

Сравнение парных групп с использованием программного обеспечения SPSS и STATA: непараметрические критерии

Гржибовский А.М.¹⁻⁴ PhD, заведующий центральной научной исследовательской лабораторией
Унгурияну Т.Н.¹ д.м.н., доцент кафедры гигиены и медицинской экологии
Горбатова М.А.¹ к.м.н., научный сотрудник центральной научной исследовательской лаборатории

1 — ГБОУ ВПО «Северный государственный медицинский университет» Минздрава России
163000, г. Архангельск, проспект Троицкий, д. 51

2 — Национальный Институт Общественного Здравоохранения
Осло, Норвегия

3 — Северо-Восточный Федеральный Университет
Якутск, Россия

4 — Международный Казахско-Турецкий Университет им. Х.А. Ясави
Туркестан, Казахстан

Автор для корреспонденции: Гржибовский Андрей Мечиславович; e-mail: Andrej.Grjibovski@gmail.com

Финансирование. Исследование не имело спонсорской поддержки.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 12.07.2017.

Рассмотрены способы сравнения количественных данных в парных группах с помощью непараметрических критериев. Описано сравнение двух парных групп с помощью парного критерия Вилкоксона, трех парных групп — с помощью критерия Фридмана. Представлены примеры расчетов вышеупомянутых критериев с использованием программного обеспечения SPSS и STATA. Даются рекомендации о том, как следует представлять результаты сравнения данных с помощью рассмотренных непараметрических критериев в научных публикациях.

Ключевые слова: статистический анализ, парные выборки, парный критерий Вилкоксона, критерий Фридмана, SPSS, STATA.

Введение

Статья продолжает серию публикаций, посвященных статистическому анализу данных биомедицинских исследований [7—9]. В предыдущем выпуске практикума были представлены параметрические методы, применяемые для сравнения количественных данных в трех и более зависимых выборках. Рассматривались парный критерий Стьюдента и однофакторный дисперсионный анализ повторных измерений, основным условием применения которых помимо наличия зависимых выборок, является нормальность распределения данных.

Цель настоящей статьи — познакомить читателей с непараметрическими критериями для сравнения зависимых групп. Рассмотрены одновыборочный критерий Вилкоксона (*Wilcoxon signed rank test*) для сравнения двух зависимых групп и критерий Фридмана (*Friedman test*), предназначенный для сравнения трех и более зависимых групп.

Зависимые или парные выборки могут быть в ситуации, когда изучаемый признак измеряется у одних и тех

же участников исследования в два разных момента времени, например, до и после лечебного вмешательства, — исследования типа «до — после». Также парными данными будут результаты измерения одного и того же показателя у одной и той же группы пациентов, выполненные в двух разных лабораториях или двумя разными исследователями. Парными также считаются данные, полученные для двух групп в результате исследования типа «случай — контроль» с использованием метода подобранных пар [19].

В случае парных измерений существует связь между повторно измеренными значениями параметра для каждого из объектов исследования: значения исследуемого параметра у одного и того же объекта исследования будут ближе друг к другу, чем результаты измерений у нескольких разных объектов, а, следовательно, и дисперсия значений при повторных измерениях будет меньше, чем в случае сравнения независимых групп [1, 4]. Поэтому парные данные требуют несколько иного подхода к их анализу по сравнению с данными, полученными при исследовании независимых выборок.

Непараметрические критерии используются для сравнения средних значений (в данном случае медианных) между группами, если распределение данных во всех группах или хотя бы в одной группе является асимметричным. При расчете непараметрических критериев не используются параметры нормального распределения (среднее арифметическое и стандартное отклонение), а их расчет основан на рангах, что позволяет нивелировать эффект выскакивающих величин. Непараметрические критерии не требуют соблюдения условия нормальности распределения, что позволяет их применять даже при сильно смещенных распределениях, однако распределения данных в сравниваемых группах не должны сильно отличаться друг от друга. Алгоритм расчета тестовой статистики одновыборочного критерия Вилкоксона и критерия Фридмана описан в большинстве учебников по биostatистике [1, 3, 12, 13, 18] и в данной статье не рассматривается.

Следует отметить, что методы непараметрической статистики могут быть использованы и при наличии нормального распределения количественного признака, но в таком случае они будут иметь меньшую мощность по сравнению с параметрическими методами, то есть могут не выявить имеющиеся различия между группами там, где различия действительно существуют. Перед тем как приступить к статистической обработке данных, целесообразно ознакомиться с литературой по эпидемиологии и дизайнам исследований [2, 17].

Парный критерий Вилкоксона или одновыборочный критерий Вилкоксона (Wilcoxon signed rank test) используется для сравнения двух парных выборок с распределением отличным от нормального для проверки нулевой гипотезы о том, что парные наблюдения не отличаются, т.е. разность медиан для парных данных равна нулю. Парный критерий Вилкоксона основан на ранжировании абсолютных величин разности значений в парных выборках. Для того, чтобы можно было достигнуть 5% уровня вероятности, должно быть не меньше шести разностей, неравных нулю. Критерий учитывает, как направление (знак), так и величину разностей между наблюдениями.

Так как условие нормальности распределения данных не является обязательным, то парный критерий Вилкоксона можно применить не обращая внимания на распределение изучаемого признака в группах. Однако надо помнить, что непараметрические критерии являются менее чувствительными, чем параметрические [3], поэтому если данные подчиняются закону нормального распределения, то корректнее использовать парный критерий Стьюдента. Помимо количественных переменных, не подчиняющихся закону нормального распределения, этот критерий можно применять и для порядковых данных, например пятибалльная оценка

в школе, десятибалльная оценка состояния новорожденного по шкале Апгар и т.д. [6].

