

# Лекция 5. Количественные методы сравнения различных групп данных

---

# **Сравнение 2-х выборок с использованием параметрических критериев**

---

# Статистические гипотезы

---

**Нулевая гипотеза**, обозначаемая  $H_0$ , формулируется как гипотеза об отсутствии отличий: о сходстве двух распределений, о равенстве средних арифметических двух выборок и т.п. Нулевой она называется потому, что содержит 0:  $X_1 - X_2 = 0$ , где  $X_1$  и  $X_2$  - значения признаков. Нулевая гипотеза утверждает, к примеру, что результаты выполнения задания экспериментальной группой и контрольной не различаются.

**Альтернативная гипотеза**  $H_1$  противоположна по смыслу нулевой, она утверждает наличие отличий в выборках, в параметрах их распределений и так далее (результаты экспериментальной группы значимо отличаются от результатов контрольной группы).

Статистическая гипотеза может быть направленной или ненаправленной

---

**Ненаправленная** гипотеза фиксирует только наличие или отсутствие различий:

ненаправленная нулевая гипотеза:

- $H_0$  - результаты экспериментальной группы значимо не отличаются от контрольной

ненаправленная альтернативная гипотеза:

- $H_a$  - результаты экспериментальной группы значимо отличаются от контрольной

Статистическая гипотеза может быть направленной или ненаправленной

---

**Направленная** гипотеза фиксирует только наличие или отсутствие различий:

направленная нулевая гипотеза:

- $H_0$  - результаты экспериментальной группы не превышают результаты контрольной

направленная альтернативная гипотеза:

- $H_a$  - результаты экспериментальной группы выше результатов контрольной группы

# Статистические критерии

---

**Статистический критерий** - это правило, которое позволяет принимать истинную и отклонять ложную гипотезу с высокой степенью вероятности.

Математически критерий представляет собой формулу, по которой мы рассчитываем некоторое число

# Статистические критерии

---

статистические критерии делятся на параметрические и непараметрические.

***Параметрическими*** называются критерии, в формулу расчета которых входят параметры распределения

(чаще всего это *среднее арифметическое и стандартное отклонение*)

***Непараметрические*** критерии, соответственно, параметры распределение в формулу расчета не включают, они оперируют только *частотами* или *рангами*

Действия исследователя	Состояние нулевой гипотезы	
	Истинное	Ложное
Принимается $H_0$	Принято правильное решение ( $p=1-\alpha$ )	Совершена ошибка 2-го рода ( $p= \beta$ )
Отклоняется $H_0$	Совершена ошибка 1-го рода ( $p=\alpha$ )	Принято правильное решение ( $p= 1-\beta$ )

Вероятность ошибки 1-го рода обозначается  $\alpha$ , соответственно вероятность правильного решения будет  $1-\alpha$ .

$\alpha$  называют уровнем значимости различий

Вероятность  $1-\alpha$  называется доверительной вероятностью.

В каждом исследовании указывают вероятность ошибки  $\alpha$  (либо доверительную вероятность  $1-\alpha$ ) или в виде десятичной дроби ( $\alpha=0.05$ ), или в процентах ( $\alpha=5\%$ )



# Порядок решения задач с использованием статистических критериев

---

1. Формулируем нулевую и альтернативную гипотезы
2. Выбираем критерий, соответствующий типу задачи и удовлетворяющий условиям
3. Определяем показатели выборки, необходимые для вычисления критерия
4. Вычисляем эмпирическое значение критерия
5. Выписываем табличные (критические) значения, соответствующие уровню значимости 0,01 и 0,05
6. Выбираем, какая из гипотез является верной на основании сравнения эмпирического значения критерия с критическими
7. Формулируем ответ исходя из требований исследования (задачи)

# Выборки

---

Статистика делит выборки на:

1. Зависимые – получаются путем повторных или парных измерений
2. Независимые – все остальные выборки.

