

УДК 519.23

СРАВНЕНИЕ ДВУХ НЕСВЯЗАННЫХ ВЫБОРОК С ИСПОЛЬЗОВАНИЕМ ПАКЕТА СТАТИСТИЧЕСКИХ ПРОГРАММ STATA: НЕПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

© 2014 г. ¹О. А. Харьков, ^{1,2}А. М. Гржибовский¹Северный государственный медицинский университет, г. Архангельск²Норвежский институт общественного здравоохранения, г. Осло, Норвегия

В статье представлены теоретические основы применения статистических критериев для непарных выборок при распределении данных, отличающемся от нормального, и практические рекомендации для применения вышеуказанных критериев с использованием пакета прикладных статистических программ STATA. Приведены практические примеры расчетов с помощью формул, а также алгоритм действий в использовании пакета STATA для применения критериев Манна – Уитни и Вилкоксона для непарных выборок с последующей интерпретацией результатов.

Ключевые слова:

непараметрические критерии, критерий Манна – Уитни, критерий Вилкоксона для независимых групп

В предыдущем выпуске журнала рассматривалось сравнение одной и двух несвязанных групп с помощью критерия Стьюдента. Этот критерий основан на допущении, что изучаемый признак подчиняется нормальному распределению [1] и сравниваемые совокупности имеют равные дисперсии. Однако эти условия не всегда выполняются: в одних случаях слишком велика разница дисперсий, в других распределение признака отличается от нормального. Кроме того, изучаемая переменная может быть представлена не только количественной шкалой, но и порядковой. В таком случае корректнее применять непараметрические методы анализа данных.

Непараметрические методы – это методы, не требующие какого-либо определенного распределения данных [5]. Здесь не используются параметры генеральной совокупности (например, среднее арифметическое и стандартное отклонение, как для критериев, требующих нормального распределения), а осуществляется ранжирование абсолютных значений, что позволяет нивелировать эффект выскакивающих величин и скошенности распределения [3]. При нормальном распределении генеральной совокупности параметрические методы обладают большей мощностью по сравнению с непараметрическими. Поэтому если сравниваемый количественный признак имеет нормальное распределение, то следует отдавать предпочтение параметрическим критериям. Однако если распределение отличается от нормального, то использование параметрических методов приводит как минимум к сомнительным, а то и ошибочным результатам.

Критерий Манна – Уитни и двухвыборочный критерий Вилкоксона являются непараметрическими аналогами критерия Стьюдента для несвязанных (непарных) выборок, поэтому условия применения следующие:

- количественный или порядковый признак;
- не требуется проверка на нормальность распределения;
- независимость сравниваемых выборок;
- две группы.

Для проверки соблюдения условий применения непараметрических критериев не требуется никаких математических манипуляций. Независимость двух сравниваемых выборок определяется типом исследования [2] и предполагает наличие двух взаимоисключающих групп, например группы мужчин и группы женщин, группы имеющих и группы не имеющих изучаемого заболевания, и т. п.

В действительности расчеты критериев Манна – Уитни и двухвыборочного Вилкоксона схожи, поэтому в литературе чаще можно видеть расчет критерия Манна – Уитни, хотя во многих макетах статистических программ эти два критерия представлены раздельно.

Как было сказано ранее, в основе расчета непараметрических критериев лежит ранжирование. Рассмотрим пример. У беременных женщин с отсутствием (1 группа) и наличием (2 группа) изучаемой патологии беременности изучался показатель личностной тревожности по методике Спилбергера – Ханина:

1 группа:	23	31	27	28	27	39	21	40	35
2 группа:	30	49	32	26	52	36	26	50	

Если теоретически предположить, что женщины с патологией беременности имеют более высокий показатель личностной тревожности по сравнению с женщинами без патологии течения гестационного периода, то мы будем ожидать, что более высокие ранги будут наблюдаться во второй группе, а более низкие – в первой. Критерий Манна – Уитни и двухвыборочный критерий Вилкоксона работают по такому же принципу [6]. Теперь посмотрим, как это работает на практике. Табл. 1 показывает, как осуществляется ранжирование.