Для применения одновыборочного критерия Вилкоксона необходимо соблюдение следующих условий:

- количественный или качественный порядковый признак;
- в случае количественного признака нормальное распределение разности между значениями изучаемого признака в группах не обязательно;
- наличие не более двух зависимых групп;
- каждая из изучаемых пар должна быть независима от остальных пар наблюдений.

Если одна и та же группа больных последовательно подвергается трем или большему числу методов лечения или просто наблюдается в три и более момента времени, применяют дисперсионный анализ повторных измерений. Но чтобы использование дисперсионного анализа было правомерно, данные должны подчиняться нормальному распределению. Если условие нормальности распределения данных не соблюдается, то лучше воспользоваться критерием Фридмана — непараметрическим аналогом дисперсионного анализа повторных измерений. Условия для применения критерия Фридмана такие же, как и для одновыборочного критерия Вилкоксона, за исключением того, что могут сравниваться три и более зависимых групп.

При обнаружении статистически значимых различий между группами с помощью критерия Фридмана следующим шагом будет проведение апостериорных сравнений с помощью критерия Вилкоксона для парных выборок с измененным критическим уровнем значимости. Для того, чтобы рассчитать новый критический уровень статистической значимости нужно стандартный уровень значимости 0,05 разделить на число попарных сравнений. Например, если попарно сравниваются три группы: 1 и 2, 2 и 3, 1 и 3, то скорректированный критический уровень статистической значимости составит: $0,05/3 = 0,017$. Таким образом, различия между группами будут считаться статистически значимыми, если достигнутый уровень статистической значимости будет менее 0,017. В случае четырех групп проводится 6 попарных сравнений (1—2, 1—3, 1—4, 2—3, 2—4, 3—4) и, соответственно, скорректированный уровень $p = 0,0085$ [5].

В настоящей статье для приобретения читателями практических навыков проведения статистического сравнения количественных данных в двух парных выборках использован пример гипотетического исследования по изучению изменения количества выкуриваемых сигарет у пациентов, страдающих алкоголизмом, в трезвом состоянии и в состоянии алкогольного опьянения. Сравнение количественных данных в трех зависимых выборках рассмотрено на гипотетическом примере по изучению изменения уровня содержания сахара в крови

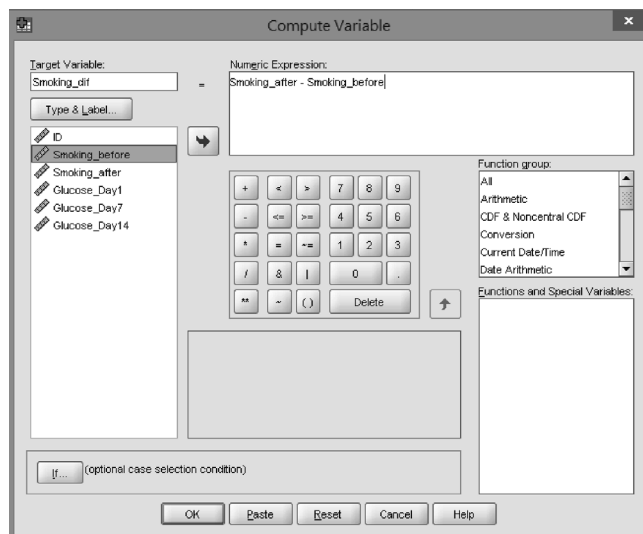


Рис. 1. Диалоговое окно «Compute Variable» для создания новых переменных с помощью математических вычислений программы SPSS 17.0.

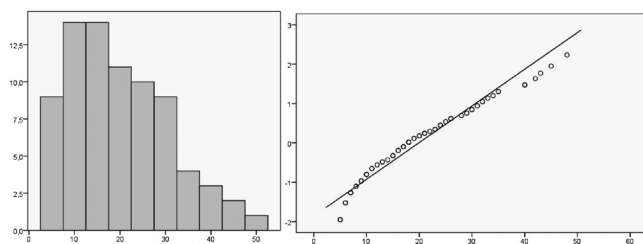


Рис. 2. Гистограмма (слева) и квантильная диаграмма (справа) распределения переменной «Smoking_dif».

у пациентов, злоупотребляющих алкоголем и находящихся на стационарном лечении по поводу сахарного диабета. С использованием программного обеспечения SPSS и STATA представлено сравнения двух зависимых групп с помощью одновыборочного критерия Вилкоксона и сравнение трех зависимых групп — с помощью критерия Фридмана. Основную информацию по использованию SPSS можно получить в пособиях [11, 14, 20], информацию по использованию программы STATA можно прочитать в практикуме [7, 13, 15, 18] и специализированной литературе [21—23].

Сравнение двух зависимых групп с использованием программы SPSS 17.0

Расчет одновыборочного критерия Вилкоксона в SPSS рассмотрен на примере данных гипотетического исследования, в котором изучалось количество выкуриваемых сигарет у 77 пациентов в трезвом состоянии (переменная «Smoking_before») и в состоянии алкогольного опьянения (переменная «Smoking_after»).

В рассматриваемом примере нулевой будет являться гипотеза об отсутствии различий между средним количеством выкуриваемых сигарет в трезвом состоянии и в состоянии алкогольного опьянения. Перед тем как начать проверку гипотезы с помощью критерия Вилкоксона для парных выборок, удостоверимся, что распределение разности между двумя измерениями не подчиняется закону нормального распределения.