# Сравнение независимых выборок критерием Стьюдента

---

*По  $t$ -критерию Стьюдента* производится сравнение средних арифметических значений, то есть уровня признака, в двух независимых выборках с распределением частот, соответствующим *нормальному типу*

# 1. Формулируем гипотезы

---

- $H_0$  – Среднее арифметическое выборок можно считать одинаковым
- $H_a$  – Среднее арифметическое 1-ой выборки значимо отличается от среднего арифметического 2-ой выборки,

## 2. Выбираем критерий

---

$t$  – критерий Стьюдента для независимых выборок  
(показатели выборки должны иметь нормальное распределение)

### 3. Показатели выборки, необходимые для вычисления критерия

---

$M_x$  – среднее арифметическое 1 выборки

$M_y$  – среднее арифметическое 2 выборки

$D_x$  – дисперсия 1 выборки

$D_y$  – дисперсия 2 выборки

$n_x$  – объем 1 выборки

$n_y$  – объем 2 выборки

## 4. Эмпирическое значение критерия

---

$$t_{\text{ЭМП}} = \frac{M_x - M_y}{\sqrt{\frac{D_x \cdot (n_x - 1) + D_y \cdot (n_y - 1)}{n_x + n_y - 2} \cdot \left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

# 5. Табличные значения

$t_{0.05}(\nu)$  и  $t_{0.01}(\nu)$

$\nu$  - число степеней  
свободы

$$\nu = n_x + n_y - 2$$

Таблица критических значений t-критерия Стьюдента.

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,893	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,179	2,681	3,055	3,930	4,318
13	1,771	2,160	2,650	3,012	3,852	4,221
14	1,761	2,145	2,624	2,977	3,787	4,140
15	1,753	2,131	2,602	2,947	3,733	4,073
16	1,746	2,120	2,583	2,921	3,686	4,015
17	1,740	2,110	2,567	2,898	3,646	3,965
18	1,734	2,101	2,552	2,878	3,610	3,922
19	1,729	2,093	2,539	2,861	3,579	3,883
20	1,725	2,086	2,528	2,845	3,552	3,850
21	1,721	2,080	2,518	2,831	3,527	3,810



## 6. Выбор гипотезы

---

Если  $t_{\text{эмп}} < t_{0.05}$  верна  $H_0$

- различия не обнаружены

Если  $t_{\text{эмп}} > t_{0.01}$  верна  $H_a$

- различия значимы (уровень значимости 0,01)

Если  $t_{0.05} \leq t_{\text{эмп}} \leq t_{0.01}$  нужны дополнительные исследования

- различия значимы (уровень значимости 0,05)

# 7. Ответ

---

**Пример** В одной из школ Ленинградской области ученики 10 и 11 класса выполнили задания ШТУР (Школьный тест умственного развития). Различаются ли ученики 10 и 11 класса по уровню общей осведомленности?

общая осведомленность		абс частота	отн частота
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

общая осведомленность		абс частота	отн частота
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%

# 1. Формулируем гипотезы

---

- $H_0$  – Среднее арифметическое показателей ШТУР у учеников 10 и 11 класса по можно считать одинаковым
- $H_a$  – Среднее арифметическое показателей ШТУР у учеников 11 класса значимо выше среднего арифметического учеников 10 класса,

## 2. Выбираем критерий

---

$t$  – критерий Стьюдента для независимых выборок  
(показатели выборки должны иметь нормальное  
распределение)

### 3. Показатели выборки, необходимые для вычисления критерия

---

<i>общая осведомленность</i>		абс частота	отн частота
<i>ь</i>			
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

<i>общая осведомленность</i>		абс частота	отн частота
<i>ь</i>			
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%

### 3. Показатели выборки, необходимые для вычисления критерия

---

$M_x$  – среднее арифметическое 1 выборки

$M_y$  – среднее арифметическое 2 выборки

$D_x$  – дисперсия 1 выборки

$D_y$  – дисперсия 2 выборки

$$n_x = 28$$

$$n_y = 31$$

### 3. Показатели выборки, необходимые для вычисления критерия

---

$M_x$  – среднее арифметическое 1 выборки

$M_y$  – среднее арифметическое 2 выборки

$D_x$  – дисперсия 1 выборки

$D_y$  – дисперсия 2 выборки

$$n_x = 28$$

$$n_y = 31$$

$$M_x = \frac{\sum x_i \cdot f_i}{n}$$



### 3. Показатели выборки, необходимые для вычисления критерия

---

<i>общая осведомленность</i>		абс частота	отн частота
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

<i>общая осведомленность</i>		абс частота	отн частота
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%