Таблица 1

Балл	21	23	26	26	27	27	28	30	31	32	35	36	39	40	49	50	52
Ранг	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Истинный ранг	1	2	3,5	3,5	5,5	5,5	7	8	9	10	11	12	13	14	15	16	17
Группа	1	1	2	2	1	1	1	2	1	2	1	2	1	1	2	2	2
Сумма рангов	1 группа: 68								2 группа: 85								

Алгоритм расчета двухвыборочного критерия Вилкоксона:

1. Ранжировать значения переменной в порядке возрастания (или убывания) независимо от принадлежности к группе.

2. Заменить абсолютные значения переменной рангами, как показано в табл. 2. В случае если несколько значений равны (например, по табл. 1 это значения «26» и «27»), то им присваивается ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны (например, в отношении значения «26» – $(3 + 4) / 2 = 3,5$; «27» – $(5 + 6) / 2 = 5,5$).

3. Подсчитать сумму рангов отдельно для каждой группы.

4. Наименьшая сумма рангов является абсолютным значением критерия Вилкоксона.

5. Далее рассчитаем z-значение для двухвыборочного критерия Вилкоксона и уровень статистической значимости:

$$\bar{W}_s = \frac{n_1(n_1 + n_2 + 1)}{2}; SE_{W_s} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}; z = \frac{W_s - \bar{W}_s}{SE_{W_s}}$$

где W_s – значение критерия Вилкоксона; \bar{W}_s – среднее значение критерия Вилкоксона; SE_{W_s} – стандартная ошибка; n_1 – количество наблюдений

в первой группе; n_2 – количество наблюдений во второй группе.

$$W_s = 68; \bar{W}_s = \frac{9 * (9 + 8 + 1)}{2} = 81; SE_{W_s} = \sqrt{\frac{9 * 8 * (9 + 8 + 1)}{12}} = 10,4; z = \frac{68 - 81}{10,4} = -1,25$$

Если полученное z-значение больше, чем 1,960 (независимо от знака), то $p < 0,05$, если z-значение больше, чем 2,576, то $p < 0,01$ и если z-значение больше, чем 3,291, то $p < 0,001$. Это говорит о том, что 95 %, 99 % и 99,9 % z-значений лежат в диапазоне от –1,960 до +1,960; от –2,576 до +2,576 и от –3,291 до +3,291 соответственно. Согласно нашим данным ($z = -1,25$), нулевую гипотезу об отсутствии различий между уровнями личностной тревожности у женщин с патологией беременности и без патологии ($p > 0,05$) отклонить нельзя, значит, мы делаем вывод об отсутствии таковых.

Алгоритм расчета критерия Манна – Уитни:

1. Осуществить ранжирование, как в случае двухвыборочного критерия Вилкоксона.

2. Для того чтобы рассчитать значение критерия Манна – Уитни ($U_{эмп}$), нужно знать количество наблюдений в каждой группе (n_1, n_2), большую из двух ранговых сумм (T_x) и количество наблюдений в группе, имеющей T_x (n_x):

$$U_{эмп} = n_1 * n_2 + \frac{n_x * (n_x + 1)}{2} - T_x$$

$$U_{эмп} = 9 * 8 + \frac{8 * (8 + 1)}{2} - 85 = 23$$

3. Далее по табл. 2 для n_1 и n_2 определить критическое значение U-критерия ($U_{кр}$) [4]. Если $U_{эмп} \leq U_{кр}$, то принимается альтернативная гипотеза (свидетельствующая о том, что есть различия) при соответствующем уровне значимости, в данном случае при $p = 0,05$.

