

Занятие 7

Непараметрические критерии.

Непараметрические критерии

До сих пор все изученные нами методы предполагали **тестирование гипотез** о **ПАРАМЕТРАХ** популяций на основе выборок.

Мы познакомились с **гипотезами** о:

- ✓ равенстве **среднего** значения чему-то (*t test*);
- ✓ равенстве 2-х **средних** (*t tests for independent & paired samples*);
- ✓ равенстве ≥ 3 **средних** (*ANOVA*);
- ✓ равенстве **дисперсий** (*Levene's test*);
- ✓ равенстве нулю **коэффициента корреляции** (*корреляции*);
- ✓ равенстве нулю **коэффициентов регрессии** (*регрессии*).

Непараметрические критерии

Чтобы тестировать гипотезу о параметре, нужно знать распределение статистики, а чтоб его знать, нужно знать распределение в популяции.

Все изученные тесты требуют **НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ** в популяции и в выборке (без **АУТЛАЕРОВ**).

А ещё в сравнениях групп и в корреляциях/регрессиях требуется **ГОМОГЕННОСТЬ ДИСПЕРСИИ**.

Пренебрежение условиями использования параметрических тестов непредсказуемо увеличивает **вероятность ошибки 1-го рода**!

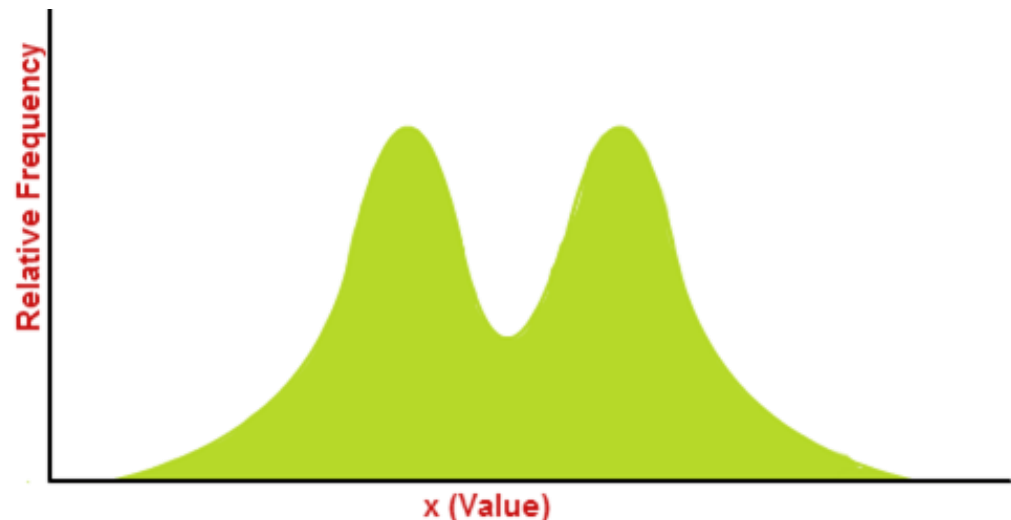
Обычно проблемы можно решить **трансформацией данных**.

Непараметрические критерии

Примечание: слабые отклонения от нормального распределения не очень страшны (в силу Центральной предельной теоремы), а **для больших выборок ими можно пренебречь** (кроме регрессионного анализа).

ANOVA устойчива к отклонениям от нормального распределения, особенно если выборки одинаковы по размеру.

Примечание 2: нет никакого толка трансформировать бимодальное (мультимодальное) распределение, надо искать фактор-причину бимодальности.



Непараметрические критерии

Грубый, но универсальный способ трансформации данных – **РАНГОВАЯ ТРАНСФОРМАЦИЯ** (присвоение рангов каждому наблюдению, от меньшего к большему).

Допустима для:

- ✓ t-критериев,
- ✓ ANOVA (для несложных моделей),
- ✓ корреляций.

Недопустима для:

- ✓ регрессионного анализа.



Конечно, при ранжировании происходит большая потеря информации. И после трансформации выборки должны удовлетворять требованиям, иначе в ней нет смысла!

Непараметрические критерии

Альтернатива ранговой трансформации –

непараметрические методы.

Многие подходят для очень малых выборок!

nonparametric methods = “distribution-free” tests.

- ✓ Свойства распределения неизвестны, **параметры распределения** (среднее, дисперсию и т. п.) **не используются**.
- ✓ Основной подход – **ранжирование** (*ranking*) наблюдений (выстраиваем их по порядку от самого маленького значения к наибольшему, или наоборот).

- ✓ Подразумевается, что **сравниваемые распределения имеют одинаковую форму** (и, иногда, дисперсию).



Непараметрические критерии

В некоторых книжках пишут, что параметрические методы не годятся, если данные **РАНГОВЫЕ**: неизвестно, насколько одно значение отличается от другого.

НО:

It is sometimes counseled that only nonparametric testing may be employed when dealing with ordinal-scale data, but such advice is based upon what Gaito (1980) calls “an old misconception”; this issue is also discussed by Anderson (1961), Gaito (1960), Savage (1957), and Stevens (1968). Interval-scale or ratio-scale measurements are not intrinsically required for the application of parametric testing procedures. Thus parametric techniques may be considered for ordinal-scale data *if* the assumptions of such methods are met—typically, random sampling from normally distributed populations with homogeneity of variances. But ordinal data often come

Zar, 2010; Quinn, Keough, 2002

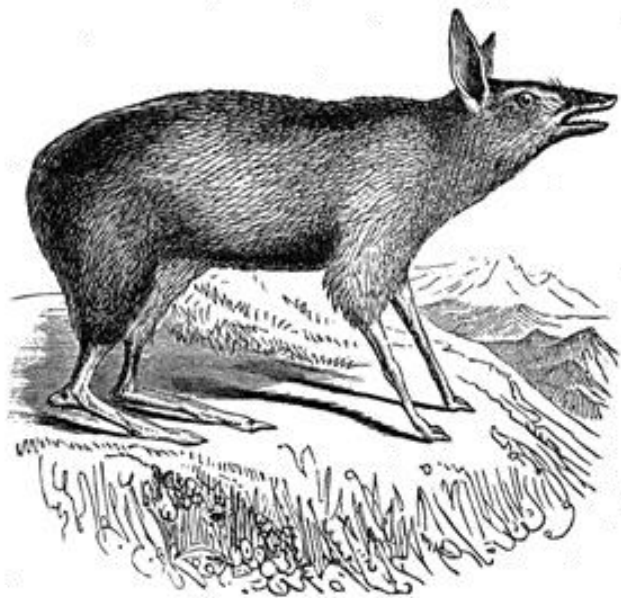
Главное - проверить данные на соответствие нормальности и гомогенности.

Непараметрические критерии

Сравнение 2-х независимых групп (вместо *t*-test for independent samples)

Мы исследуем два вида млекопитающих. Хотим сравнить размеры выводков у этих зверей.

Фактор – вид. Группы: 1. длинноухие; 2. пятнистые
Зависимая переменная – размер выводка



длинноухий



пятнистый

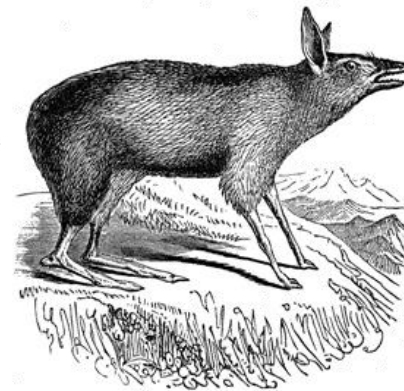
Манн-Уитни тест (*Mann-Whitney U-test*)

H_0 : размер выводка у длинноухих такой же, как и у пятнистых.

H_1 : размер выводка не одинаков у этих видов.

Мы ничего не говорим про параметры распределений!

- Тест Манна-Уитни можно использовать и для ранговых, и для непрерывных переменных;
- Можно использовать для малых выборок (от 5 измерений!).



Непараметрические критерии

длинноухие		пятнистые	
размер	ранг	размер	ранг
8	15.5	4	5
7	13	7	13
4	5	5	8.5
7	13	8	15.5
9	17.5	3	2
3	2	3	2
5	8.5	5	8.5
6	11	4	5
9	17.5		
5	8.5		
111.5		59.5	

Это непараметрический аналог двухвыборочного t-теста.

Ранжируем данные от меньшего к большему (**игнорируя** деление на группы).

Число 3 встретилось трижды (это называется связанные ранги, **tied ranks**):

ранги у них будут одинаковы = $(1+2+3)/3=2$

Непараметрические критерии

Статистика критерия:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$
$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$



n_1 и n_2 – размер выборок,
 R_1 и R_2 – суммы рангов в выборках.

Статистикой критерия U_{obs} будет **меньшее** из этих двух значений.

Причём H_0 мы отвергнем в случае, если оно будет **МЕНЬШЕ** критического значения U_{cv} . (т.е., это исключение среди прочих критериев).

Непараметрические критерии

Если выборки удовлетворяют требованиям для параметрических тестов, мощность теста Манна-Уитни = 95% от мощности t-теста.

М-У тест один из самых мощных среди непараметрических тестов!

Альтернативная процедура – применить **ранговую трансформацию** к исходным данным (т.е., проранжировать измерения как для М-У теста) и провести двухвыборочный t-тест уже над рангами (Zar, 2010, Quinn, Keough, 2002).