### 3. Показатели выборки, необходимые для вычисления критерия

общая осведомленность		абс частота	отн частота
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

$$M_x = \frac{23 \cdot 3 + 26 \cdot 7 + 29 \cdot 12 + 32 \cdot 5 + 35 \cdot 1}{28}$$

$$M_x = 28,357$$

$$n_x = 28$$

$$n_y = 31$$

### 3. Показатели выборки, необходимые для вычисления критерия

---

$$M_y = \frac{23 \cdot 4 + 26 \cdot 5 + 29 \cdot 6 + 32 \cdot 12 + 35 \cdot 4}{31}$$

$$M_y = 29,677$$

$$n_x = 28$$

$$n_y = 31$$

общая осведомленность		абс частот а	отн частота
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%

### 3. Показатели выборки, необходимые для вычисления критерия

---

$$M_x = 28,357$$

$$M_y = 29,677$$

$D_x$  – дисперсия 1 выборки

$D_y$  – дисперсия 2 выборки

$n_x$  – объем 1 выборки

$n_y$  – объем 2 выборки

### 3. Показатели выборки, необходимые для вычисления критерия

---

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x -$$

$$D_y -$$

$$n_x = 28$$

$$n_y = 31$$

Дисперсия 1

$$D_x = \sigma_x^2 = \frac{\sum (x_i - M_x)^2 \cdot f_i}{n - 1}$$

Дисперсия 2

(получается из дисперсии 1, если раскрыть скобки и привести подобные)

$$D = \sigma_x^2 = \frac{\sum (x_i)^2 \cdot f_i - n \cdot M_x^2}{n - 1}$$

### 3. Показатели выборки, необходимые для вычисления критерия

---

<i>общая осведомленность</i>		абс частота	отн частота
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

<i>общая осведомленность</i>		абс частота	отн частота
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%

### 3. Показатели выборки, необходимые для вычисления критерия

---

общая осведомленность		абс частота	отн частота
интервал	ср точка		
22 - 24	23	3	10,71%
25 - 27	26	7	25,00%
28 - 30	29	12	42,86%
31 - 33	32	5	17,86%
34 - 36	35	1	3,57%

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x = \frac{23^2 \cdot 3 + 26^2 \cdot 7 + 29^2 \cdot 12 + 32^2 \cdot 5 + 35^2 \cdot 1 - 28.357^2 \cdot 28}{28 - 1}$$

$$D_x = 8,905$$

$$n_x = 28$$

$$n_y = 31$$

### 3. Показатели выборки, необходимые для вычисления критерия

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x = 8,905$$

$$D_y = \frac{23^2 \cdot 4 + 26^2 \cdot 5 + 29^2 \cdot 6 + 32^2 \cdot 12 + 35^2 \cdot 4 - 29.677^2 \cdot 31}{31 - 1}$$

$$D_y = 14,226$$

$$n_x = 28$$

$$n_y = 31$$

общая осведомленность		абс частота	отн частота
интервал	ср точка		
22 - 24	23	4	12,90%
25 - 27	26	5	16,13%
28 - 30	29	6	19,35%
31 - 33	32	12	38,71%
34 - 36	35	4	12,90%



### 3. Показатели выборки, необходимые для вычисления критерия

---

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x = \frac{23^2 \cdot 3 + 26^2 \cdot 7 + 29^2 \cdot 12 + 32^2 \cdot 5 + 35^2 \cdot 1 - 28.357^2 \cdot 28}{28 - 1}$$

$$D_y =$$

$$n_x = 28$$

$$n_y = 31$$

### 3. Показатели выборки, необходимые для вычисления критерия

---

$$D_x = \frac{23^2 \cdot 3 + 26^2 \cdot 7 + 29^2 \cdot 12 + 32^2 \cdot 5 + 35^2 \cdot 1 - 28.357^2 \cdot 28}{28 - 1}$$

$$D_y = \frac{23^2 \cdot 4 + 26^2 \cdot 5 + 29^2 \cdot 6 + 32^2 \cdot 12 + 35^2 \cdot 4 - 29.677^2 \cdot 31}{31 - 1}$$

### 3. Показатели выборки, необходимые для вычисления критерия

---

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x = 8,905$$

$$D_y = 14,226$$

$$n_x = 28$$

$$n_y = 31$$

## 4. Эмпирическое значение критерия

---

$$t_{\text{ЭМП}} = \frac{M_x - M_y}{\sqrt{\frac{D_x \cdot (n_x - 1) + D_y \cdot (n_y - 1)}{n_x + n_y - 2} \cdot \left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