Таблица 2

Критические значения критерия Манна – Уитни ($U_{кр}$) при $p = 0,05$

n_2	n_1															
	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8		
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13		
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20		
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27		
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34		
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41		
9	12	15	17	20	23	26	28	31	34	37	39	42	45	48		
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55		
11	16	19	23	26	30	33	37	40	44	47	51	55	58	62		
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69		
13	20	24	28	33	37	41	45	50	54	59	63	67	72	76		
14	22	26	31	36	40	45	50	55	59	64	67	74	78	83		
15	24	29	34	39	44	49	54	59	64	70	75	80	85	90		

Продолжение таблицы 2

	n_1													
n_2	7	8	9	10	11	12	13	14	15	16	17	18	19	20
16	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	34	41	48	55	62	69	76	83	90	98	105	112	119	127

При $n_1 = 9$ и $n_2 = 8$ $U_{кр} = 15$. Следовательно, для нашего примера принимается нулевая гипотеза об отсутствии различий, так как $U_{эмп} > U_{кр}$.

Рассмотрим пример использования двухвыборочно-го критерия Вилкоксона (и соответственно критерия Манна – Уитни) в программе STATA [7, 8]. Возьмем тот же самый пример. У беременных женщин с отсутствием (значение переменной pregnancy – norm) и наличием изучаемой патологии (значение переменной pregnancy – pathology) беременности (переменная pregnancy) изучался показатель личностной тревожности (переменная anxiety) по методике Спилбергера – Ханина (рис. 1).

Проверим, отличается ли личностная тревожность (переменная anxiety) у женщин с нормальным (значение переменной pregnancy – norm) и патологическим (значение переменной pregnancy – pathology) течением беременности. Для того чтобы исключить возможность применения параметрического критерия Стьюдента для несвязанных выборок, проверим распределение изучаемого признака в обеих группах с помощью критерия Shapiro-Wilk (рис. 2).

Если хотя бы в одной из изучаемых групп распределение признака ненормальное, целесообразнее применять непараметрические критерии. Ввиду отклонения распределения переменной anxiety (личностная тревожность) в группе женщин с патологическим течением беременности (pregnancy = pathology) от нормального ($p = 0,026$) для сравнения групп лучше использовать двухвыборочный критерий Вилкоксона или критерий Манна – Уитни.

Для применения непараметрического критерия для

	pregnancy	anxiety
1	norm	23
2	norm	31
3	norm	27
4	norm	28
5	norm	27
6	norm	39
7	norm	21
8	norm	40
9	norm	35
10	pathology	30
11	pathology	49
12	pathology	32
13	pathology	26
14	pathology	52
15	pathology	36
16	pathology	26
17	pathology	50

Рис. 1. Окно STATA Data Editor для примера использования двухвыборочного критерия Вилкоксона/критерия Манна – Уитни

двух несвязанных выборок в STATA следует зайти в диалоговое окно Wilcoxon rank-sum test (что означает двухвыборочный критерий Вилкоксона), которое открывается при помощи меню Statistics – Summaries, tables, and tests – Nonparametric tests of hypothesis – Wilcoxon rank-sum test (рис. 3). В статистической программе STATA представлен только двухвыборочный критерий Вилкоксона, так как и он, и критерий Манна – Уитни дают идентичные результаты.

Во вкладке Main в окошке Variable выбирается изучаемая переменная anxiety (личностная тревожность); в окошке Grouping variable – группирующая переменная pregnancy (беременность) и Ok (рис. 4). Результаты сравнения групп с помощью критерия Вилкоксона представлены на рис. 5.

