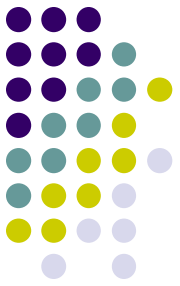
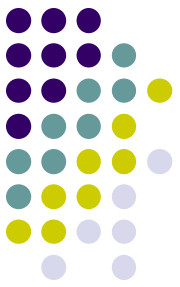


H-критерий Краскела – Уоллиса



- Предназначен для оценки различий по степени выраженности анализируемого признака одновременно между двумя, тремя и более выборками
- Критерий позволяет установить, что уровень признака изменяется при переходе от выборке к выборке, но не указывает на направление этих изменений
- Гипотезы
 - H_0 : между выборками 1, 2, 3 и т.д. существуют лишь случайные различия по уровню исследуемого признака
 - H_1 : между выборками 1, 2, 3 и т.д. существуют неслучайные различия по уровню исследуемого признака

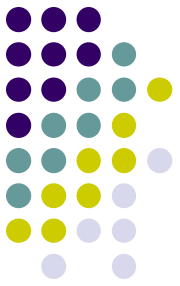
H-критерий Краскела – Уоллиса



Ограничения:

- если $\alpha=0,05$, то при сопоставлении трех выборок допускается, чтобы в одной выборке было 3 значения, а во второй и третьей тогда по 2 значения
- если $\alpha=0,01$, то необходимо чтобы в каждой выборке было не менее 3 наблюдений, или чтобы по крайней мере в одной из выборок было 4 наблюдения, а в двух других – по 2
- при большом количестве выборок и испытуемых в каждой выборке необходимо пользоваться таблицей критических значений критерия χ^2 при числе степеней свободы $\nu = c-1$
- при множественном сопоставлении выборок достоверные различия между какой-либо конкретной парой (или парами) их могут оказаться стертыми

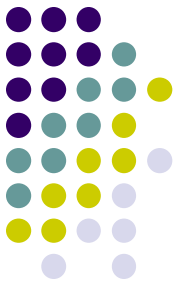
H-критерий Краскела – Уоллиса



Алгоритм *H*-теста:

- 1) обе выборки соединяются в единую выборку. При этом запоминается, к какой выборке относится каждый элемент массива
- 2) общая выборка ранжируется по возрастанию
- 3) разбить единую выборку на прежние две выборки
- 4) подсчитать сумму рангов отдельно по каждой выборке.
Проверить, совпадает ли общая сумма рангов с расчетной

H-критерий Краскела – Уоллиса



Алгоритм *H*-теста:

5) найти значение *H*

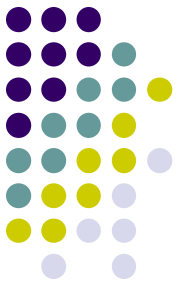
$$H = \frac{12}{N(N+1)} \sum_{i=1}^c \frac{R_i^2}{n_i} - 3(N+1),$$

где $N = \sum n_i$ – общий количество наблюдений,

R_i – ранговая сумма i -той группы,

n_i – объем i -той группы

H-критерий Краскела – Уоллиса



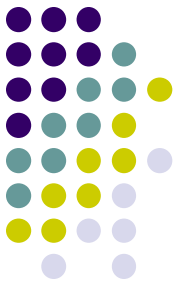
Алгоритм *H*-теста:

6) найти эмпирическое значение $H_{эмп}$, внося поправку на одинаковые ранги

$$H_{эмп} = \frac{H}{1 - \frac{\sum (t^3 - t)}{N^3 - N}},$$

где t – объем группы одинаковых рангов

H-критерий Краскела – Уоллиса



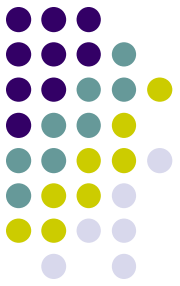
Алгоритм **H**-теста:

- 7) Найти критическое значение критерия $H_{крит}$ при числе групп $c=3$ – по таблице критических значений **H**-критерия при числе групп $c>3$ – по таблице критических значения критерия χ^2

Таблица критических значений критерия H Краскала – Уоллиса для разных сочетаний n_1, n_2, n_3

n_1	n_2	n_3	H	ρ	n_1	n_2	n_3	H	ρ	n_1	n_2	n_3	H	ρ
2	1	1	2,7000	0,500	3	3	3	7,2000	0,004	4	4	2	7,0364	0,006
								6,4889	0,011				6,8727	0,011
								5,6889	0,029				5,4545	0,046
								5,6000	0,050				5,2364	0,052
								5,0667	0,086				4,5545	0,098
								4,6222	0,100				4,4455	0,103
2	2	1	3,6000	0,200	4	1	1	3,5714	0,200	4	4	3	7,1439	0,010
													7,1364	0,011
													5,5985	0,049
													5,5758	0,051
													4,5455	0,099
													4,4773	0,102
2	2	2	4,5714	0,067	4	2	1	4,8214	0,057	4	4	4	7,6538	0,008
								4,5000	0,076				7,5385	0,011
								4,0179	0,114				5,6923	0,049