## 4. Эмпирическое значение критерия

---

$$M_x = 28,357$$

$$M_y = 29,677$$

$$D_x = 8,905$$

$$D_y = 14,226$$

$$n_x = 28$$

$$n_y = 31$$

$$t_{\text{эмп}} = \frac{M_x - M_y}{\sqrt{\frac{D_x \cdot (n_x - 1) + D_y \cdot (n_y - 1)}{n_x + n_y - 2} \cdot \left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$
$$t_{\text{эмп}} = \frac{28.357 - 29.677}{\sqrt{\frac{8.905 \cdot (28 - 1) + 14.226 \cdot (31 - 1)}{28 + 31 - 2} \cdot \left(\frac{1}{28} + \frac{1}{31}\right)}}$$

## 4. Эмпирическое значение критерия

---

$$t_{\text{эмп}} = \frac{28.357 - 29.677}{\sqrt{\frac{8.905 \cdot (28 - 1) + 14.226 \cdot (31 - 1)}{28 + 31 - 2} \cdot \left(\frac{1}{28} + \frac{1}{31}\right)}}$$

$$t_{\text{эмп}} = 1,480$$

## 5. Табличные значе

$t_{0.05}(v)$  и  $t_{0.01}(v)$

$v$  - число степеней  
свободы

$$v = 28 + 31 - 2 = 57$$

$$t_{0.05}(57) = 1,960 \quad \text{и}$$

$$t_{0.01}(57) = 2,576$$

Таблица критических значений t-критерия Стьюдента.

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610

23	1,714	2,069	2,500	2,807	3,485	3,767
24	1,711	2,064	2,492	2,797	3,467	3,745
25	1,708	2,060	2,485	2,787	3,450	3,725
26	1,706	2,056	2,479	2,779	3,435	3,707
27	1,703	2,052	2,473	2,771	3,421	3,690
28	1,701	2,048	2,467	2,763	3,408	3,674
29	1,699	2,045	2,462	2,756	3,396	3,659
30	1,697	2,042	2,457	2,750	3,385	3,646
$\infty$	1,645	1,960	2,326	2,576	3,090	3,291

## 6. Выбор гипотезы

---

$$t_{\text{эмп}} < t_{0.05} \quad 1,480 < 1,960$$

верна  $H_0$  - различия не обнаружены

,



## 7. Ответ

---

Между девушками 10 и 11 класса значимых различий по уровню общей осведомленности не обнаружено

# **Сравнение 2-х выборок с использованием непараметрических критериев**

---

# 1. Формулируем гипотезы

---

- $H_0$  – Уровень выраженности исследуемого показателя в выборках примерно одинаковый
- $H_a$  – Уровень выраженности исследуемого показателя в 1-ой выборке значимо отличается от уровня 2-ой выборки,

## 2. Выбираем критерий

---

U – критерий Манна-Уитни

(показатели выборки могут иметь любое распределение)

# 3. Показатели выборки, необходимые для вычисления критерия

---

Проранжировать все измерения, *объединив результаты двух выборок*. Объем 1-ой выборки  $n_1$ , объем 2-ой выборки  $n_2$

Правила ранжирования следующие:

- Наименьшему значению из всех присваивается ранг 1, наибольшему -  $n_1+n_2$ .
- Если два или большее количество значений равны, то для них рассчитывается средний ранг. Например, если три наименьших значения в выборке равны, то их средний ранг был бы  $(1+2+3)/3=2$ .
- следующий за ними человек получает ранг с учетом всех, которые должны быть заняты, т.е. за рангом 2, 2, 2 (должны были быть 1, 2, 3) идет ранг 4

А если равны 10 и 11 значения, то их средний ранг  $(10+11)/2=10.5$ . следующий за ними человек получает ранг 12.

## 4. Эмпирическое значение критерия

---

Проранжировать все измерения, *объединив результаты двух выборок*.

Подсчитать сумму рангов отдельно для первой ( $\Sigma R_1$ ) и для второй ( $\Sigma R_2$ ) выборки.

Проверить, совпадает ли общая сумма рангов с расчетной по формуле  $\Sigma R_1 + \Sigma R_2 = N(N+1)/2$ , где  $N = n_1 + n_2$  - общее число ранжируемых значений.