В таблице рис. 5 представлено количество наблюдений в каждой группе (obs), сумма рангов (rank

```

-> pregnancy = norm

                Shapiro-Wilk W test for normal data

+-----+-----+-----+-----+-----+
Variable | Obs   W       V       z       Prob>z
+-----+-----+-----+-----+-----+
anxiety  |   9  0.94185  0.854  -0.257  0.60150
+-----+-----+-----+-----+-----+

-> pregnancy = pathology

                Shapiro-Wilk W test for normal data

+-----+-----+-----+-----+-----+
Variable | Obs   W       V       z       Prob>z
+-----+-----+-----+-----+-----+
anxiety  |   8  0.79627  2.838  1.941  0.02611
+-----+-----+-----+-----+-----+
    
```

Рис. 2. Результаты проверки распределения переменной anxiety (личностная тревожность) в группах женщин с нормальным (pregnancy = norm) и патологическим (pregnancy = pathology) течением беременности с помощью критерия Shapiro-Wilk

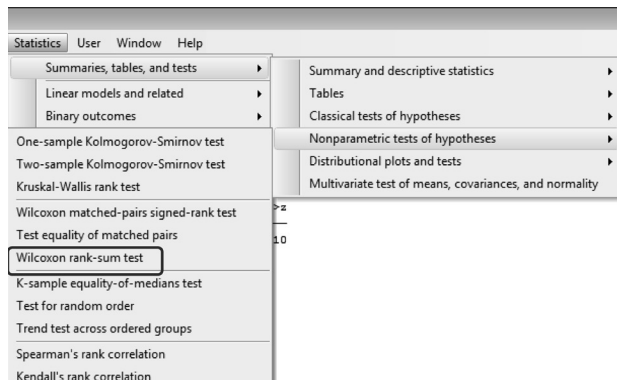


Рис. 3. Алгоритм поиска двухвыборочного критерия Вилкоксона в программе STATA

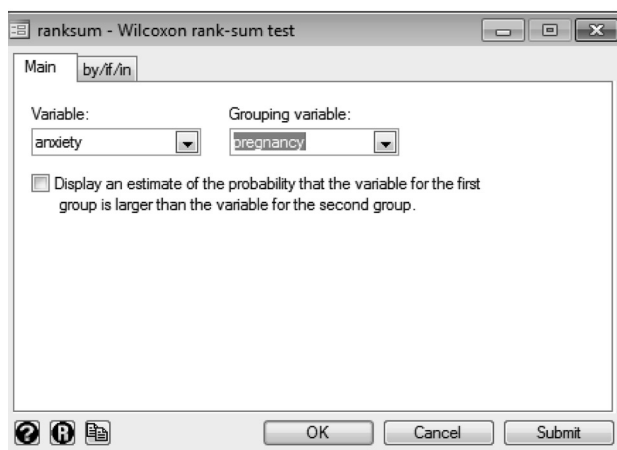


Рис. 4. Диалоговое окно Wilcoxon rank-sum test

sum), ожидаемая сумма рангов (expected), чуть ниже – z-значение и уровень статистической значимости ($Prob > |z|$), который свидетельствует об отсутствии различий между изучаемыми группами, так как достигнутый уровень значимости ($p = 0,2104$) больше критического ($p = 0,050$).

Представляя результаты применения двухвыборочного критерия Вилкоксона, рекомендуется указывать z-значение и абсолютную величину достигнутого уровня значимости (p). Пример такого представления в табл. 3.

Таблица 3

Динамические особенности динамометрии правой руки у девочек младшего школьного возраста в разные периоды наблюдения, Me (Q_1 ; Q_3)

Признак	Женщины		z	p
	с нормальным течением беременности	с патологическим течением беременности		
Личностная тревожность, балл	28,0 (27,0; 35,0)	34,0 (28,0; 49,5)	-1,252	0,210

В целях экономии места в диссертации или журнальной статье информацию табл. 3 можно представить в тексте.

В следующем выпуске Практикума будут рассмотрены параметрические методы анализа данных при наличии трех и более независимых групп.