Для проверки соблюдения условия нормальности распределения разности между значениями изучаемого признака в группах необходимо создать новую переменную, обозначим ее «Smoking_dif», значения которой будут представлять собой разность между количеством выкуриваемых сигарет в состоянии алкогольного опьянения и в трезвом состоянии. Создается новая переменная путем открытия диалогового окна «Compute» (рис. 1) в меню «Transform».

Проверка распределения проводится с использованием графических методов или статистических критериев. Пошаговый алгоритм проверки распределения количественных данных на основе описательной статистики, графиков распределения и статистических критериев рассматривался в предыдущих выпусках практикума [7—9], поэтому в настоящей статье не представлен.

Проверка распределения показала, что новая переменная не подчиняется закону нормального распределения (табл. 1 и рис. 2), значит, среднее количество выкуриваемых сигарет в трезвом состоянии и в состоянии алкогольного опьянения лучше сравнивать с помощью критерия Вилкоксона для парных выборок.

Для сравнения пар переменных с помощью критерия Вилкоксона необходимо открыть диалоговое окно «Two Related Samples Test» (рис. 3), зайдя в меню «Analyze», затем «Nonparametric Tests», в котором следует выбрать «2 Related Samples». В левой области окна следует выбрать пару (или пары) переменных, ко-

Таблица 1

Результаты проверки распределения переменной «Smoking_dif» с помощью статистических критериев

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Smoking_dif	,104	77	,040	,947	77	,003

a. Lilliefors Significance Correction

торые планируется сравнить. В данном случае это пара, как показано на рис. 3, после чего в меню «Options» нужно выбрать «Descriptives» и «Quartiles» для получения описательной статистики. Возврат в окно «2 Related Samples» осуществляется нажатием кнопки «Continue», а запуск анализа — кнопки «ОК».

Описательная статистика для количества выкуренных сигарет для каждой из групп представлена в табл. 2. Поскольку распределение переменных отличается от нормального, предпочтительнее описывать данные не с помощью средних арифметических и стандартных отклонений, а с помощью медиан (столбец «Median»), первого (столбец «25th») и третьего (столбец «75th») квартилей.

Суммы рангов, а также средние значения для положительных и отрицательных рангов представлены в табл. 3. У всех участников исследования количество

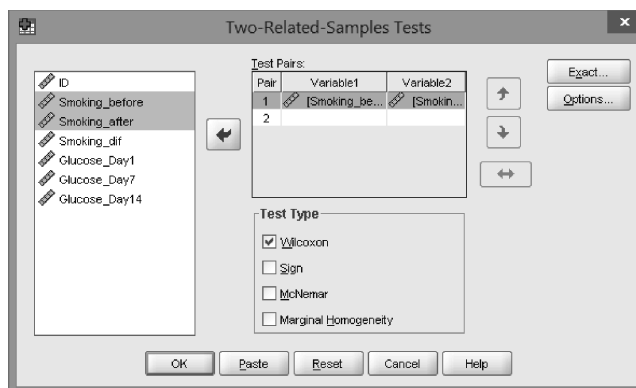


Рис. 3. Диалоговое окно «Two-Related-Samples Tests» программы SPSS 17.0.

Таблица 2

Описательная статистика для сравниваемых групп
Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Smoking_before	77	22,4	11,284	5	55	13	21,00	29
Smoking_after	77	42,32	11,238	21	80	36	42,00	47

Таблица 3

Результаты применения одновыборочного критерия Вилкоксона при сравнении двух связанных групп
Ranks

	N	Mean Rank	Sum of Ranks
Smoking_after - Smoking_before	Negative Ranks	0 ^a	,00
	Positive Ranks	77 ^b	3003,00
	Ties	0 ^c	
	Total	77	

- a. Smoking_after < Smoking_before
- b. Smoking_after > Smoking_before
- c. Smoking_after = Smoking_before

Таблица 4

Уровень значимости при сравнении двух связанных групп с помощью одновыборочного критерия Вилкоксона
Test Statistics^b

	Smoking_after - Smoking_before
Z	-7,625 ^a
Asymp. Sig. (2-tailed)	,000

- a. Based on negative ranks.
- b. Wilcoxon Signed Ranks Test

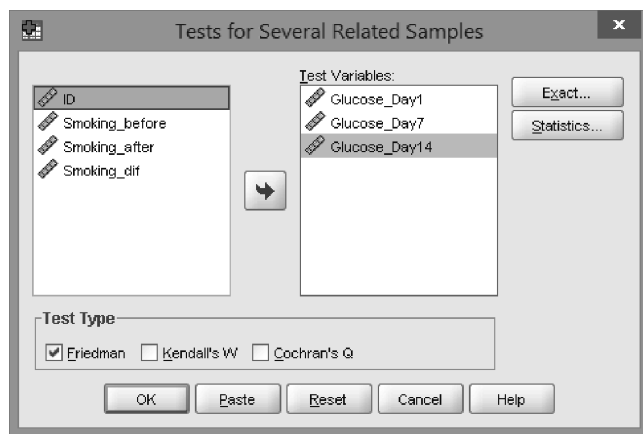


Рис. 4. Диалоговое окно «Tests for Several Related Samples» программы SPSS 17.0.

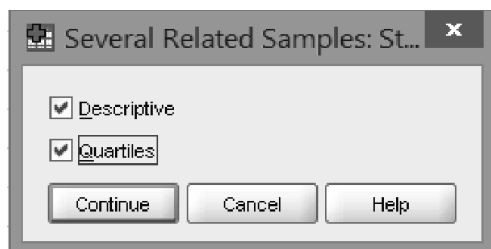


Рис. 5. Диалоговое окно «Several Related Samples: Statistics» программы SPSS 17.0.