Рассчитать эмпирические значения критерия Манна-Уитни для выборок:

$$U_1 = n_1 \cdot n_2 + n_1(n_1 + 1)/2 - \Sigma R_1$$

$$U_2 = n_1 \cdot n_2 + n_2(n_2 + 1)/2 - \Sigma R_2$$

Эмпирическим значением является меньшее из двух рассчитанных значений ( $U_1$  и  $U_2$ ).

---

Оно сопоставляется с критическими значениями, для  $\alpha=0.05$  и  $\alpha=0.01$ .

Таблица КРИТЕРИЙ МАННА-УИТНИ (ВИЛКОКСОНА-МАННА-УИТНИ) (U)

В таблице показано максимальное значение(U) при котором различия могут считаться статистически значимыми с доверительной вероятностью 95% (уровень значимости=0,05)

$n_1 \backslash n_2$	4	5	6	7	8	9	10	11	12	14	15
4	1										
5	2	4									
6	3	5	7								
7	4	6	8	11							
8	5	8	10	13	15						
9	6	9	12	15	18	21					
10	7	11	14	17	20	24	27				
11	8	12	16	19	23	27	31	34			
12	9	13	17	21	26	30	34	38	42		
13	10	15	19	24	28	33	37	42	47	51	
14	11	16	21	26	31	36	41	46	51	56	61



Таблица КРИТЕРИЙ МАННА-УИТНИ (ВИЛКОКСОНА-МАННА-УИТНИ) (U)

В таблице показано максимальное значение(U) при котором различия могут считаться статистически значимыми с доверительной вероятностью 99% (уровень значимости=0,01)

$n_1 \backslash n_2$	4	5	6	7	8	9	10	11	12	14	15
4	-										
5	0	1									
6	1	2	3								
7	1	3	4	6							
8	2	4	6	7	9						
9	3	5	7	9	11	14					
10	4	6	8	11	13	16	19				
11	4	7	9	12	15	18	22	25			
12	5	8	11	14	17	21	24	28	31		
13	5	9	12	16	20	23	27	31	35	39	
14	6	10	13	17	22	26	30	34	38	43	47

## 6. Выбор гипотезы

---

Оно сопоставляется с критическими значениями, для  $\alpha=0.05$  и  $\alpha=0.01$ .

Чем меньше значение  $U$ , тем **выше достоверность различий**, а значит,

если  $U_{\text{эмп}} \geq U_{\text{крит}}$ , то принимается нулевая гипотеза  $H_0$ ,

если  $U_{\text{эмп}} \leq U_{\text{крит}}$ , то принимается альтернативная гипотеза.

# Пример:

Сравнить девочек  
и мальчиков по  
направленности  
на себя

№ п/п	пол	направленность на себя
1	муж	27
2	муж	34
3	муж	33
4	муж	24
5	муж	26
6	муж	23
7	жен	38
8	жен	29
9	муж	30
10	муж	22
11	жен	38
12	муж	29
13	жен	40
14	жен	31
15	жен	36
16	жен	37
17	жен	33
18	муж	32
19	жен	35
20	жен	33
21	муж	30

# 1. Формулируем гипотезы

---

- $H_0$  – Уровень направленности на себя у девочек и мальчиков примерно одинаковый
- $H_a$  – Уровень направленности на себя у девочек значительно отличается от уровня направленности на себя у мальчиков

## 2. Выбираем критерий

---

U – критерий Манна-Уитни

### 3. Показатели выборки, необходимые для вычисления критерия

---

Проранжировать все измерения, *объединив результаты двух выборок.*

Объем 1-ой выборки (мальчики)  $n_1=11$

объем 2-ой выборки (девочки)  $n_2 =10$

№ п/п	пол	направленность на себя			
		1-выборка (м)		2-выборка (ж)	
		баллы	ранг	баллы	ранг
10	муж	22	1		
6	муж	23	2		
4	муж	24	3		
5	муж	26	4		
1	муж	27	5		
8	жен			29	6,5
12	муж	29	6,5		
9	муж	30	8,5		
21	муж	30	8,5		
14	жен			31	10
18	муж	32	11		
3	муж	33	13		
17	жен			33	13
20	жен			33	13
2	муж	34	15		
19	жен			35	16
15	жен			36	17
16	жен			37	18
7	жен			38	19,5
11	жен			38	19,5
13	жен			40	21