Только М-У тест (но не t-тест) пригоден для проверки необычной гипотезы о том, что значения в одной группе отличаются от значений в другой группе в **a** раз (достаточно умножить значения в соответствующей группе на **a** и дальше провести стандартный М-У тест).

Непараметрические критерии

Если размеры выборок больше 20, распределение статистики U приближается к **нормальному** со средним

$$\mu_U = \frac{n_1 n_2}{2}$$

Поэтому считается значение $z_{obs} = \frac{U_{obs} - \mu_U}{\sigma_U}$

И сравнивается с критическим значением для нормального распределения **Z** (наблюдаемое z должно быть по модулю больше критического).



Поэтому для маленьких выборок в статье можно приводить только **U**, а для больших выборок нужно приводить **и U, и z**.

Непараметрические критерии

Сравнение 2-х независимых групп:

✓ **Тест Колмогорова-Смирнова** (*Kolmogorov-Smirnov two-sample test* – не тот, что для проверки на соответствие нормальному распределению): отличается от М-У теста тем, что М-У более чувствителен к различиям средних значений, медианы и т.п., а К-С тест более чувствителен к различиям распределений по форме.

✓ **Тест Вальда-Вольфовица** (*Wald-Wolfowitz Runs Test*) – данные сортируются игнорируя деление на группы и оцениваются последовательности элементов (*runs*) из разных групп. Как и К-С тест, чувствителен к различиям распределений по форме.

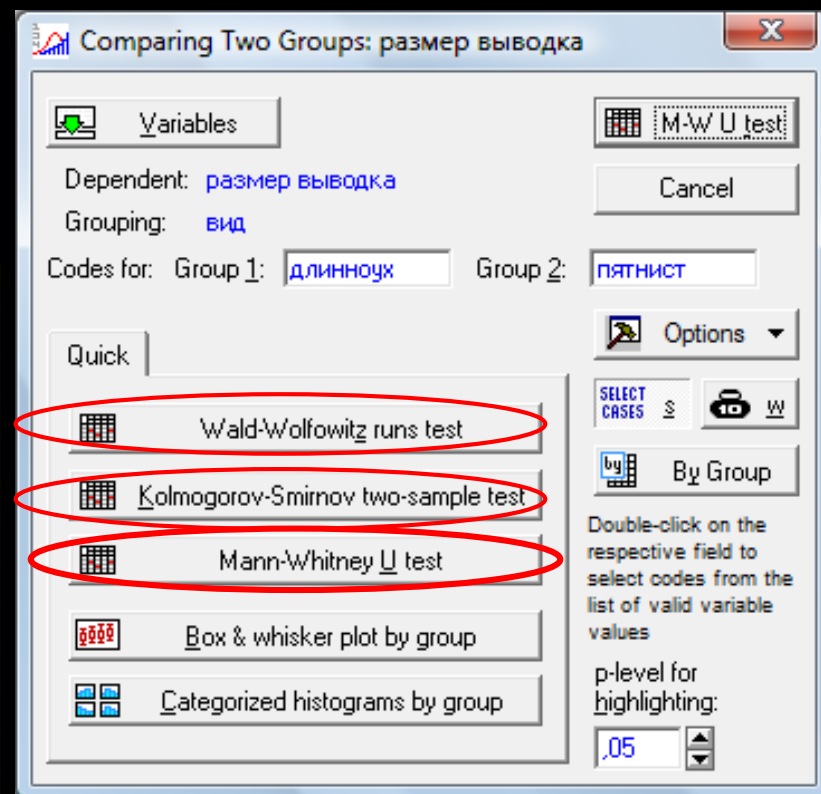
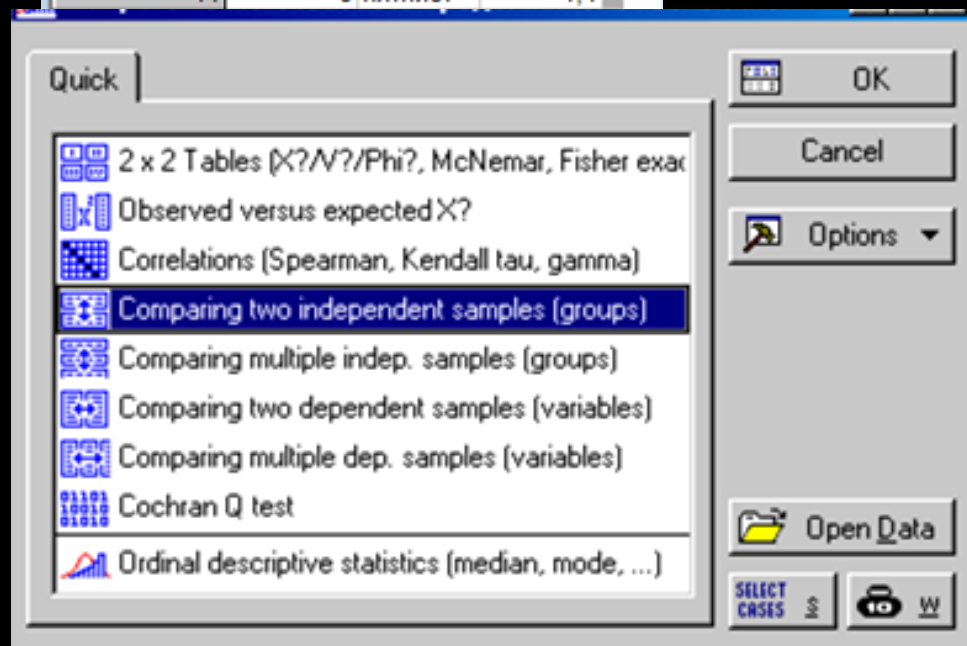
Манн-Уитни тест более мощный, чем эти тесты.

Непараметрические критерии

Mann-Whitney U-test
Kolmogorov-Smirnov two-sample test
Wald-Wolfowitz Runs Test

Data: размер вывода* (3v by 24c)

	1 размер выводка	2 вид	3 пицца, кг
1	5	длинноух	3,4
2	8	длинноух	3,2
3	10	длинноух	2,8
4	6	длинноух	3,1
5	12	длинноух	2,5
6	8	длинноух	1,9
7	9	длинноух	2,4
8	13	длинноух	3
9	14	пятнист	4,2
10	12	пятнист	3,9
11	9	пятнист	2,8
12	10	пятнист	2,6
13	13	пятнист	3,5
14	8	пятнист	4,1



Отвергаем H_0 : М-У тест показал, что размеры выводов у разных видов **неодинаковые**

-Whitney U Test (размер вывода)										
Mann-Whitney U Test (размер вывода) By variable вид Marked tests are significant at $p < .05000$ Exclude condition: $v2=103$										
variable	Rank Sum длинноух	Rank Sum пятнист	U	Z	p-level	Z adjusted	p-level	Valid N длинноух	Valid N пятнист	2*1sided exact p
размер вывода	45,50000	90,50000	9,500000	-2,36297	0,018130	-2,37700	0,017455	8	8	0,014763

Просто результаты
М-У теста

Результаты М-У теста
с поправкой на
наличие tied ranks

Точное значение p для
небольших выборок
без tied ranks

В отличие от К-С и В-В тестов.

Kolmogorov-Smirnov Test (размер вывода)									
By variable вид Marked tests are significant at $p < .05000$ Exclude condition: $v2=103$									
variable	Max Neg Differnc	Max Pos Differnc	p-level	Mean длинноух	Mean пятнист	Std.Dev. длинноух	Std.Dev. пятнист	Valid N длинноух	Valid N пятнист
размер вывода	-0,500000	0,00	$p > .10$	8,875000	12,62500	2,748376	2,386719	8	8

Wald-Wolfowitz Runs Test (размер вывода)										
By variable вид Marked tests are significant at $p < .05000$ Exclude condition: $v2=103$										
Variable	Valid N длинноух	Valid N пятнист	Mean длинноух	Mean пятнист	Z	p-level	Z adjstd	p-level	No. of Runs	No. of ties
размер вывода	8	8	8,875000	12,62500	-0,517549	0,604773	0,258775	0,795809	8	4

Непараметрические критерии

Сравнение 2-х связанных групп (*вместо t -test for independent samples*)

Критерий Вилкоксона (*Wilcoxon matched pair test*)

Типовая задача: сравнение 2-х повторных измерений в одной группе объектов (зверьков, деревьев, ручьёв, ...).

Скажем, мы изучаем утконосов, и хотим знать — различается ли отношение самки к самцу и самца к самке в парах

Мы считаем частоту дружелюбных контактов со стороны самки к самцу и наоборот. У *каждого самца* есть по *жене*, а у каждой *самки* — по *мужу*.

Непараметрические критерии

H_0 : количество контактов в популяции, из которой мы получили выборку самцов, **такое же**, как и в популяции, из которой выборка самок.

H_1 : количество контактов не одинаково.

Фактор – самцы и самки, т.е., пол (связанные парами!)

Зависимая переменная – частота инициирования дружелюбных контактов.



Непараметрические критерии

	самец	самка
1 пара	356	363
2 пара	351	361
3 пара	353	358
4 пара	355	356
5 пара	354	359
6 пара	355	355

$$D_i = X_{i1} - X_{i2}$$

1. Считают разности между значениями в парах;
2. исключают нулевые разности;
3. присуждают абсолютным значениям (по модулю) разностей ранги;
4. суммируют отдельно **ранги** положительных и отрицательных разностей;
5. Наименьшая из этих сумм - статистика **T**.
6. Отвергаем H_0 , если T меньше T_{cv} .