выкуриваемых сигарет в состоянии опьянения выше, чем в трезвом состоянии. Оценка статистической значимости полученных результатов показана в табл. 4, где представлена величина Z , которая рассчитывается из наименьшей суммы рангов (T), а также уровень значимости ($Asymp. Sig. (2-tailed)$) для двустороннего теста. Так как достигнутый уровень статистической значимости менее 0,05, то нулевая гипотеза отклоняется, поэтому можно сделать вывод, что количество выкуриваемых сигарет в состоянии опьянения статистически значимо выше, чем в трезвом состоянии.

Помимо критерия Вилкоксона можно сравнивать количественные данные с помощью критерия знаков, который обозначается как «Sign» в диалоговом окне

«Two-Related-Samples Tests» (рис. 3). Однако критерий знаков не учитывает значения изменений переменных, а учитывает лишь их направление и обладает меньшей статистической мощностью, чем критерий Вилкоксона.

Представляя результаты использования критерия Вилкоксона, рекомендуется указывать значение Z и достигнутый уровень значимости (p). Рассмотренный выше пример можно представить следующим образом: В состоянии алкогольного опьянения количество выкуриваемых сигарет за сутки выше ($Me = 42$ шт. в сутки, $Q1 = 36$; $Q3 = 47$), чем в трезвом состоянии ($Me = 21$ шт. в сутки, $Q1 = 13,0$; $Q3 = 29$). Различия были статистически значимы ($Z = -7,625$, $p < 0,001$).

Сравнение трех зависимых групп с использованием программы SPSS 17.0

Применение критерия Фридмана в SPSS рассмотрено на примере данных гипотетического исследования, в котором изучалось содержание сахара в крови у группы пациентов, страдающих алкоголизмом и поступивших в стационар для компенсации сахарного диабета. Группу исследования составили 33 пациента, у которых измерялся уровень сахара в крови при поступлении в стационар (переменная «Glucose_Day1»), на 7-е сутки переменная «Glucose_Day7») и на 14-е сутки переменная «Glucose_Day14»).

Рассмотрим применение критерия Фридмана для данного примера, сравнивая средние значения содержания сахара в три момента измерения. Будет проверяться нулевая гипотеза о том, что нет различий в среднем содержании сахара в крови при трех измерениях. Перед применением критерия Фридмана следует убедиться, что данные не подчиняются закону нормального распределения. Проверка распределения показала, что распределение переменных «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14» отличается от нормального распределения (табл. 5).

Чтобы применить критерий Фридмана необходимо открыть диалоговое окно «Tests for Several Repeated Samples» (рис. 4) путем последовательного выбора

Таблица 5

Результаты проверки распределения переменных «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14» с помощью статистических критериев

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Glucose_Day1	,163	33	,026	,925	33	,025
Glucose_Day7	,178	33	,009	,907	33	,008
Glucose_Day14	,154	33	,046	,913	33	,012

a. Lilliefors Significance Correction

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Glucose_Day1	33	9,9358	1,33303	8,12	13,00	8,9850	9,6500	10,8150
Glucose_Day7	33	8,1985	1,35110	5,54	12,56	7,5400	7,9800	8,7950
Glucose_Day14	33	5,6688	,92709	4,34	8,12	5,0000	5,3400	6,0150

меню «Analyze» → «Nonparametric Tests» → «K Repeated samples». В выбранном окне следует выделить и перенести сравниваемые переменные (в данном случае «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14») из левого поля в правое; отметить «Friedman» в графе «Test Type». Для получения описательной статистики отметить в меню «Statistics» (рис. 5) «Descriptive» и «Quartiles».

Результаты сравнения групп с помощью критерия Фрийдмана представлены в табл. 6, 7 и 8. В табл. 6 представлены данные описательной статистики, из которых предпочтительнее использовать медианы и квартили. В табл. 7 представлены средние ранги для каждой из групп, а в табл. 8 — значение критерия Фрийдмана (Chi-square), количество степеней свободы (df) и уровень значимости ρ (Asymp. Sig). Критерий Фрийдмана имеет распределение типа хи-квадрат (χ^2), поэтому запись «Chi-square» не должна вызывать удивления. Вместе с тем, не стоит путать этот критерий с критерием хи-квадрат Пирсона, который предназначен для сравнения качественных переменных.

Помимо критерия Фрийдмана SPSS предлагает оценить различия между связанными группами с помощью критерия W Кендалла (Kendall's W) и критерия Q Кокрена (Cochrane's Q). Эти критерии для анализа данных в примере не подходят, так как критерий Кендалла предназначен для оценки согласия между группами, а критерий Кокрена используется для сравнения дихотомических переменных.

Представляя результаты, полученные с помощью критерия Фрийдмана, следует указывать величину критерия, количество степеней свободы и достигнутый уровень значимости. Для рассматриваемого примера запись результатов теста может выглядеть так:

$\chi^2 = 49,7$, $df = 2$, $\rho < 0,001$. При обнаружении статистически значимых различий между тремя группами с помощью критерия Фрийдмана следующим шагом будет проведение апостериорных сравнений с помощью рассмотренного выше критерия Вилкоксона с измененным критическим уровнем значимости. При проведении трех попарных сравнений в примере скорректированный критический уровень статистической значимости составляет 0,017. В табл. 9 показаны результаты попарных сравнений. Установлено, что достигнутый уровень статистической значимости для всех сравнений ($\rho < 0,001$) меньше критического ($\rho = 0,017$), следовательно, средний уровень сахара крови между тремя измерениями статистически значимо различается.