№ п/п	пол	направленность на себя			
		1-выборка (м)		2-выборка (ж)	
		баллы	ранг	баллы	ранг
10	муж	22	1		
6	муж	23	2		
4	муж	24	3		
5	муж	26	4		
1	муж	27	5		
8	жен			29	6,5
12	муж	29	6,5		
9	муж	30	8,5		
21	муж	30	8,5		
14	жен			31	10
18	муж	32	11		
3	муж	33	13		
17	жен			33	13
20	жен			33	13
2	муж	34	15		
19	жен			35	16
15	жен			36	17
16	жен			37	18
7	жен			38	19,5
11	жен			38	19,5
13	жен			40	21
<b>Суммы</b>			<b>77,5</b>		<b>153,5</b>



## 4. Эмпирическое значение критерия

---

$$\Sigma R_1 = 77,5$$

$$\Sigma R_2 = 153,5$$

$$U_1 = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_2 = 11 \cdot 10 + \frac{11 \cdot 12}{2} - 77,5 = 98,5$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2(n_2 + 1)}{2} - R_1 = 11 \cdot 10 + \frac{10 \cdot 11}{2} - 153,5 = 11,5$$

$$U_{\text{эмп}} = \min(U_1, U_2) = 11,5$$

Таблица КРИТЕРИЙ МАННА-УИТНИ (ВИЛКОКСОНА-МАННА-УИТНИ) (U)

В таблице показано максимальное значение(U) при котором различия могут считаться статистически значимыми с доверительной вероятностью 95% (уровень значимости=0,05)

$n_1 \backslash n_2$	4	5	6	7	8	9	10	11	12	14	15
4	1										
5	2	4									
6	3	5	7								
7	4	6	8	11							
8	5	8	10	13	15						
9	6	9	12	15	18	21					
10	7	11	14	17	20	24	27				
11	8	12	16	19	23	27	31	34			
12	9	13	17	21	26	30	34	38	42		
13	10	15	19	24	28	33	37	42	47	51	
14	11	16	21	26	31	36	41	46	51	56	61

Таблица КРИТЕРИЙ МАННА-УИТНИ (ВИЛКОКСОНА-МАННА-УИТНИ) (U)

В таблице показано максимальное значение(U) при котором различия могут считаться статистически значимыми с доверительной вероятностью 99% (уровень значимости=0,01)

$n_1 \backslash n_2$	4	5	6	7	8	9	10	11	12	14	15
4	-										
5	0	1									
6	1	2	3								
7	1	3	4	6							
8	2	4	6	7	9						
9	3	5	7	9	11	14					
10	n2	6	8	11	13	16	19				
11	4	7	9	12	15	18	22	25			
12	5	8	11	14	17	21	24	28	31		
13	5	9	12	16	20	23	27	31	35	39	
14	6	10	13	17	22	26	30	34	38	43	47

## 6. Выбор гипотезы

---

$$U_{0.05}(11, 10) = 31,$$

$$U_{0,01}(11, 10) = 22$$

$$U_{\text{эмп}} = 11.5$$

$U_{\text{эмп}} \leq U_{\text{крит}}$ , принимается альтернативная гипотеза

Ответ: направленность на себя у девочек выше, чем у мальчиков  
(уровень значимости 0,01)

# Сравнение зависимых выборок критерием Стьюдента

---

## Определение значимого сдвига в показателях



# 1. Формулируем гипотезы

---

- $H_0$  – Среднее арифметическое выборок можно считать одинаковым
- $H_a$  – Среднее арифметическое 1-ой выборки значимо отличается от среднего арифметического 2-ой выборки,
- $H_0$  – в исследуемой выборке значимых изменений (сдвига в показателях) не произошло
- $H_a$  – в исследуемой выборке произошли значимые изменения (сдвиг в показателях)

## 2. Выбираем критерий

---

$t$  – критерий Стьюдента для зависимых выборок  
(показатели выборки должны иметь нормальное распределение)

### 3. Показатели выборки, необходимые для вычисления критерия

---

№ п/п	$x_i$	$y_i$	$d_i$	
1				
2				
3				

$x_i$  – показатель 1 человека в паре (показатель испытуемого в 1 срезе)

$y_i$  – показатель 2 человека в паре (показатель испытуемого во 2 срезе)