Аналог t-теста для двух связанных выборок, мощность – около 95% мощности t-теста . При числе пар >100 T аппроксимируется нормальным распределением.

В статье приводим: T, N (число пар), p.



Сравнение 2-х связанных групп: Знаковый тест (*Sign test*)

Считают разности в парах, но не ранжируют их, а просто определяют число положительных и отрицательных разностей (нули исключают). Сравнивают их соотношение с 1:1 (биномиальным тестом, см. лекцию 8)

Подходит для случаев, когда точные значения переменной не известны.

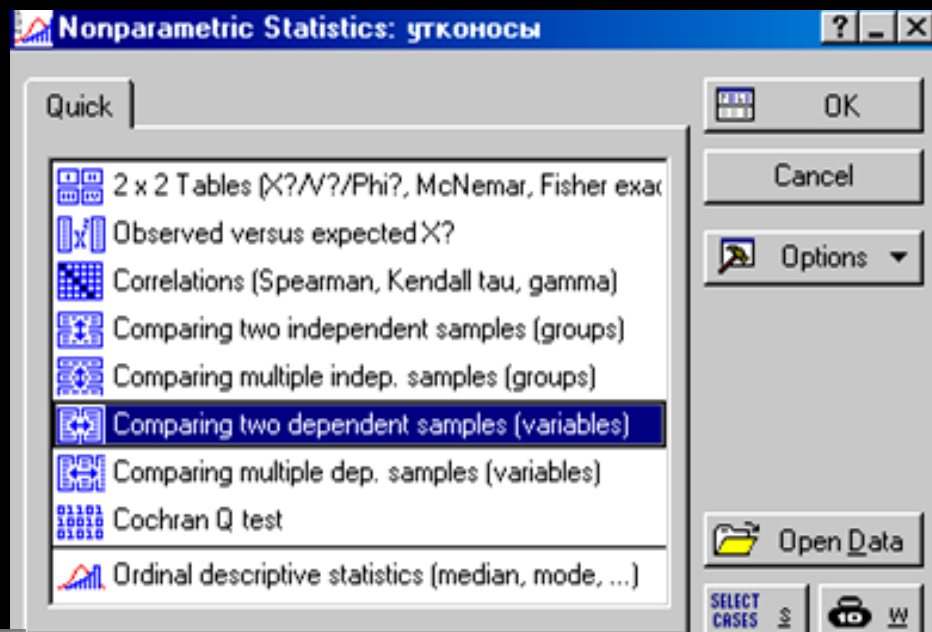
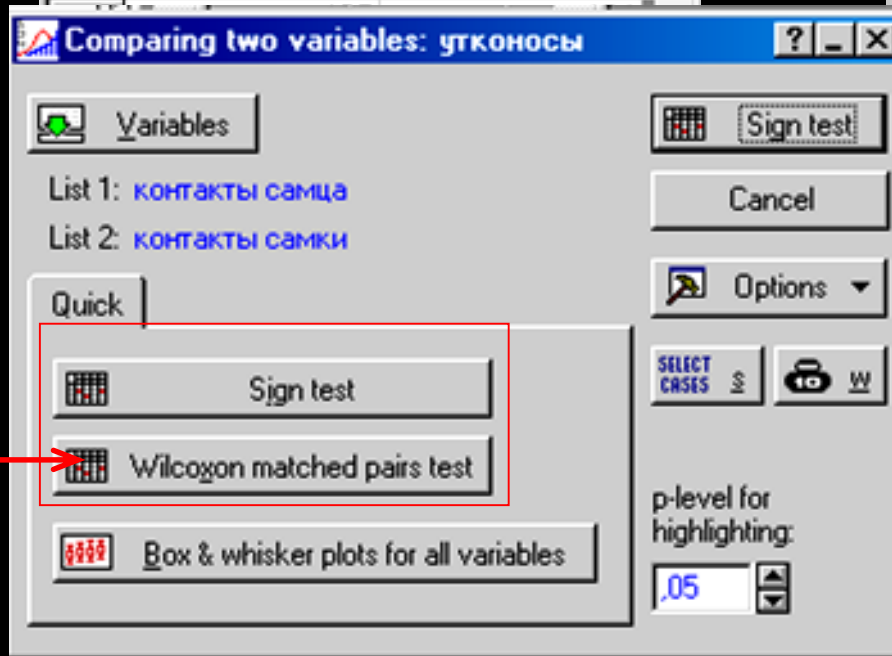
Имеет **низкую мощность**, поэтому применяется только в больших выборках (больше 20 пар).

Непараметрические критерии

Wilcoxon matched pair test

Data: утконосы.sta (2v by 10c)

	1	2
	контакты самца	контакты самки
1	151	78
2	145	85
3	99	73
4	123	91
5	134	85
6	147	79



on Matched Pairs Test (утконосы)

Wilcoxon Matched Pairs Test (утконосы)
Marked tests are significant at p < .05000

	Valid N	T	Z	p-level
са & контакты самки	10	0,00	2,803060	0,005062

Число дружелюбных контактов у самцов и самок в парах было **неодинаковым**

Непараметрические критерии

Сравнение ≥ 3 -х независимых групп (вместо *One-way ANOVA*)

Тест Крускала-Уоллиса (*Kruskal-Wallis test*)

Мы хотим сравнить привлекательность фруктов для 3-х редких видов. Каждому зверю предложили еду, и записали, какую долю от съеденной пищи составили фрукты (3 группы особей; доли – коварные переменные!).

Фактор – вид. 1. длинноухие; 2. пятнистые; 3. хвостатые

Зависимая переменная – доля фруктов от общей массы съеденного.



Критерий Крускал-Уоллиса (*Kruskal-Wallis test*)

- ✓ Непараметрический аналог One-way ANOVA
- ✓ на 95% настолько же мощный, как и ANOVA;
- ✓ для 2-х групп идентичен Манн-Уитни тесту;
- ✓ подразумевает сходство форм распределений и равенство дисперсий в группах (определяется по усатым ящикам и гистограммам!).



Непараметрические критерии

1. все значения ранжируются от меньшего к большему (игнорируя деление на группы);
2. Считается сумма рангов в каждой группе;
3. считается статистика $H(df, N)$.

H_0 : **распределение** в популяциях, из которых мы получили выборки, **одинаковое**.

H_1 : распределения не одинаковые.

$$H = \frac{12}{N(N+1)} \sum \frac{R_j^2}{n_j} - 3(N+1)$$

общий размер выборки

сумма рангов в каждой группе

размер группы

Критерий Крускал-Уоллиса (*Kruskal-Wallis test*)

При маленьких выборках и 3-5-и групп считается H-статистика.

Для больших выборок (или >5-и групп) H аппроксимируется распределением χ^2 .

Это единственный способ (по крайней мере, в Statistica) сравнить между собой несколько групп: невозможно провести многофакторный анализ.

В статье приводим: ***H***, ***N*** (в каждой группе), ***p***.



Непараметрические критерии

Сравнение ≥ 3 -х независимых групп:

Медианный тест (*Median test*)

Считается **общая медиана** для всех групп (получается, что это не непараметрический тест, а distribution-free).

Затем считают, сколько в каждой группе оказалось значений, которые больше и которые меньше общей медианы; соотношения этих чисел (табличка $2 \times k$, где k число групп) сравнивают критерием χ^2 (см. лекция 8).

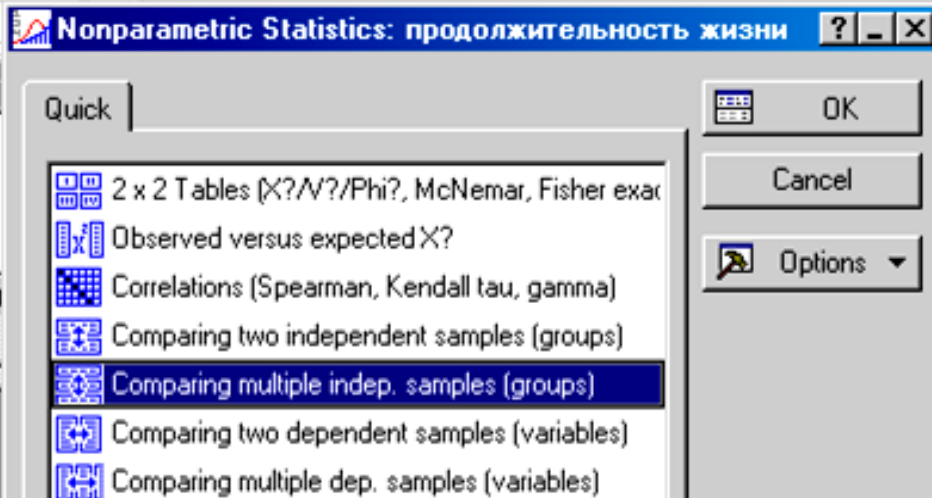
Подходит для выборок, в которых часть наблюдений выходит за пределы шкалы (или их точные значения неизвестны).

Но имеет **очень низкую мощность** – лишь 67% мощности Манн-Уитни теста или теста Крускалла-Уоллеса.
Рекомендуется для групп с $n \geq 20$.

