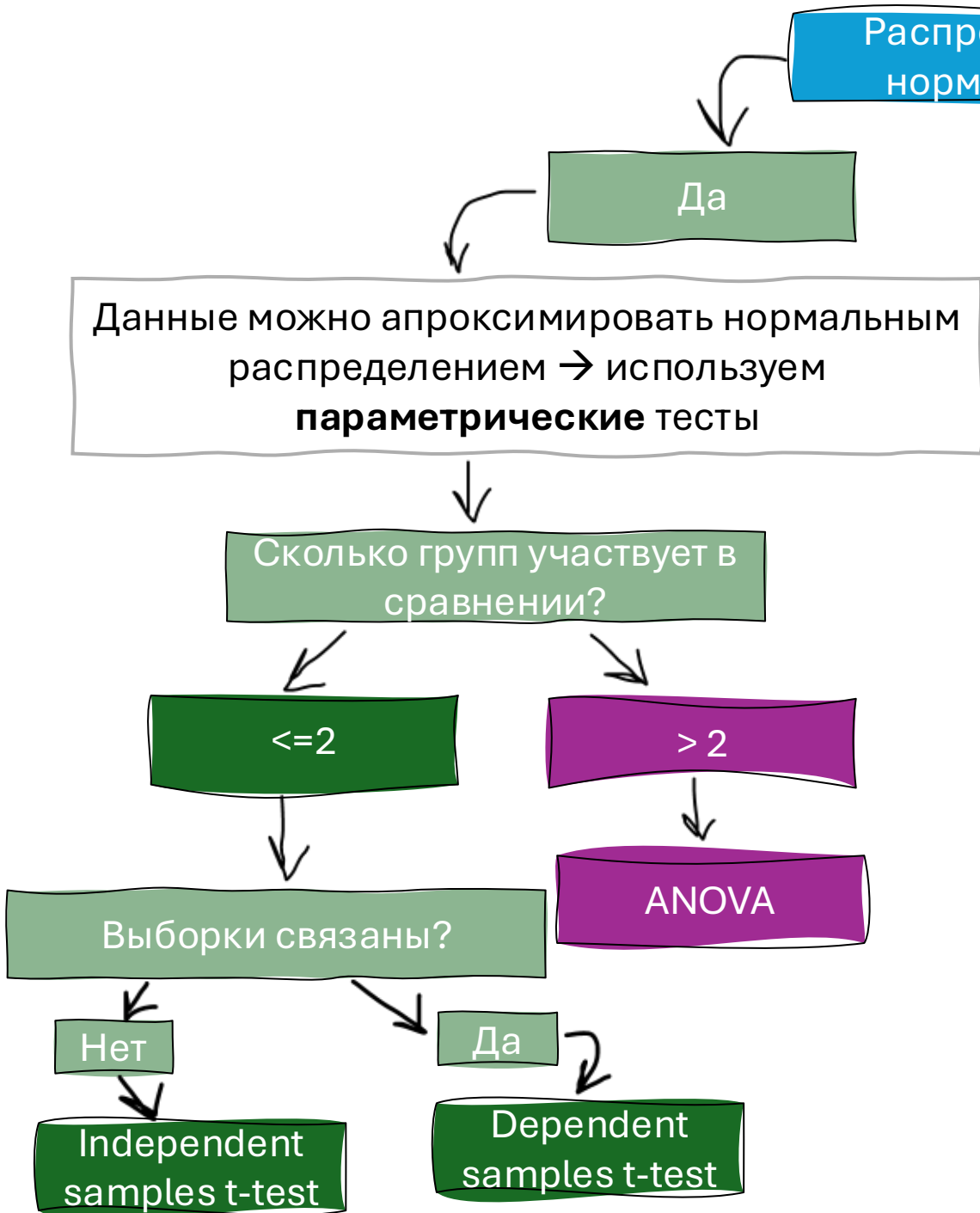
## Output:

Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = Score ~ Group, data = data)

\$Group

|     | diff      | lwr        | upr      | p adj     |
|-----|-----------|------------|----------|-----------|
| B-A | 5.670882  | 0.08719096 | 11.25457 | 0.0459886 |
| C-A | 13.844931 | 8.26123959 | 19.42862 | 0.0000042 |
| C-B | 8.174049  | 2.59035750 | 13.75774 | 0.0032446 |



# Mann-Whitney U-test

Непараметрический аналог t-теста для независимых(!) выборок. Используется, если распределение переменной отлично от нормального

# Wilcoxon signed-rank test

Непараметрический аналог t-теста для зависимых выборок. Используется, если распределение переменной отлично от нормального

# Kruskal-Wallis test

Непараметрический аналог ANOVA. Используется, если распределение переменной отлично от нормального

## Параметрический тест

## Непараметрический тест

Независимые выборки

Зависимые выборки

Независимые выборки

Зависимые выборки

`t_test(..., paired = FALSE)`

`t_test(..., paired = TRUE)`

`wilcox_test(..., paired=FALSE)`

`wilcox_test(..., paired=TRUE)`

## Если выборок >2

Независимые выборки

Зависимые выборки

Независимые выборки

Зависимые выборки

`anova_test(...)`

`anova_test(.., wid=ID)`

`kruskal_test(...)`

`friedman_test(...)`

# Задания на семинаре

датасет **parkinsons\_disease\_data.csv**

Проверьте следующие гипотезы:

1. Различается ли ИМК у курящих и некурящих людей?
2. Связана ли депрессия с результатами теста MoCA?
3. Связано ли наличие речевых нарушений у пациента с паркинсонизмом?