Список литературы

1. Гланц С. Медико-биологическая статистика / пер. с англ. М., Практика, 1998. 459 с.
2. Гржибовский А. М. Анализ количественных данных для двух независимых групп // Экология человека. 2008. № 2. С. 54–61.
3. Гржибовский А. М. Одномерный анализ повторных измерений // Экология человека. 2008. № 4. С. 51–60.
4. Статистический анализ эмпирических исследований [Электронный ресурс]. URL: <http://statexpert.org> (дата обращения 26.07.2013).
5. Титкова Л. С. Математические методы в психологии. Владивосток : Изд-во Дальневосточного университета, 2002. 85 с.

```
. ranksum anxiety, by(pregnancy)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

+-----+-----+-----+
pregnancy |      obs   rank sum   expected
+-----+-----+-----+
      norm |         9         68         81
      pathology |         8         85         72
+-----+-----+-----+
      combined |        17        153        153
+-----+-----+-----+

unadjusted variance      108.00
adjustment for ties      -0.26
+-----+-----+-----+
adjusted variance      107.74
+-----+-----+-----+

Ho: anxiety(pregna~y==norm) = anxiety(pregna~y==pathology)
      z =  -1.252
      Prob > |z| =  0.2104
```

Рис. 5. Результаты сравнения личностной тревожности в группах женщин с нормальным и патологическим течением беременности (с помощью двухвыборочного критерия Вилкоксона)

6. Field A. Discovering statistics using SPSS. 2nd ed. London : SAGE Publication, 2005. 781 p.
7. Hamilton L. C. Statistics with STATA: Updated for Version 10. USA, 2009. 400 p.
8. Stata: Release 12. Statistical Software. College Station, TX : StataCorp LP, 2011.

References:

1. Glantz S. *Basic biostatistics (translated from English into Russian)*. Moscow, Praktika Publ., 1998, 459 p.
2. Grjibovski A. M. Analysis of quantitative data for two independent groups. *Ekologiya cheloveka* [Human Ecology]. 2008, 2, pp. 54-61. [in Russian]
3. Grjibovski A. M. Univariate analysis of repeated measurements. *Ekologiya cheloveka* [Human Ecology]. 2008, 4, pp. 51-60. [in Russian]
4. *Statisticheskii analiz empiricheskikh issledovaniya* [Statistical analysis of empirical research]. Available at: <http://statexpert.org> (Accessed 26.07.2013).
5. Titkova L. S. *Matematicheskie metody v psikhologii* [Mathematical methods in psychology]. Vladivostok, Far East University Publ., 2002, 85 p.
6. Field A. Discovering statistics using SPSS. 2nd ed. SAGE Publication, London, 2005. 781 p.
7. Hamilton L. C. Statistics with STATA: Updated for Version 10. USA, 2009, 400 p.
8. Stata: Release 12. Statistical Software. College Station, TX, StataCorp LP, 2011.

ANALYSIS OF TWO INDEPENDENT SAMPLES USING STATA SOFTWARE: NON-PARAMETRIC CRITERIA

¹O. A. Kharkova, ^{1,2}A. M. Grjibovski

¹*International School of Public Health, Northern State Medical University, Arkhangelsk, Russia*

²*Department of International Public Health, Norwegian Institute of Public Health, Oslo, Norway*

In this paper we present theoretical principles of the use of non-parametric criteria for comparing two independent groups. Practical examples are presented with detailed explanations of the use of formulas as well as step-by-step algorithm on how to use STATA software for Mann-Whitney and two-sample Wilcoxon tests with subsequent interpretation of the outputs.

Key words: non-parametric criteria, Mann-Whitney test, two-sample Wilcoxon test

Контактная информация:

Гржибовский Андрей Мечиславович – профессор, доктор медицины, старший советник Норвежского института общественного здоровья, г. Осло, Норвегия; директор Архангельской международной школы общественного здоровья ГБОУ ВПО «Северный государственный медицинский университет», г. Архангельск

Адрес: Nasjonalt folkehelseinstitutt, Pb 4404 Nydalen, 0403 Oslo, Norway

Тел.: +47 22048319, e-mail: angr@fhi.no