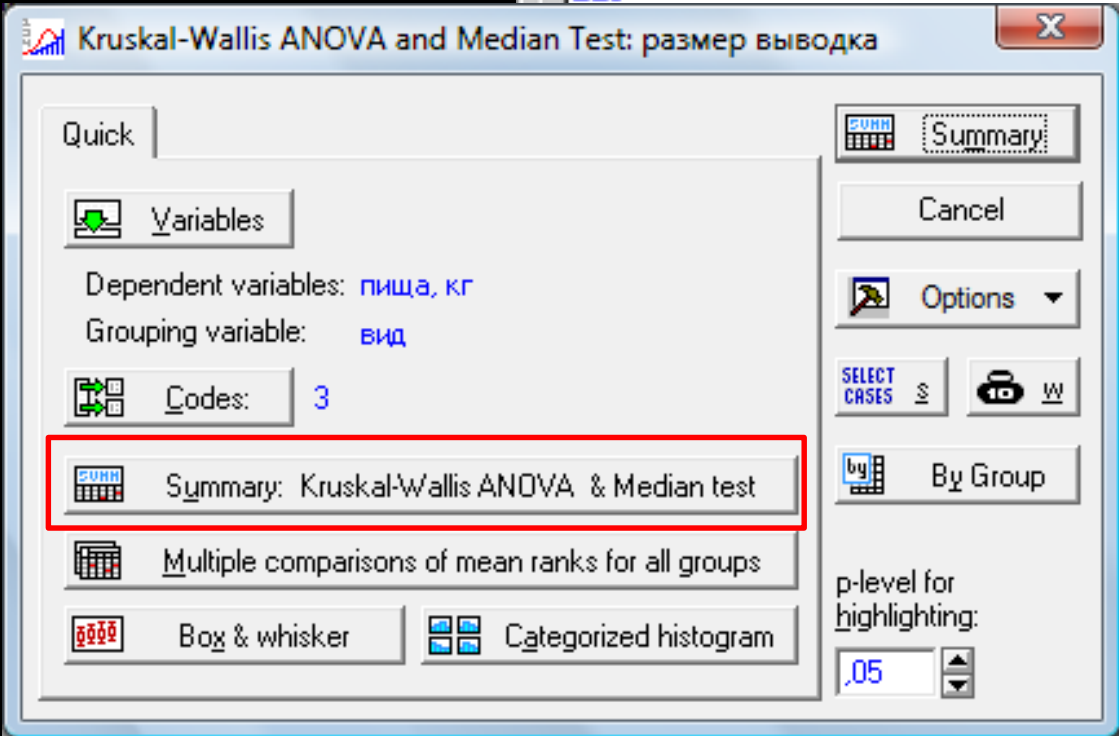
Непараметрические критерии

Kruskal-Wallis test

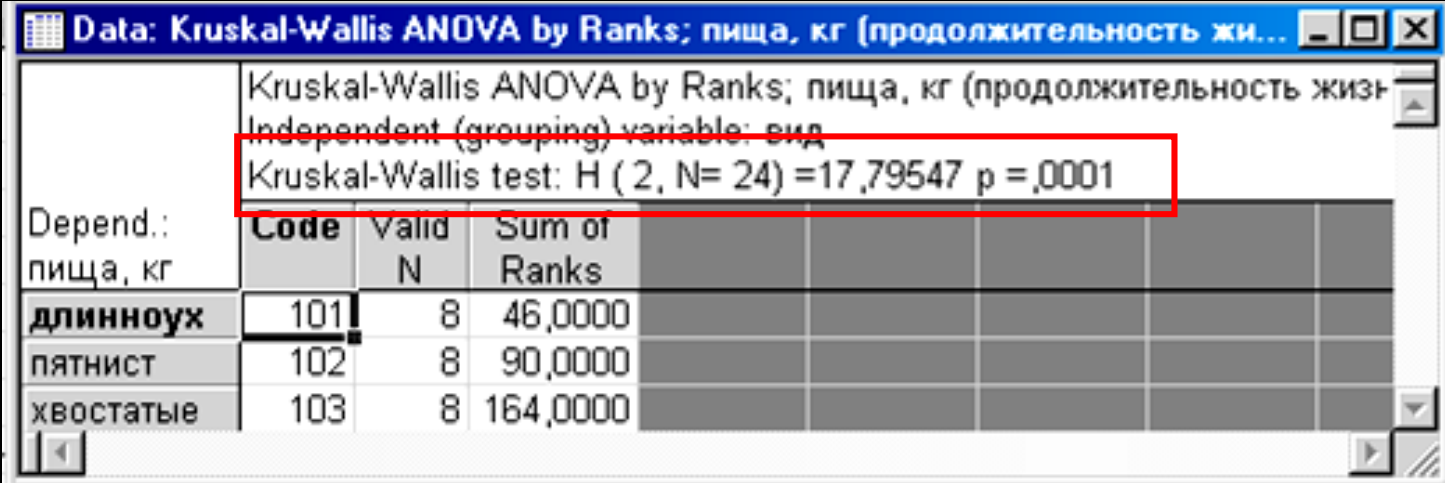
Median test



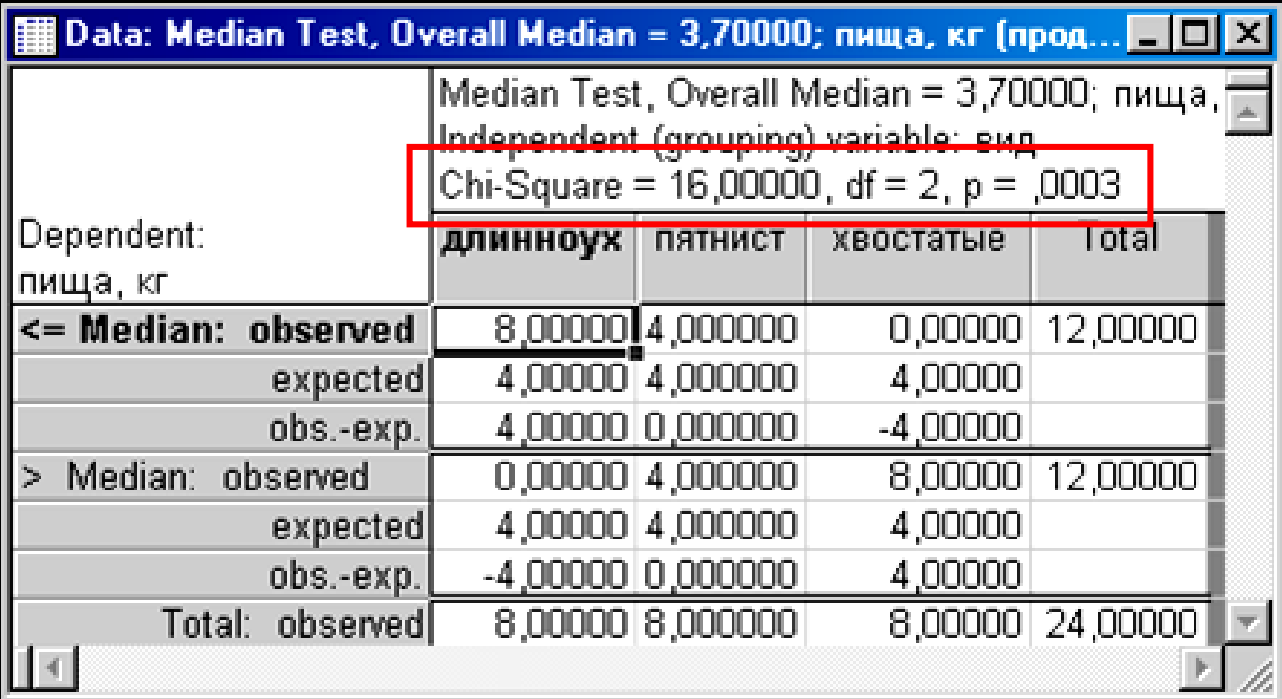
Data: продолжительность...		
	2	3
ос*	вид	пища, кг
20	длинноу>	3,4
21	длинноу>	3,2
34	длинноу>	2,8
51	длинноу>	3,1
19	длинноу>	2,5
24	длинноу>	1,9
37	длинноу>	2,4
40	длинноу>	3
49	пятнист	4,2
50	пятнист	3,9
38	пятнист	2,8
43	пятнист	2,6
39	пятнист	3,5
42	пятнист	4,1
47	пятнист	4
16	46 пятнист	3,4
17	хвостаты	10,2
18	хвостаты	9,8
19	хвостаты	10,5
20	хвостаты	11
21	хвостаты	8,6
22	хвостаты	9,2
23	хвостаты	9,9
24	хвостаты	10



Непараметрические критерии



Доля фруктов отличалась между разными видами



Непараметрические критерии

Если критерий Крускал-Уоллиса показал достоверные различия в группах:

Как и в ANOVA, после сравнения нескольких групп имеет смысл провести **пост-хок тест** (апостериорное сравнение), по аналогии с тестом Тьюки, чтобы выяснить какие же группы различаются.