Таблица 7

Средний ранг при сравнении трех зависимых групп с помощью критерия Фрийдмана

	Mean Rank
Glucose_Day1	2,85
Glucose_Day7	2,03
Glucose_Day14	1,12

Таблица 8

Результаты сравнения трех связанных групп с помощью критерия Фрийдмана

	Mean Rank
Glucose_Day1	2,85
Glucose_Day7	2,03
Glucose_Day14	1,12

Таблица 9

Результаты попарных сравнений с помощью одновыборочного критерия Вилкоксона

Test Statistics^b

	Glucose_Day7 - Glucose_Day1	Glucose_Day14 - Glucose_Day1	Glucose_Day14 - Glucose_Day7
Z	-4,432 ^a	-5,012 ^a	-4,815 ^a
Asymp. Sig. (2-tailed)	,000	,000	,000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Сравнение двух зависимых групп с использованием программы STATA 14.0

Сравнение количественных данных между двумя зависимыми группами с использованием программы STATA выполнено на данных, содержащихся в файле «5_Smoking_Glucose_STATA», который можно загрузить с сайта журнала «Наркология». Расчет одновыборочного критерия Вилкоксона в STATA рассмотрен на примере данных гипотетического исследования, в котором изучалось количество выкуриваемых сигарет у 77 пациентов в трезвом состоянии (переменная «Smoking_before») и в состоянии алкогольного опьянения (переменная «Smoking_after»).

Следует проверить соблюдение условия нормальности распределения разности между значениями изучаемого признака в группах. Для того чтобы рассчи-

тать разность между парными наблюдениями, создадим новую переменную «Smoking_dif» в STATA. Для создания новой переменной, получаемой путем арифметического вычисления, необходимо открыть Data > Create or change data > Create new variable. В появившемся окне «generate — Create a new variable» в поле «Variable name» вписать название создаваемой переменной, например, «Smoking_dif». Затем нажать на кнопку Create и в появившемся окне «Expression builder» в поле Category выбрать Variables. Выбирая нужные переменные из списка (дважды кликнув на переменную) и арифметические действия, отображаемые в окне справа, вводим формулу расчета новой переменной, например, Smoking_after — Smoking_before (рис. 6). После нажатия на «ОК» в базе данных появится новый столбец с переменной «Smoking_dif». Создать новую переменную можно с помощью команды «generate». Для этого в нижней части главного диалогового окна STATA в поле Commands необходимо ввести: `. generate Smoking_dif = Smoking_after — Smoking_before`.

Проверка распределения проводится с использованием описательной статистики, графических методов и статистических критериев. Алгоритм построения гистограммы, квантильной диаграммы и расчета статистических тестов на нормальное распределение в STATA подробно был изложен в предыдущих номерах журнала [7 — 9] и в практикуме [13], поэтому в настоящей статье не описывается.

Результаты проверки распределения переменной «Smoking_dif» с помощью статистического критерия Shapiro-Wilk представлены в табл. 10. Достигнутый уровень статистической значимости (в данном случае обозначается как Prob>z) составил 0,002, следовательно, распределение переменной «Smoking_dif» отличается от нормального распределения. Поэтому среднее количество выкуриваемых сигарет в трезвом состоянии и в состоянии алкогольного опьянения лучше сравнивать с помощью критерия Вилкоксона для парных выборок.

Для вычисления парного критерия Вилкоксона в STATA необходимо в меню «Statistics» выбрать

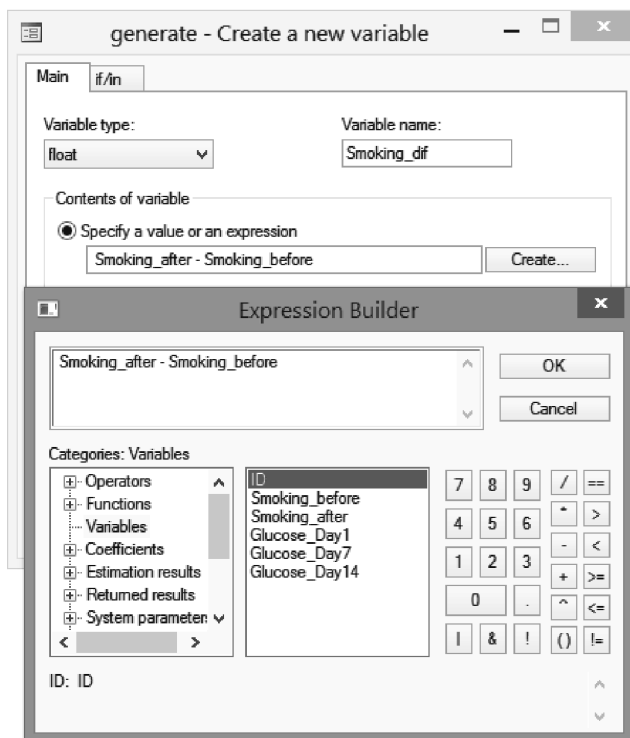


Рис. 6. Диалоговое окно для создания новой переменной путем вычисления в STATA 14.0.