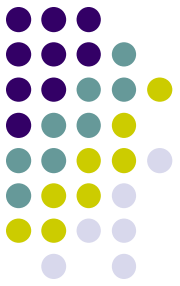
H-критерий Краскела – Уоллиса



Алгоритм *H*-теста:

- 8) Сравнить найденное эмпирическое значение критерия $H_{эмп}$ и критическое значение критерия $H_{крит}$
- если $H_{эмп} < H_{крит}$ ($p > \alpha$), то нулевая гипотеза H_0 принимается
 - если уровень значимости $H_{эмп} \geq H_{крит}$ ($p \leq \alpha$), то нулевая гипотеза H_0 отклоняется

H-критерий Краскела – Уоллиса



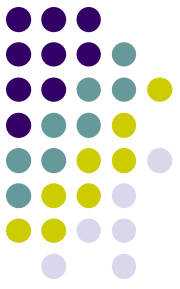
Пример:

В эксперименте по исследованию интеллектуальной настойчивости 22 испытуемым предъявлялись неразрешимые анаграммы. Показатели длительности попыток в решении представлены в таблице

	Группа 1: анаграмма ФОЛИТОН	Группа 2: анаграмма КАМУСТО	Группа 3: анаграмма СНЕРАКО	Группа 4: анаграмма ГРУТОСИЛ
1	145	145	128	60
2	194	210	283	2361
3	731	236	469	2416
4	1200	385	482	3600
5		720	1678	
6		848	2081	
7		905		
8		1080		

Можно ли утверждать, что длительность попыток решения каждой из 4 неразрешимых анаграмм примерно одинакова?

H-критерий Краскела – Уоллиса



Гипотезы

H_0 : четыре группы испытуемых, получившие разные неразрешимые анаграммы, не различаются по длительности попыток их решения

H_1 : четыре группы испытуемых, получившие разные неразрешимые анаграммы, различаются по длительности попыток их решения

Проверка гипотез в STATISTICA при помощи *H*-критерия Краскела – Уоллиса



1) Занести все показатели в один столбец

Промаркировать показатели:

1 – для первой выборки,

2 – для второй выборки и т.д.

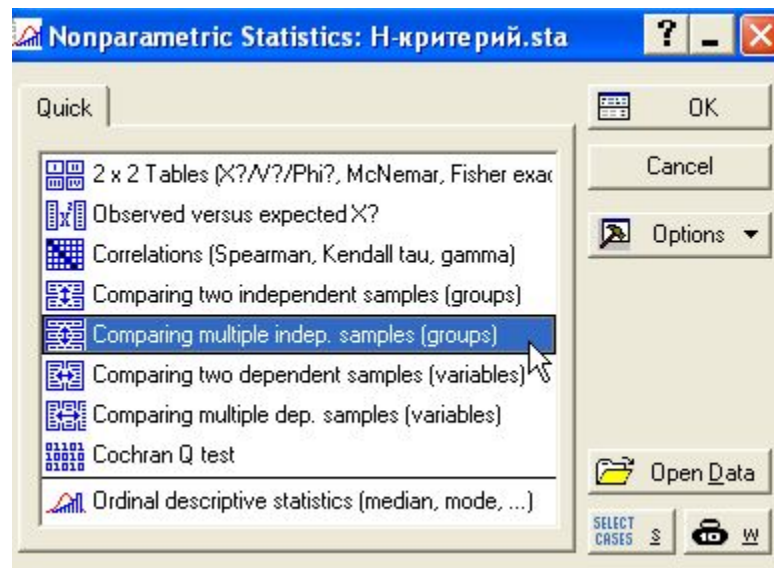
Data: H-критерий.sta (2v...)

	1 Длительность	2 Маркеры
1	145	1
2	194	1
3	731	1
4	1200	1
5	145	2
6	210	2
7	236	2
8	385	2
9	720	2
10	848	2
11	905	2
12	1080	2
13	128	3
14	283	3
15	469	3
16	482	3
17	1678	3
18	2081	3
19	60	4
20	2361	4
21	2416	4
22	3600	4

Проверка гипотез в STATISTICA при помощи H -критерия Краскела – Уоллиса

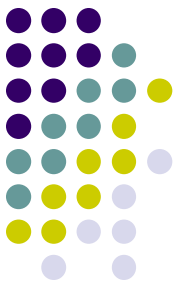


- 2) Выполнить последовательность команд *Статистика* → *Непараметрические данные*
- 3) В окне *Nonparametric Statistics* выбрать *Comparing multiple indep. samples (groups)*