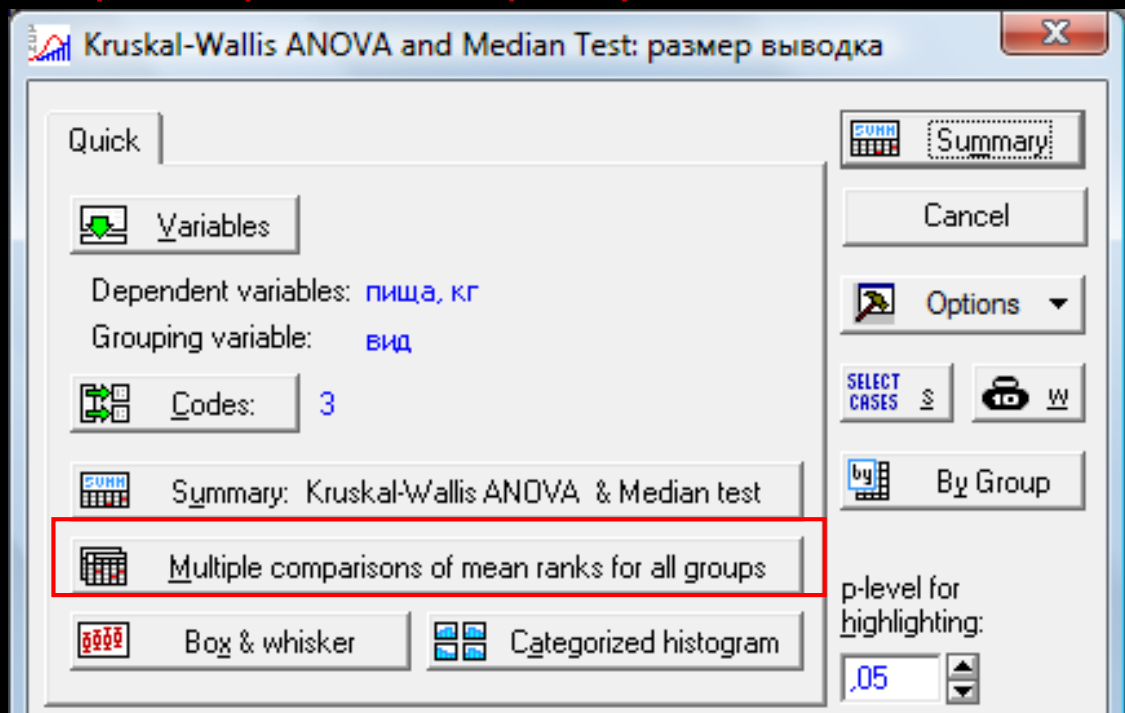
Такие тесты существуют – Nemenyi test, **Dunn's test** (Zar, 1999; 2010).

И они есть в Statistica начиная с 8-й версии!

Теперь мы можем выяснить, какие именно виды отличаются между собой долей съеденных фруктов.



Непараметрические критерии



Пост-хок тест для непараметрической ANOVA

Multiple Comparisons p values (2-tailed); пища, кг (размер вывода)

Depend.: пища, кг

Independent (grouping) variable: вид

Kruskal-Wallis test: $H(2, N=24) = 17,79547$ $p = ,0001$

	длинноух R:5,7500	пятнист R:11,250	хвостатые R:20,500			
длинноух		0,359385	0,000091			
пятнист	0,359385		0,026667			
хвостатые	0,000091	0,026667				

Multiple Comparisons z' values; пища, кг (размер вывода)

Multiple Comparisons p values

Непараметрические критерии

Сравнение ≥ 3 связанных групп (вместо *Repeated measures ANOVA*)

Критерий Фридмана (*Friedman ANOVA*)

У утконосов родились детёныши, и мы хотим знать, изменилась ли упитанность самок после беременности и после выкармливания.

1. состояние до беременности;
2. после появления детей;
3. после выкармливания детёнышей



Непараметрические критерии

Критерий Фридмана (*Friedman ANOVA*)

- ✓ для двух групп эквивалентен Знаковому тесту (sign test);
- ✓ по сравнению с аналогичными параметрическими тестами, для 2-х групп имеет всего 64% мощности, для 3-х – 72%, для 100 стремится к 95%.

Основан на том, что значения ранжируются меньшего к большему **внутри каждой строки**. Потом суммируют ранги для каждого столбца и считают статистику χ^2_r , которая имеет распределение χ^2 .

H_0 : **распределение** в популяциях, из которых мы получили выборки, **одинаковое**.

H_1 : распределения не одинаковые.

В статье приводим: χ^2 , **N** (число особей), **p**

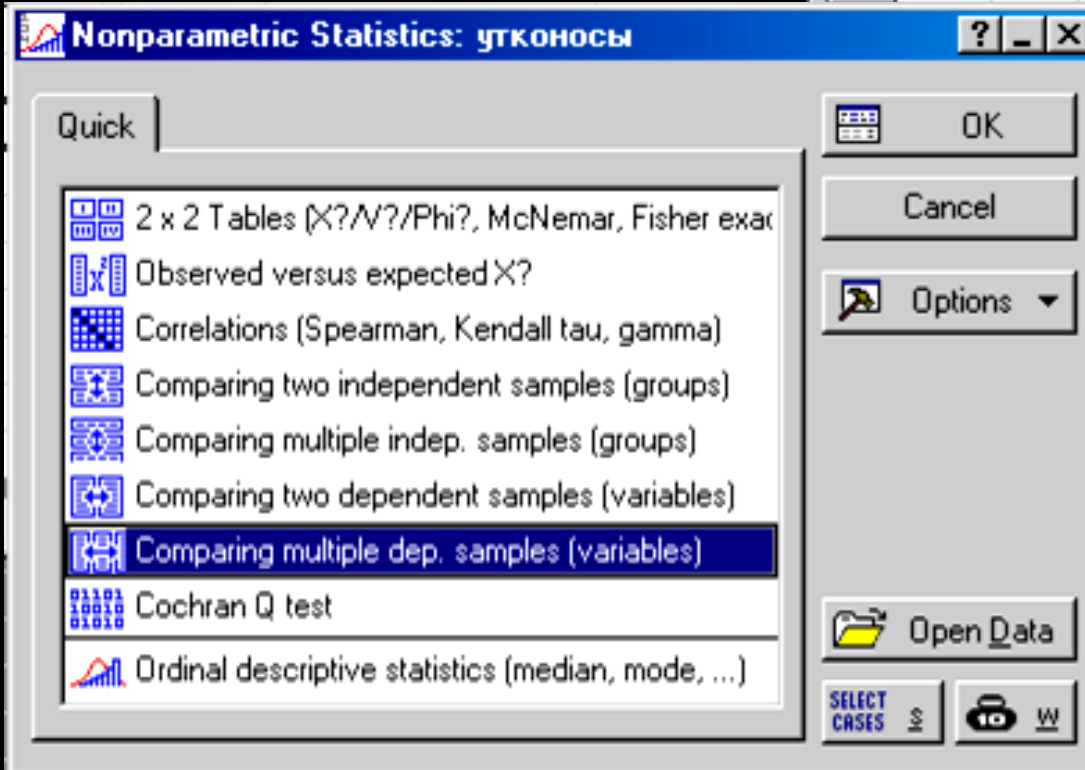


Непараметрические критерии Friedman ANOVA



Data: утконосы* (5v by 10c)

	1	2	3	4	5
	акты с	акты с	состояние самки	после беременности	после выкармли
1	151	78	125	98	
2	145	85	134	102	
3	99	73	129	92	
4	123	91	138	96	
5	134	85	135	100	
6	147	79	139	89	
			137	96	
			140	98	
			138	95	
			131	103	



Непараметрические критерии

Отвергаем H_0 –
состояние самок
изменялось

Friedman ANOVA by Ranks: утконосы

? _ x

Quick

Variables:

3-5

Summary: Friedman ANOVA & Kendall's concordance

Box & whisker plot for all variables

Summary

Cancel

Options

SELECT CASES

W

Data: Friedman ANOVA and Kendall Coeff. of Concordance (утконосы)						
Friedman ANOVA and Kendall Coeff. of Concordance (y1)						
ANOVA Chi Sqr. (N = 10, df = 2) = 20,00000 p < ,00005						
Coeff. of Concordance = 1,0000 Aver. rank r = 1,0000						
Variable	Average Rank	Sum of Ranks	Mean	Std.Dev.		
масса самки до	3,000000	30,00000	134,6000	4,880801		
после беременности	2,000000	20,00000	96,9000	4,306326		
после выкармливания	1,000000	10,00000	52,3000	4,056545		

Ранговые корреляции (вместо параметрической корреляции)

Трансформация данных в регрессионном анализе и корреляциях:

Применяется таким же образом, как и для других критериев; основанием для применения должны служить несоответствие нормальному распределению и гетерогенность дисперсий, а **не нелинейность** связи:

если распределения нормальны и дисперсии гомогенны, бессмысленно использовать трансформацию данных для получения линейной регрессии из нелинейной: нарушится нормальность/гомогенность.

Непараметрические критерии

Коэффициент корреляции **Спирмана** (*Spearman rank order correlation*)

Позволяет оценить силу **ПАРНОЙ ЛИНЕЙНОЙ** связи – насколько совместно изменяются ДВЕ переменные.

Связана ли дистанция расселения с индексом упитанности у мышей?

