

Критерии различий

**Сравнение более двух
выборок**

Сравнение более двух независимых выборок

Критерий H Краскала-Уоллеса

- Критерий является непараметрическим.
- Аналог – однофакторный дисперсионный анализ ANOVA для независимых выборок.

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

где

N – суммарная численность всех выборок

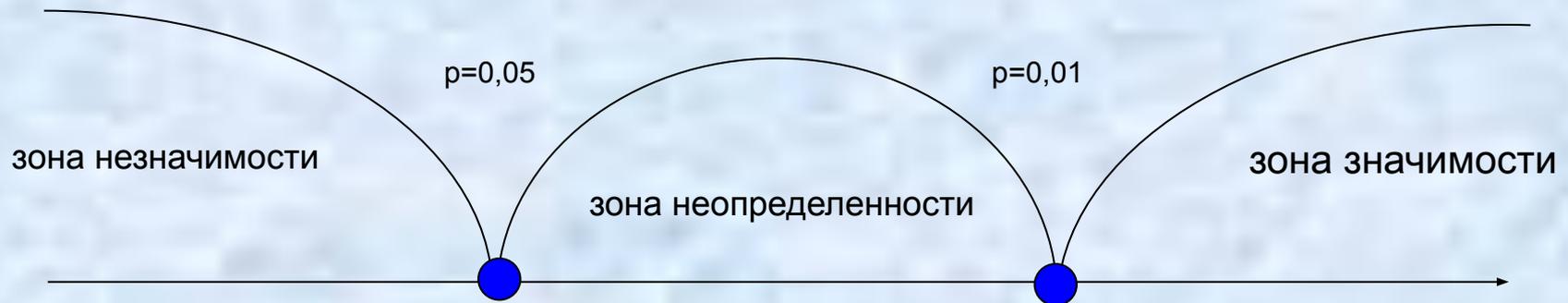
k – количество сравниваемых выборок

R_i – сумма рангов для выборки i

n_i – численность выборки i

- H_0 : Отсутствуют статистически достоверные различия между выборками.
- H_1 : Имеются статистически достоверные различия между выборками.

- Чем сильнее различаются выборки, тем больше вычисленное значение N и тем меньше p -уровень значимости.



- Если сравниваются 3 выборки и объем каждой выборки меньше 5, то пользуются таблицей критических значений H-Краскала-Уоллеса.
- Если объем хотя бы одной выборки больше 5 *либо* количество выборок больше 3, то пользуются таблицей критических значений χ^2 для $df=k-1$ (k – число выборок)

№	Выборка 1	Выборка 2	Выборка 3
1	3	5	14
2	4	9	16
3	6	12	17
4	7	15	
5	8	19	
6	10		
7	11		
8	13		

Значения	Выборка	Ранги	Ранги 1	Ранги 2	Ранги 3
3	1	1	1		
4	1	2	2		
5	2	3		3	
6	1	4	4		
7	1	5	5		
8	1	6	6		
9	2	7		7	
10	1	8	8		
11	1	9	9		
12	2	10		10	
13	1	11	11		
14	3	12			12
15	2	13		13	
16	3	14			14
17	3	15			15
19	2	16		16	
Суммы рангов			46	49	41

- 1. Значения выборок объединяются в один ряд, упорядоченный в порядке возрастания или убывания.
Обозначается принадлежность каждого значения к той или иной выборке.
- 2. Значения выборок ранжируются и выписываются отдельно ранги для каждой выборки.
- 3. Вычисляются суммы рангов.

- 4. Вычисляется N по формуле.
 - 5. Определяется p -уровень значимости.
 - 6. Принимается статистическое решение.
-
- $N=6,575$

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КРИТЕРИЯ χ^2

df	p				df	p			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
1	2,706	3,842	6,635	10,829	46	58,641	62,841	71,221	81,431
2	4,605	5,992	9,211	13,817	47	59,774	64,013	72,463	82,752
3	6,251	7,815	11,346	16,269	48	60,907	65,183	73,703	84,069
4	7,779	9,488	13,278	18,470	49	62,038	66,351	74,940	85,384
5	9,236	11,071	15,088	20,519	50	63,167	67,518	76,175	86,694
6	10,645	12,593	16,814	22,462	51	64,295	68,683	77,408	88,003
7	12,017	14,068	18,478	24,327	52	65,422	69,846	78,638	89,308
8	13,362	15,509	20,093	26,130	53	66,548	71,008	79,866	90,609
9	14,684	16,921	21,669	27,883	54	67,673	72,168	81,092	91,909
10	15,987	18,309	23,213	29,594	55	68,796	73,326	82,316	93,205



Подтверждается гипотеза
 H_1
Имеются значимые
различия между
выборками

Сравнение более двух зависимых выборок

- Непараметрический критерий χ^2 -Фридмана (хи-квадрат).
- Аналог – однофакторный дисперсионный анализ ANOVA для повторных измерений.

$$\chi^2 = \left[\frac{12}{Nk(k+1)} \cdot \sum_{i=1}^k R_i^2 \right] - 3N(k+1)$$

$$df = k - 1$$

где

N – число объектов (испытуемых)

k – количество условий (повторных измерений)

R_i – сумма рангов для условия i

- Если $k=3$, $N>9$ или $k>3$, $N>4$, то пользуются обычной таблицей для χ^2 .
- Если $k=3$, $N<10$ или $k=4$, $N<5$, то пользуются дополнительными таблицами критических значений χ^2 -Фридмана.

№	Условие 1		Условие 2		Условие 3		Условие 4	
	X	Ранг	X	Ранг	X	Ранг	X	Ранг
1	6	2	14	3.5	5	1	14	3.5
2	11	3	5	2	4	1	12	4
3	12	4	8	2	7	1	10	3
4	8	1	10	2	11	3	12	4
5	5	1	14	3.5	10	2	14	3.5
6	10	3	7	2	6	1	12	4
Сумма рангов:		14		15		9		22

1. Для каждого объекта условия ранжируются (по строке).
2. Вычисляется сумма рангов для каждого условия.
3. Вычисляется эмпирическое значение χ^2 по формуле.
4. Определяется уровень значимости.
5. Принимается статистическое решение.

- $\chi^2=8,6$

- $df=3$

КРИТИЧЕСКИЕ ЗНАЧЕНИЯ КРИТЕРИЯ χ^2

df	p				df	p			
	0,10	0,05	0,01	0,001		0,10	0,05	0,01	0,001
1	2,706	3,842	6,635	10,829	46	58,641	62,841	71,221	81,431
2	4,605	5,992	9,211	13,817	47	59,774	64,013	72,463	82,752
3	6,251	7,815	11,346	16,269	48	60,907	65,183	73,703	84,069
4	7,779	9,488	13,278	18,470	49	62,038	66,351	74,940	85,384
5	9,236	11,071	15,088	20,519	50	63,167	67,518	76,175	86,694
6	10,645	12,593	16,814	22,462	51	64,295	68,683	77,408	88,003
7	12,017	14,068	18,478	24,327	52	65,422	69,846	78,638	89,308
8	13,362	15,509	20,093	26,130	53	66,548	71,008	79,866	90,609
9	14,684	16,921	21,669	27,883	54	67,673	72,168	81,092	91,909
10	15,987	18,309	23,213	29,594	55	68,796	73,326	82,316	93,205



Подтверждается гипотеза
 H_1
 Имеются значимые
 различия между
 выборками