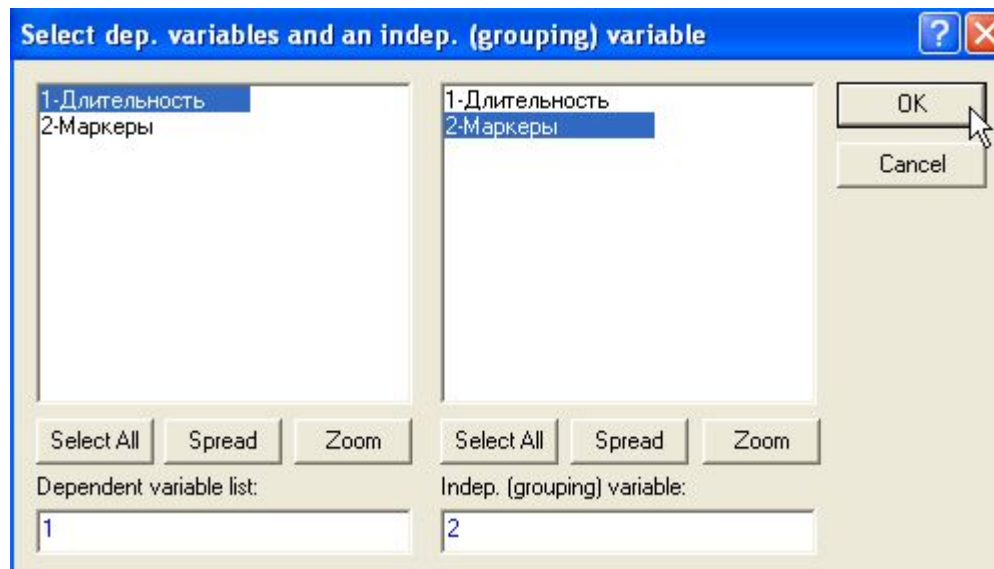


и нажать **ОК**

Проверка гипотез в STATISTICA при помощи *H*-критерия Краскела – Уоллиса



- 4) Нажать кнопку *Variable* и выбрать в левом окне (зависимая переменная) – первую переменную, в правом (группированная переменная) – вторую переменную

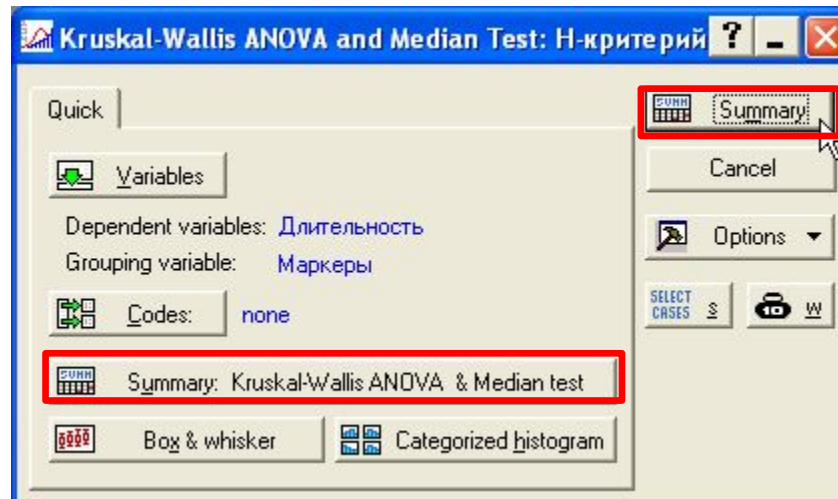


и нажать **ОК**

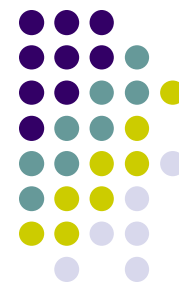
Проверка гипотез в STATISTICA при помощи H -критерия Краскела – Уоллиса



5) Нажать на одну из кнопок *Summary*



Проверка гипотез в STATISTICA при помощи H -критерия Краскела – Уоллиса



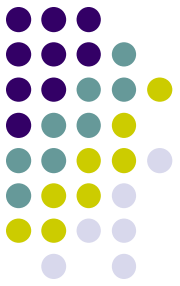
6) Результаты

Depend.:	Code	Valid N	Sum of Ranks
Длительность			
Grp.1	1	4	38,50000
Grp.2	2	8	82,50000
Grp.3	3	6	68,00000
Grp.4	4	4	64,00000

Если значения в этой таблице будут отмечены красным цветом, это будет означать, что нулевая гипотеза H_0 отклоняется

Так как в данном случае нет отмеченных красным результатов, то нулевая гипотеза H_0 принимается на уровне значимости $\alpha=0,05$

Проверка гипотез в STATISTICA при помощи *H*-критерия Краскела – Уоллиса



Вывод: не выявлено статистически значимых различий в четырех группах испытуемых, получивших разные неразрешимые анаграммы. Исследователь может в дальнейшем использовать эти неразрешимыми анаграммами как равнозначные по сложности