Переменные – 1. дистанция расселения от выводковой норы; 2. индекс упитанности



Непараметрические критерии

1. Ранжируем данные для **каждой переменной** от меньшего к большему;
2. Если встретились одинаковые значения (***tied ranks***), присваиваем им средние ранги;
3. Считаем разности рангов в **каждой паре** данных;
4. Считаем коэффициент r_s

$$r_s = 1 - \frac{6 \sum D_i^2}{n(n^2 - 1)}$$

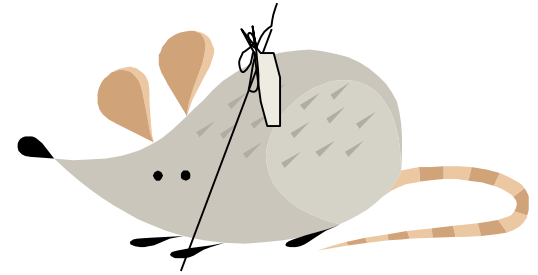
разности
рангов

число строк
(размер
выборки)

Непараметрические критерии

$$H_0 : \rho_s = 0$$

$$H_1 : \rho_s \neq 0$$



Статистика критерия – сам коэффициент корреляции Спирмана (имеет t-распределение)

Коэффициент **Спирмана** – аналог **коэффициента корреляции Пирсона**, стремится к нему в больших выборках. Мощность – около 91% коэффициента Пирсона. Знак показывает направление связи, абсолютное значение – силу связи. От -1 до +1.

Лучший для дробных количественных признаков. Размер выборки ≥ 10 .

В статье приводим: r_s , N , p .



Непараметрические критерии

Spearman Rank Order Correlations

Data: расселение мышей.sta (10v by 20c)			
	1	2	3
	№ мыши	дистанция, м	упитанность
1	1	96,00	15
2	2	43,00	8
3	3	67,05	11
4	4	40,00	12
5	5	55,93	10
6	6	102,00	19
7	7	42,00	5
8	8	38,00	12
9	9	35,00	11
10	10	41,00	10
11	11	40,00	10
12	12	15,00	9
13	13	14,00	8
14	14	78,00	6
15	15	58,00	7
16	16	49,00	11
17	17	52,88	15
18	18	120,00	17
19	19	25,00	10

Nonparametric Statistics: расселение мышей

Quick

2 x 2 Tables (X²/V²/Phi², McNemar, Fisher exact)

Observed versus expected X²

Correlations (Spearman, Kendall tau, gamma)

Comparing two independent samples (groups)

Comparing multiple indep. samples (groups)

Comparing two dependent samples (variables)

Comparing multiple dep. samples (variables)

Cochran Q test

Ordinal descriptive statistics (median, mode, ...)

OK

Cancel

Options

Open Data

SELECT CASES

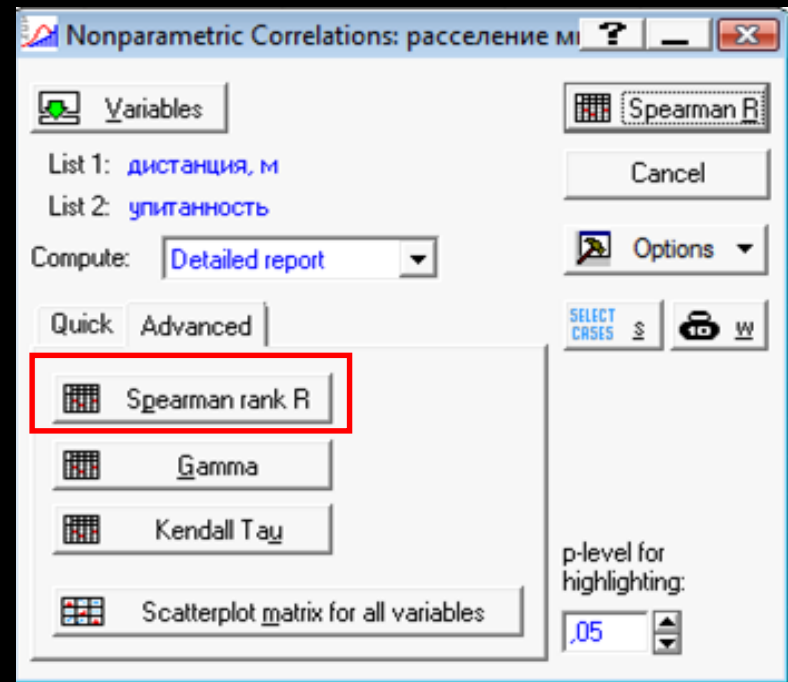
W

Непараметрические критерии

Spearman Rank Order Correlations

Отвергаем H_0 :

Оказалось, что дистанция расселения положительно связана с упитанностью у мыши.



Spearman Rank Order Correlations (расселение мышей)

MD pairwise deleted

Marked correlations are significant at $p < .05000$

Pair of Variables	Valid N	Spearman R	t(N-2)	p-level
дистанция, м & упитанность	20	0.573795	2.972419	0.008160

Correlations (расселение мышей 10v*20c) | Spearman Rank Order Correlations (расселение мышей) | Com

Непараметрические критерии

Коэффициент корреляции Кендалла (*Kendall's coefficient of rank correlation, Kendall- τ*)

Он оценивает разность между вероятностью того, что порядок данных в обеих переменных одинаков, и вероятностью того, что порядки разные.

Считается совсем не так, как коэффициент Спирмана (совсем другой математический принцип).

Оценивает парную корреляцию. Мощность его НИЖЕ, чем у коэф. Спирмана.



Дистанция
расселения -
упитанность



Только для **ранговых** переменных! Для количественных лучше коэффициент Спирмана, особенно для больших выборок

Непараметрические критерии

Гамма-статистика (*Gamma*)

Почти как коэффициент корреляции Кендалла (тоже оценивает вероятность совпадений)

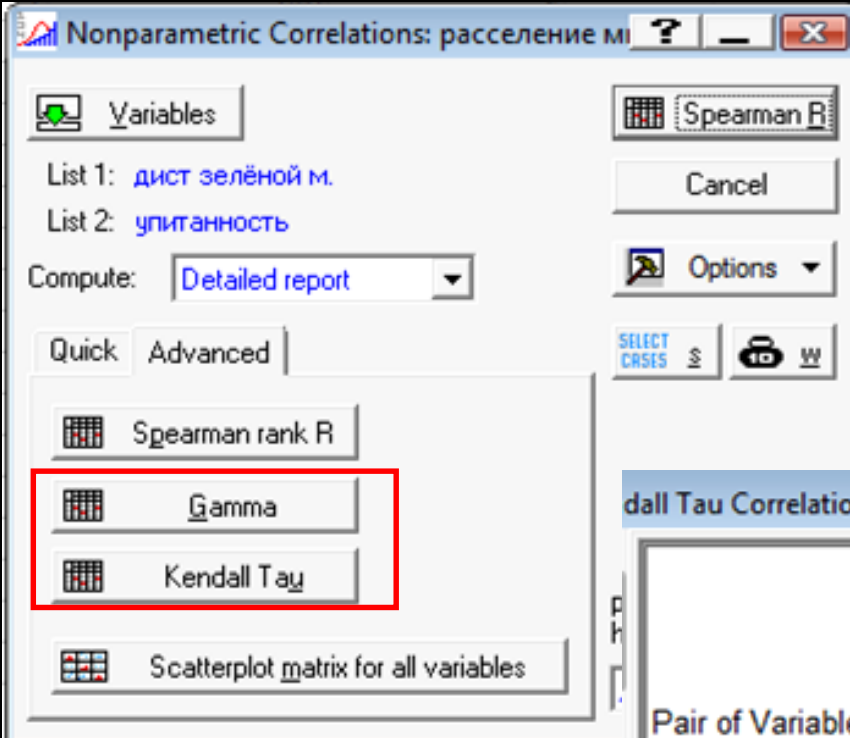
Её лучше использовать, если в выборке много **совпадающих значений** (tied ranks) - она не так страдает от них.

Мощность НИЖЕ, чем у коэффициента Спирмана.



Непараметрические критерии

Отвергаем H_0 : дистанция расселения у зелёных мышей отрицательно связана с упитанностью.



Data: расселение мышей.sta (10v by 20c)

	5	6	7
	№ зел мыши	дист зелёной м.	упитанность
1	1	122	23
2	2	245	9
3	3	101	17
4	4	287	7
5	5	98	31
6	6	50	36
7	7	264	10
8	8	175	13
9			

Kendall Tau Correlations (расселение мышей)

Kendall Tau Correlations (расселение мышей)
MD pairwise deleted
Marked correlations are significant at p < .05000

Pair of Variables	Valid N	Kendall Tau	Z	p-level	p-exact 1-tailed
дист зелёной м. & упитанность	8	-0,857143	-2,96923	0,002985	,001



В статье: τ , N, p

Коэффициент конкордантности Кендалла (*Kendall's coefficient of concordance*)

Оценка **корреляции** между ≥ 3 ранговыми переменными. Значения ранжируются внутри каждой переменной, и считается общая статистика W , она примерно соответствует среднему коэффициенту корреляции Спирмана для всех пар переменных.

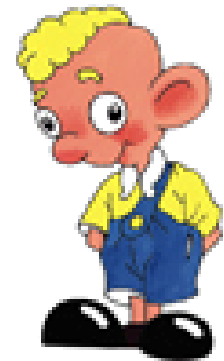
Важно: в Statistica переменные должны быть в строках, а не в столбцах!