$d_i$  – разность показателей  $d_i = x_i - y_i$  (или  $d_i = y_i - x_i$ )

$n$  – количество пар



### 3. Показатели выборки, необходимые для вычисления критерия

---

№ п/п	$x_i$	$y_i$	$d_i$	$d_i^2$
1				
2				
3				

$x_i$  – показатель 1 человека в паре (показатель испытуемого в 1 срезе)

$y_i$  – показатель 2 человека в паре (показатель испытуемого во 2 срезе)

$d_i$  – разность показателей  $d_i = x_i - y_i$  (или  $d_i = y_i - x_i$ )

$n$  – количество пар

## 4. Эмпирическое значение критерия

---

$$t_{\text{эмп}} = \frac{\sum d_i}{\sqrt{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}} \cdot \sqrt{\frac{n-1}{n}}$$

# 5. Табличные значения

$t_{0.05}(\nu)$  и  $t_{0.01}(\nu)$

$\nu$  - число степеней  
свободы

$$\nu = n_x - 1$$

Таблица критических значений t-критерия Стьюдента.

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,893	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,179	2,681	3,055	3,930	4,318
13	1,771	2,160	2,650	3,012	3,852	4,221
14	1,761	2,145	2,624	2,977	3,787	4,140
15	1,753	2,131	2,602	2,947	3,733	4,073
16	1,746	2,120	2,583	2,921	3,686	4,015
17	1,740	2,110	2,567	2,898	3,646	3,965
18	1,734	2,101	2,552	2,878	3,610	3,922
19	1,729	2,093	2,539	2,861	3,579	3,883
20	1,725	2,086	2,528	2,845	3,552	3,850
21	1,721	2,080	2,518	2,831	3,527	3,810

## 6. Выбор гипотезы

---

Если  $t_{\text{эмп}} < t_{0.05}$  верна  $H_0$

- различия не обнаружены ,

Если  $t_{\text{эмп}} > t_{0.01}$  верна  $H_a$

- различия значимы (уровень значимости 0,01)

Если  $t_{0.05} \leq t_{\text{эмп}} \leq t_{0.01}$  нужны дополнительные исследования

.  
- различия значимы (уровень значимости 0,05)

# 7. Ответ

---

**Пример** Определить есть ли сдвиг в показателях готовности к сотрудничеству в конфликтной ситуации в экспериментальном классе

№п/п		Экспериментальный класс	
		до	после
1	Марина К.	4	8
2	Тамара К.	8	11
3	Олег К.	3	9
4	Алексей К.	6	8
5	Антонина К.	10	8
6	Полина К.	6	8
7	Наталья Л.	5	7
8	Глеб О.	7	9
9	Влад П.	3	10
10	Андрей П.	6	8

# 1. Формулируем гипотезы

---

$H_0$ : уровень готовности к сотрудничеству в конфликтной ситуации учеников экспериментального после эксперимента значимо не изменился.

$H_a$ : уровень готовности к сотрудничеству в конфликтной ситуации учеников экспериментального после эксперимента значимо вырос

## 2. Выбираем критерий

---

$t$  – критерий Стьюдента для зависимых выборок  
(показатели выборки должны иметь нормальное распределение)



### 3. Показатели выборки, необходимые для вычисления критерия

---

№ п/п	$x_i$	$y_i$	$d_i$	
1				
2				
3				

$x_i$  – показатель 1 человека в паре (показатель испытуемого в 1 срезе)

$y_i$  – показатель 2 человека в паре (показатель испытуемого во 2 срезе)

$d_i$  – разность показателей  $d_i = x_i - y_i$  (или  $d_i = y_i - x_i$ )

$n$  – количество пар

### 3. Показатели выборки, необходимые для вычисления критерия

---

№ п/п	$x_i$	$y_i$	$d_i$	$d_i^2$
1				
2				
3				

$x_i$  – показатель 1 человека в паре (показатель испытуемого в 1 срезе)

$y_i$  – показатель 2 человека в паре (показатель испытуемого во 2 срезе)

$d_i$  – разность показателей  $d_i = x_i - y_i$  (или  $d_i = y_i - x_i$ )