Результаты проверки распределения переменной «Smoking_dif» с помощью статистического критерия Shapiro-Wilk (команда: `. swilk`)

Таблица 10

```
. swilk Smoking_dif
```

Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	z	Prob>z
Smoking_dif	77	0.94596	3.595	2.798	0.00257

«Summaries, tables, and tests», далее выбрать «Nonparametric tests of hypotheses», а затем «Wilcoxon matched-pairs signed-rank test». В диалоговом окне «Wilcoxon matched-pairs signed-rank test» (рис. 7) во вкладке «Main» в окне «Variable» выбирается переменная «Smoking_after». Рядом с окном «Expression» нажать на кнопку «Create» и в появившемся окне «Expression builder» в поле Category выбрать Variables. Затем следует выбрать из появившегося списка переменную «Smoking_before» (дважды кликнув на переменную) и нажать ОК. После этого переменная «Smoking_before» отобразится в окне «Expression» (рис. 7). Сравнение переменных запускается кнопкой ОК. Для расчета парного критерия Вилкоксона также можно воспользоваться командой: `. signrank Smoking_after = Smoking_before.`

Результаты сравнения групп с помощью парного критерия Вилкоксона представлены в табл. 11. В табл. 11 представлено количество (obs) положительных (positive), отрицательных (negative) и нулевых (zero) изменений, а также сумма рангов для каждого из них (sum ranks). В нижней части таблицы 11 показано значение $z = 7,625$ и уровень статистической значимости ($\text{Prob} > |z| = 0,000$). Значение достигнутого уровня статистической значимости для двустороннего критерия составило $p < 0,001$, что позволяет отклонить нулевую гипотезу об отсутствии различий между группами, следовательно, среднее количество выкуриваемых сигарет в трезвом состоянии и в состоянии алкогольного опьянения на популяционном уровне различается.

Сравнение трех зависимых групп с использованием программы STATA 14.0

Сравнение количественных данных между тремя зависимыми группами с использованием программы STATA выполнено на данных, содержащихся в файле «5_Smoking_Glucose_STATA», который можно загрузить с сайта журнала «Наркология». Рассмотрим применение критерия Фридмана в STATA на примере данных гипотетического исследования, в котором изучалось содержание сахара в крови у группы пациентов, страдающих алкоголизмом и поступивших в стационар для компенсации сахарного диабета. Группу исследования составили 33 пациентов, у которых измерялся уровень сахара в крови при поступлении в стационар (переменная «Glucose_Day1»), на 7-е сутки переменная «Glucose_Day7») и на 14-е сутки переменная «Glucose_Day14»). В рассматриваемом примере нулевой будет являться гипотеза об отсутствии различий между средним содержанием сахара при трех измерениях.

Перед применением критерия Фридмана необходимо убедиться, что распределение данных в каждой группе отличается от нормального распределения. Проверка рас-

пределения с использованием критерия Shapiro-Wilk показала, что распределение переменных «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14» не подчиняется закону нормального распределения (табл. 12), поэтому, для сравнения трех парных групп используем критерий Фридмана.

Для расчета критерия Фридмана в STATA необходимо переставить данные так, чтобы субъекты (участники) были по столбцам, а переменные «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14» — по строкам.

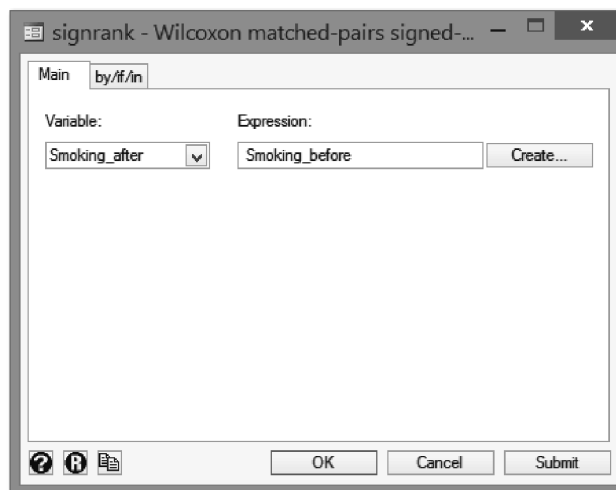


Рис. 7. Диалоговое окно «Wilcoxon matched-pairs signed-rank test» программы STATA 14.0.

Таблица 11

Результаты применения одновыборочного критерия Вилкоксона

```
. signrank Smoking_after = Smoking_before

Wilcoxon signed-rank test
```

sign	obs	sum ranks	expected
positive	77	3003	1501.5
negative	0	0	1501.5
zero	0	0	0
all	77	3003	3003

```
unadjusted variance      38788.75
adjustment for ties      -11.88
adjustment for zeros      0.00
-----
adjusted variance        38776.88

Ho: Smoking_after = Smoking_before
      z =      7.625
      Prob > |z| =    0.0000
```


Результаты проверки распределения переменных «Glucose_Day1», «Glucose_Day7» и «Glucose_Day14» с помощью критерия Shapiro-Wilk

. swilk Glucose_Day1 Glucose_Day7 Glucose_Day14						
Shapiro-Wilk W test for normal data						
Variable	Obs	W	V	z	Prob>z	
Glucose_Day1	33	0.92519	2.554	1.950	0.02558	
Glucose_Day7	33	0.90817	3.135	2.376	0.00874	
Glucose_D~14	33	0.91429	2.926	2.233	0.01277	

Таблица 13

Результаты сравнения трех зависимых групп с помощью критерия Фридмана

. . friedman v1-v33	
Friedman =	49.2727
Kendall =	0.7466
P-value =	0.0000

Для перестановки данных следует воспользоваться командой: `. xpose, clear`. После перестановки данных для расчета критерия Фридмана можно применить команду: `. friedman v1-v33`. Результаты расчета критерия представлены в табл. 13.