Петя



Гриша



Вася

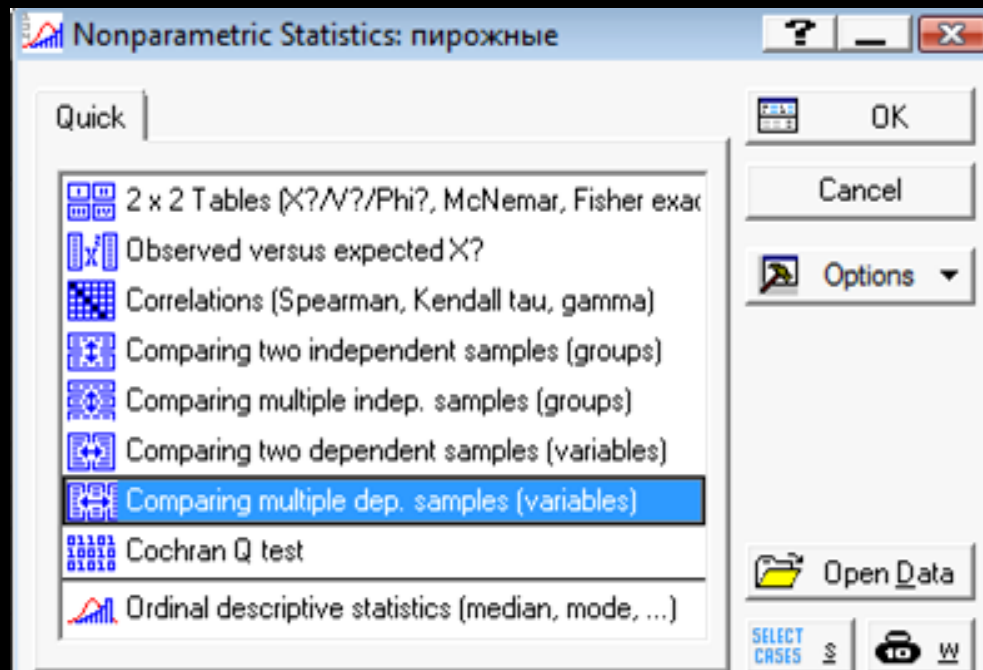
Одна из задач – оценка согласия экспертов, например, детей, оценивающих вкус 6 типов пирожных. Не даёт ответ о парных связях переменных, только **общая корреляция**.

Непараметрические критерии

Коэффициент конкордантности Кендалла

Data: пирожные* (7v by 3c)							
	1 ребёнок	2 эклер	3 заварное	4 картошка	5 орешек	6 корзинка	7 трюфель
1	Петя	1	3	2	4	5	6
2	Гриша	1	2	3	5	4	6
3	Гурвинек	1	3	4	3	6	5

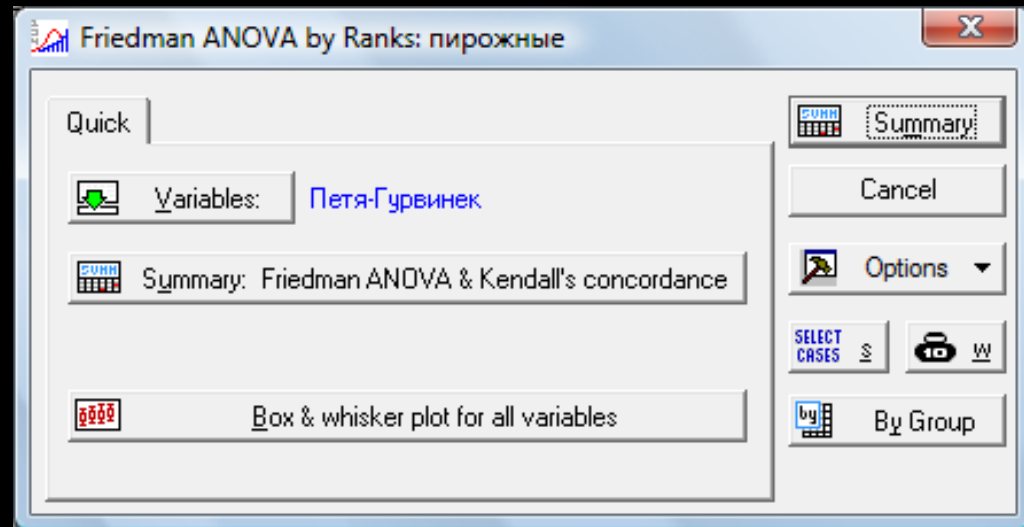
Нестандартная организация таблицы!



$$0 \leq W \leq 1$$

Чем ближе коэффициент к 1, тем выше корреляция.

Чем ближе к нулю, тем меньше связь переменных (например, согласие экспертов).



Friedman ANOVA and Kendall Coeff. of Concordance (пирожные)

Friedman ANOVA and Kendall Coeff. of Concordance (пирожные)
ANOVA Chi Sqr (N = 3, df = 5) = 12,59615 p = ,02747
Coeff. of Concordance = ,83974 Aver. rank r = ,75962

Variable	Average Rank	Sum of Ranks	Mean	Std.Dev.			
эклер	1,000000	3,00000	1,000000				
заварное	2,500000	7,50000	2,666667	0,577350			
картошка	3,000000	9,00000	3,000000	1,000000			
орешек	3,833333	11,50000	4,000000	1,000000			
корзинка	5,000000	15,00000	5,000000	1,000000			
трюфель	5,666667	17,00000	5,666667	0,577350			

Итак, при выборе теста важно, что:

1. **Параметрические** тесты более **мощные**, чем непараметрические;
2. Возможности анализа взаимосвязей переменных у параметрических тестов несравнимо выше;
3. Непараметрические безопаснее в плане ошибки 1-го рода;
4. Чем больше размер выборки, тем менее критичны требования к распределению (по Центральной предельной теореме); для выборок $N \geq 100$ используют параметрические тесты даже при больших отклонениях от нормального распределения (кроме регрессий).
5. АНОВА не очень чувствительна к отклонениям от нормального распределения (особенно для одинаковых по размеру групп).
6. **Надо бороться за использование параметрического анализа в исследованиях!**

Дополнение на всякий случай

Сравнение 2-х индексов Шеннона

Считают t-статистику (Zar, 2010, p. 174)

$$t = \frac{H'_1 - H'_2}{s_{H'_1 - H'_2}},$$

Где

$$s_{H'_1 - H'_2} = \sqrt{s_{H'_1}^2 + s_{H'_2}^2}.$$

И сравнивают с табличным t_{cv}

$$s_{H'}^2 = \frac{\sum f_i \log^2 f_i - (\sum f_i \log f_i)^2 / n}{n^2}$$

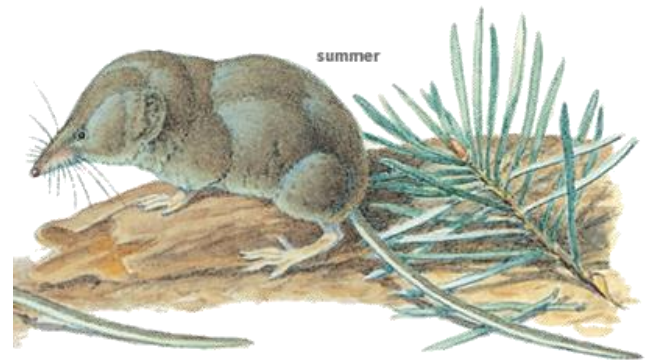
$$df = \nu = \frac{(s_{H'_1}^2 + s_{H'_2}^2)^2}{\frac{(s_{H'_1}^2)^2}{n_1} + \frac{(s_{H'_2}^2)^2}{n_2}}$$

Самые основные аналитические задачи и тесты, которые мы изучили.

1. Сравнение среднего значения в популяции с заданным числом.

$$H_0: \mu = a;$$

$$H_1: \mu \neq a$$



Например, со значением из справочника.

Одновыборочный t-критерий Стьюдента (one-sample t-test)

Статистика - t

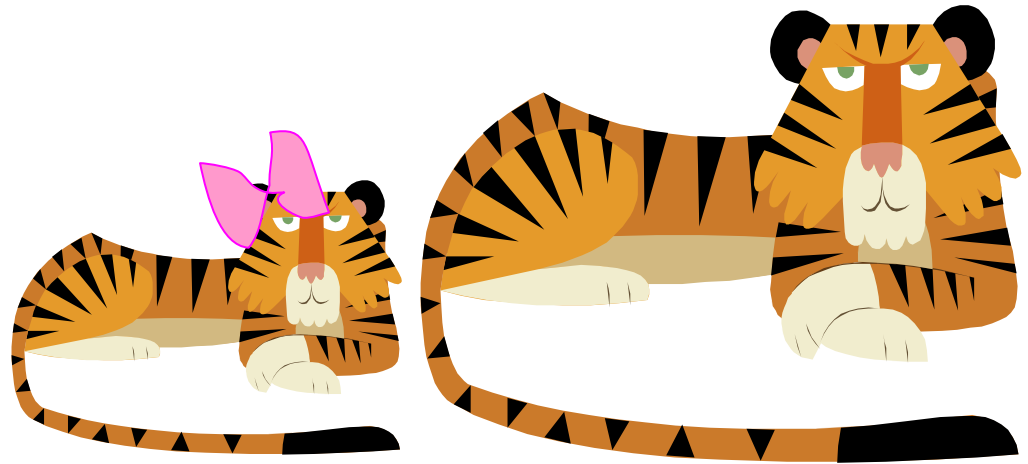
2. Сравнение между собой **средних значений 2-х независимых** выборок.