$n$  – количество пар

### 3. Показатели выборки, необходимые для вычисления критерия

№п/п		Экспериментальный класс		d	d <sup>2</sup>
		до	после		
1	Марина К.	4	8		
2	Тамара К.	8	11		
3	Олег К.	3	9		
4	Алексей К.	6	8		
5	Антонина К.	10	8		
6	Полина К.	6	8		
7	Наталья Л.	5	7		
8	Глеб О.	7	9		
9	Влад П.	3	10		
10	Андрей П.	6	8		

### 3. Показатели выборки, необходимые для вычисления критерия

№п/п		Экспериментальный класс		d	d <sup>2</sup>
		до	после		
1	Марина К.	4	8	4	
2	Тамара К.	8	11	3	
3	Олег К.	3	9	6	
4	Алексей К.	6	8	2	
5	Антонина К.	10	8	-2	
6	Полина К.	6	8	2	
7	Наталья Л.	5	7	2	
8	Глеб О.	7	9	2	
9	Влад П.	3	10	7	
10	Андрей П.	6	8	2	

### 3. Показатели выборки, необходимые для вычисления критерия

№п/п		Экспериментальный класс		d	d <sup>2</sup>
		до	после		
1	Марина К.	4	8	4	16
2	Тамара К.	8	11	3	9
3	Олег К.	3	9	6	36
4	Алексей К.	6	8	2	4
5	Антонина К.	10	8	-2	4
6	Полина К.	6	8	2	4
7	Наталья Л.	5	7	2	4
8	Глеб О.	7	9	2	4
9	Влад П.	3	10	7	49
10	Андрей П.	6	8	2	4

## 4. Эмпирическое значение критерия

---

$$t_{\text{эмп}} = \frac{\sum d_i}{\sqrt{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}} \cdot \sqrt{\frac{n-1}{n}}$$

### 3. Показатели выборки, необходимые для вычисления критерия

		Экспериментальный класс		<b>d</b>	<b>d<sup>2</sup></b>
		<b>до</b>	<b>после</b>		
1	Марина К.	4	8	4	16
2	Тамара К.	8	11	3	9
3	Олег К.	3	9	6	36
4	Алексей К.	6	8	2	4
5	Антонина К.	10	8	-2	4
6	Полина К.	6	8	2	4
7	Наталья Л.	5	7	2	4
8	Глеб О.	7	9	2	4
9	Влад П.	3	10	7	49
10	Андрей П.	6	8	2	4
				<b>28</b>	<b>134</b>

## 4. Эмпирическое значение критерия

---

$$t_{\text{эмп}} = \frac{\sum d_i}{\sqrt{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}} \cdot \sqrt{\frac{n-1}{n}}$$

$$t_{\text{эмп}} = \frac{28}{\sqrt{134 - \frac{28^2}{10}}} \cdot \sqrt{\frac{10-1}{10}} =$$



## 4. Эмпирическое значение критерия

---

$$t_{\text{эмп}} = \frac{\sum d_i}{\sqrt{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}} \cdot \sqrt{\frac{n-1}{n}}$$

$$t_{\text{эмп}} = \frac{28}{\sqrt{134 - \frac{28^2}{10}}} \cdot \sqrt{\frac{10-1}{10}} = 3.562$$

# 5. Табличные значения

$t_{0.05}(v)$  и  $t_{0.01}(v)$

$v$  - число степеней свободы

$$v = n_x - 1 = 9$$

Таблица критических значений t-критерия Стьюдента.

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,893	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,177	2,681	3,055	3,920	4,310

## 5. Табличные значения

$t_{0.05}(v)$  и  $t_{0.01}(v)$

$v$  - число степеней свободы

$$v = n_x - 1 = 9$$

$$t_{0.05}(9) = 2,262 \quad \text{и}$$

$$t_{0.01}(9) = 3,250$$

Таблица критических значений t-критерия Стьюдента.

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,314	12,71	31,82	63,66	318,3	636,6
2	2,920	4,303	6,965	9,925	22,33	31,60
3	2,353	3,182	4,541	5,841	10,21	12,92
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,893	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,177	2,681	3,055	3,920	4,310

## 6. Выбор гипотезы

---

$$3,562 > 3,250$$

$$t_{\text{эмп}} > t_{0.01} \quad \text{верна } H_a$$

- различия значимы (уровень значимости 0,01)

## 7. Ответ

---

уровень готовности к сотрудничеству в конфликтной ситуации учеников экспериментального после эксперимента значительно вырос ( $t$  Стьюдента = 3,562, уровень значимости 0,01).