При представлении результатов, полученных с помощью критерия Фридмана, рекомендуется указывать величину критерия и достигнутый уровень значимости. В данном случае: $\chi^2 = 49,27$; $p < 0,001$, то есть нулевая гипотеза отклоняется, следовательно средний уровень сахара крови для трех измерений статистически значимо различается. Как выше было сказано, при обнаружении статистически значимых различий между тремя группами с помощью критерия Фридмана следующим шагом должно быть проведение апостериорных сравнений с помощью одновыборочного критерия Вилкоксона с измененным критическим уровнем значимости. Подробнее о представлении результатов одномерных методов статистического анализа можно ознакомиться в [10, 16].

Заключение

В настоящей статье рассмотрены способы сравнения количественных данных в парных группах с помощью непараметрических критериев. Описано срав-

нение двух парных групп с помощью одновыборочного критерия Вилкоксона, трех парных групп — с помощью критерия Фридмана. Данные методы являются непараметрическими, поэтому применяются в случае, если не соблюдается условие нормальности распределения данных. В следующем выпуске практикума будет рассмотрено сравнение корреляционного и простого регрессионного анализа.

Список литературы

1. Банержи А. Медицинская статистика понятным языком: вводный курс. М.: Практическая медицина, 2007. 287 с.
2. Власов В. В. Эпидемиология: учебное пособие для вузов. М.: ГЭОТАР-МЕД, 2004. 464 с.
3. Гланц С. Медико-биологическая статистика. М.: Практика, 1998. 459 с.
4. Гржибовский А.М. Одномерный анализ повторных измерений. *Экология человека* 2008; (4): 51-60.
5. Гржибовский А.М. Анализ трех и более независимых групп количественных данных. *Экология человека* 2008; (3): 50-58.
6. Гржибовский А.М. Выбор статистического критерия для проверки гипотез. *Экология человека* 2008; (11): 48-57.
7. Гржибовский А.М., Унгурияну Т.Н., Горбатова М.А. Описательная статистика с использованием пакетов статистических программ SPSS и STATA. *Наркология* 2017; (4): 36-51.
8. Гржибовский А.М., Унгурияну Т.Н., Горбатова М.А. Анализ непарных выборок в биомедицинских исследованиях с использованием программного обеспечения SPSS и STATA: параметрические критерии. *Наркология* 2017; (5): 18-36.
9. Гржибовский А.М., Унгурияну Т.Н., Горбатова М.А. Сравнение непарных групп с использованием программного обеспечения SPSS и STATA: непараметрические критерии. *Наркология* 2017; (6): 19-30.
10. Ланг Т.А., Сесик М. Как описывать статистику в медицине. Аннотированное руководство для авторов, редакторов и рецензентов. Перевод с англ. под ред. В.П. Леонова. М.: Практическая медицина, 2011. 480 с.
11. Наследов А. Д. SPSS: Компьютерный анализ данных в психологии и социальных науках. СПб.: Питер, 2007. 416 с.
12. Петри А., Сэбин К. Наглядная медицинская статистика. М.: ГЭОТАР-Медиа, 2009. 168 с.
13. Субботина А.В., Гржибовский А.М. Описательная статистика и проверка нормальности распределения количественных данных. *Экология человека* 2014; (2): 60-63.

14. Таганов Д. SPSS: Статистический анализ в маркетинговых исследованиях. СПб.: Питер, 2005. 192 с.
15. Унгуряну Т.Н., Гржибовский А.М. Программное обеспечение для статистической обработки данных STATA: введение. *Экология человека* 2014; (1): 60-63.
16. Унгуряну Т.Н., Гржибовский А.М. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях. *Экология человека* 2011; (5): 55-60.
17. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология: Основы доказательной медицины. М.: МедиаСфера, 1998. 345 с.
18. Харьковская О.А., Гржибовский А.М. Сравнение двух парных выборок с помощью пакета статистических программ STATA: непараметрические критерии. *Экология человека* 2014; (12): 55-60.
19. Холматова К.К., Гржибовский А.М. Применение исследований «случай-контроль» в медицине и общественном здравоохранении. *Экология человека* 2016; (8): 53-60.
20. Field A. *Discovering statistics using SPSS*. London, 2005. 781 p.
21. Hamilton L.C. *Statistics with STATA: Updated for Version 10*. Brooks/Cole, Cengage Learning, 2009. 491 p.
22. Kohler U. Kreute F. *Data Analysis Using Stata*. USA, Texas: Stata Press, 2005. 378 p.
23. Rabe-Hesketh, S., Everitt, Brian. *A Handbook of Statistical Analyses Using Stata*. New York: Chapman & Hall, 2007. 352 p.