Зависимая переменная — количественная.

Независимая (фактор) — определяет нахождение в той или иной группе.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$



Двухвыборочный t-критерий Стьюдента (two-sample t-test),
статистика - t

Непараметрический аналог — тест Манна-Уитни

3. Сравнение между собой **средних значений 2-х СВЯЗАННЫХ** выборок.

Зависимая переменная – количественная.

Независимая (фактор) – уровни = повторные измерения.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$



Критерий Стьюдента для связанных выборок (t-test for dependent samples); статистика - t

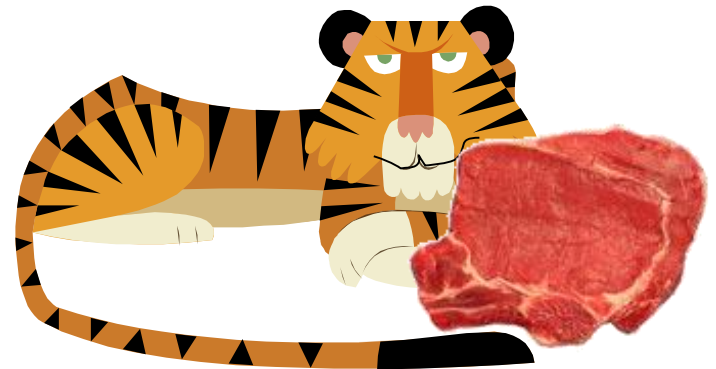
Непараметрический аналог – тест Вилкоксона

4. Сравнение между собой **средних значений ≥ 3 -х независимых** выборок.

Зависимая переменная – количественная.

Независимая (фактор) – определяет нахождение в группе.

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$



Однофакторный дисперсионный анализ (One-way ANOVA);
статистика - F

Непараметрический аналог – тест Краскал-Уоллиса

5. Сравнение средних значений ≥ 3 -х независимых выборок: действие **нескольких факторов**.

Зависимая переменная — количественная.

Независимые (факторы) — их ≥ 2 ;
определяют нахождение в группах.

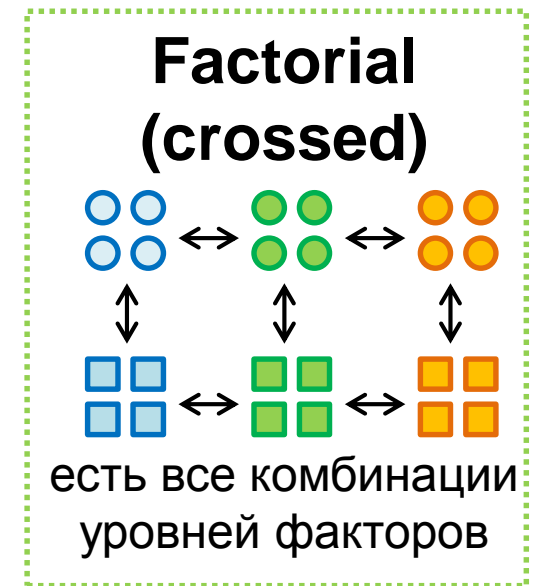
$$H_{0_1} : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_{0_2} : \mu_1 = \mu_2 = \mu_3$$

H_{0_3} : нет **взаимодействия** между факторами.

Многофакторный дисперсионный анализ (Factorial ANOVA); статистика - F

Непараметрический аналог — нет



5. Сравнение средних значений ≥ 3 -х связанных выборок.

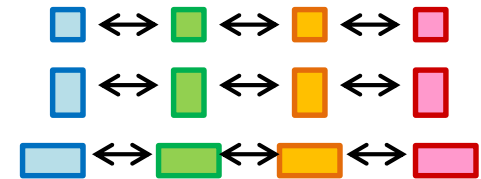
Зависимая переменная – количественная.

Независимые (факторы) – их ≥ 2 ; уровни = повторные измерения.

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$



Repeated measures



Уровни **K** факторов – связанные
выборки+1 response+interactions

Дисперсионный анализ связанных групп (Repeated measures ANOVA); статистика - F

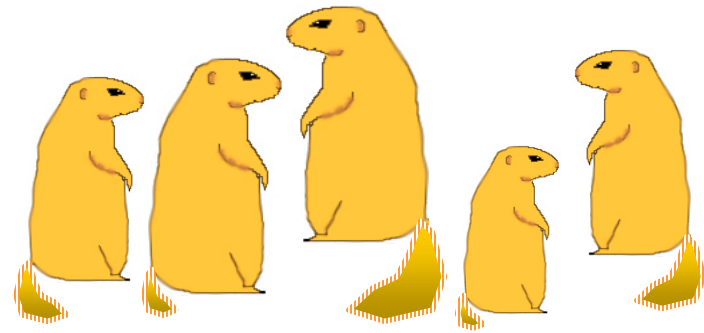
Непараметрический аналог – тест Фридмана

6. Оценка, в какой степени **две** переменные **совместно** **изменяются**.

Зависимая и независимая переменные –
количественные и неотличимы одна от другой

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$



Коэффициент корреляции Пирсона (Pearson product-moment correlation coefficient r); статистика – $t(r)$

Непараметрический аналог – коэффициент корреляции Спирмана

7. Оценка влияния одной количественной переменной на другую.

Зависимая переменная – количественная.

Независимая – количественная.

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$



Простая линейная регрессия (linear regression);

статистика – t , F

Непараметрический аналог – нет (есть, но это сложная тема)

8. Оценка влияния нескольких количественных переменных на другую переменную.

Зависимая переменная – количественная.

Независимые – ≥ 2 , количественные.

$$H_o : \beta_1 = \beta_2 = \dots = \beta_m = 0$$

$$H_{01} : \beta_1 = 0$$

$$H_{12} : \beta_2 \neq 0$$



Множественная линейная регрессия (multiple linear regression); статистика – t, F

Непараметрический аналог – нет (только множественная корреляция, коэффициент конкордантности Кенделла)

Задачи на выбор критерия

1. В институте каждый год проводят экзамен по статистике, и средний результат = 80 баллов. Преподаватель решил удвоить количество домашних заданий и посмотреть, каким станет результат.

- ✓ Какими будут нулевая и альтернативная гипотезы?
- ✓ Статистический критерий?

2. зоолог обнаружил две изолированные популяции белок – северную и южную. Ему кажется, что в северной популяции масса белок больше. Он хочет проверить своё предположение статистически.

- ✓ Какими будут нулевая и альтернативная гипотезы?
- ✓ Какой статистический тест учёный будет использовать для анализа?

3. Д-р Симонс решил узнать, как утренняя зарядка влияет на самочувствие людей. Он померил ЧСС у 52 человек и заставил их 8 недель каждое утро делать зарядку, после чего померил ЧСС вновь.

- ✓ Какими будут нулевая и альтернативная гипотезы?
- ✓ Какой статистический тест доктор будет использовать для анализа?

4. крыс обучают ставить лесенку и доставать корм. В 3-х группах крыс используют разный корм: сыр, морковь и мясо. Затем часть крыс из каждой группы помещают в знакомую среду, а часть – в новые условия. Исследователь хочет узнать, как зависит время добычи корма: 1) в разных группах; 2) в разных условиях.

✓ H_0 ? Тип статистического анализа? Статистика критерия?

5. мы хотим изучить, как уровень глюкозы в крови у кошек зависит от времени суток. У нас есть 10 особей, мы взяли у каждой кровь утром, днём и вечером.

✓ H_0 ? Тип статистического анализа? Статистика критерия?

6. телефонная компания поставила новые телефоны в аэропорты и размышляет, стоит ли оборудовать места для разговоров креслом – не увеличит ли это продолжительность разговоров. В 15 местах она поставила кресла, в 15-и местах говорить можно только стоя.

✓ Как проверить, влияет ли наличие кресла на длительность разговора?
 H_0 ? Тип статистического анализа? Статистика критерия?

7. исследователь решил узнать, как зависит площадь дома у семьи от дохода семьи (в год). Собрал данные от 50 семей.

✓ H_0 ? Статистический критерий? Как изменился бы результат теста, если бы доходы каждой семьи были больше на 5000\$ в год ?

8. педиатры изучают прибавку в весе у младенцев (её оценивают как разницу в массе ребёнка в 2 мес и при рождении). При этом, в их выборке есть дети, которые вскармливаются искусственно, а есть те, которые находятся на грудном вскармливании. Кроме того, некоторые матери кормят младенцев по требованию, другие же – строго по расписанию.

✓ Как узнать, влияют ли тип пищи и распорядок вскармливания на прибавку в весе? H_0 ? Статистический критерий?

9. Д-р Стевия работает в госпитале и хочет узнать, как у её пациентов уровень сахара в крови зависит от длительности сна, температуры тела и длительности прогулок в день.

✓ H_0 ? Статистический критерий?

К практическому занятию

1. Тест Манна -Уитни – Aggressn
2. Kruskal-wallis ANOVA – crabs, Gators
3. synchron – тест Вилкоксона и знаковый тест
4. Friedman ANOVA – Cat clinic
5. корреляции – cat clinic measurements,
6. Kendall конкордантность - Mothers.