References

1. Banerjee A. *Meditsinskaya statistika ponyatnym yazykom: vvodnyi kurs* [Medical statistics made clear]. Moscow: Prakticheskaya meditsina, 2007. 287 p. (In Russ.)
2. Vlasov V.V. *Epidemiologiya: uchebnoe posobie dlya vuzov* [Epidemiology: tutorial for high schools]. Moscow: Geotar-Med, 2004. 464 p. (In Russ.)
3. Glantz S. *Mediko-biologicheskaya statistika* [Biostatistica]. Moscow: Praktika, 1998. 459 p. (In Russ.)
4. Grijbovski A.M. *Odnomernyj analiz povtornykh izmerenij*. [Univariate analysis of repeated measurements]. *Ekologiya cheloveka* [Human Ecology] 2008; (4): 51-60. (In Russ.)
5. Grijbovski A.M. *Analiz trekh i bolee nezavisimykh grupp kolichestvennykh dannykh*. [Analysis of three and more independent groups of quantitative data]. *Ekologiya cheloveka* [Human Ecology] 2008; (3): 50-58. (In Russ.)
6. Grijbovski A.M. *Vybor statisticheskogo kriterija dlya proverki gipotez*. [Choosing a statistical test for hypothesis testing]. *Ekologiya cheloveka* [Human Ecology] 2008; (11): 48-57. (In Russ.)
7. Grijbovski A.M., Unguryanu T.N., Gorbatova M.A. *Opisatel'naya statistika s ispol'zovaniem paketov statisticheskikh programm SPSS i STATA*. [Descriptive statistics with using SPSS and STATA]. *Narkologiya* [Narkology] 2017; (4): 36-51. (In Russ.)
8. Grijbovski A.M., Unguryanu T.N., Gorbatova M.A. *Analiz neparnykh vyborok v biomeditsinskih issledovaniyakh s ispol'zovaniem programm SPSS i STATA*. [Analysis of non-paired samples in biomedical studies using SPSS and STATA software: parametric tests]. *Narkologiya* [Narkology] 2017; (5): 18-36. (In Russ.)

9. Grijbovski A.M., Unguryanu T.N., Gorbatova M.A. *Sravnenie neparnykh grupp s ispol'zovaniem programm SPSS i STATA: neparametricheskie kriterii*. [Comparing of independent groups using SPSS and STATA software: non-parametric statistical tests]. *Narkologiya* [Narkology] 2017; (6): 19-30. (In Russ.)
10. Lang T.A., Secic M. *Kak opisivat' statistiku v medicine. Annotirovannoe rukovodstvo dlja avtorov, redaktorov i recenzentov*. [How to report statistics in Medicine. Annotated guidelines for authors, editors, and reviewers]. Moscow: Practical Medicine, 2011. 480 p. (In Russ.)
11. Nasledov A. D. *SPSS: Komp'yuternyy analiz dannykh v psikhologii i sotsial'nykh naukakh*. [SPSS Computer data analysis in psychology and social sciences]. Saint Petersburg: Piter, 2007. 416 p. (In Russ.)
12. Petri A., Sebin K. *Naglyadnaya meditsinskaya statistika*. [Medical statistics at glance]. Moscow: GEOTAR-Media, 2009. 168 p. (In Russ.)
13. Subbotina A. V., Grijbovski A. M. *Opisatel'naya statistika i proverka normal'nosti raspredeleniya kolichestvennykh dannykh*. [Descriptive statistics and normality testing for quantitative data]. *Ekologiya cheloveka* [Human Ecology] 2014; (2): 60-63. (In Russ.)
14. Taganov D. *SPSS: Statisticheskij analiz v marketingovykh issledovaniyakh*. [SPSS: Statistical analysis in marketing research]. Saint Petersburg: Piter, 2005. 192 p. (In Russ.)
15. Unguryanu T.N., Grijbovski A.M. *Programmnoe obecpechenie dlya statisticheskoy obrabotki dannykh STATA: vvedenie*. [Introduction to STATA-Software for statistical data analysis]. *Ekologiya cheloveka* [Human Ecology] 2014; (1): 60-63. (In Russ.)
16. Unguryanu T.N., Grijbovski A.M. *Kratkie rekomendacii po opisaniyu, statisticheskomu analizu i predstavlenu dannykh v nauchnykh publikacijah*. [Brief recommendations on description, analysis and presentation of data in scientific papers]. *Ekologiya cheloveka* [Human Ecology] 2011; (5): 55-66. (In Russ.)
17. Fletcher R., Fletcher S., Vagner E. *Klinicheskaya epidemiologiya: Osnovy dokazatel'noy meditsiny*. [Clinical Epidemiology: the essentials]. Moscow: MediaSfera, 1998. 345 p. (In Russ.)
18. Kharkova O.A., Grijbovski A.M. *Sravnenie dvuh parnykh vyborok s pomoshh'ju paketa statisticheskikh programm STATA: neparametricheskie kriterii*. [Non-parametric tests for two paired samples using STATA]. *Ekologiya cheloveka* [Human Ecology] 2014; (12): 55-60. (In Russ.)
19. Kholmatoва K.K., Grijbovski A.M. *Primenenie issledovaniy «sluchay-kontrol'» v meditsine i obshchestvennom zdravookhraneni*. [Case-control studies in medicine and public health]. *Ekologiya cheloveka* [Human Ecology] 2016; (8): 53-60. (In Russ.)
20. Field A. *Discovering statistics using SPSS*. London, 2005. 781 p.
21. Hamilton L.C. *Statistics with STATA: Updated for Version 10*. Brooks/Cole, Cengage Learning, 2009. 491 p.
22. Kohler U. Kreute F. *Data Analysis Using Stata*. USA, Texas: Stata Press, 2005. 378 p.
23. Rabe-Hesketh, S., Everitt, Brian. *A Handbook of Statistical Analyses Using Stata*. New York: Chapman & Hall, 2007. 352 p.

COMPARING OF PAIRED GROUPS USING SPSS AND STATA SOFTWARE: NON-PARAMETRIC STATISTICAL TESTS

Grjibovski A.M.¹⁻⁴, Unguryanu T.N.^{1,5}, Gorbatova M.A.¹

- 1 – Northern State Medical University
Arkhangelsk, Russia
- 2 – Norwegian Institute of Public Health
Oslo, Norway
- 3 – North-Eastern Federal University
Yakutsk, Russia
- 4 – International Kazakh-Turkish University
Turkestan, Kazakhstan
- 5 – Arctic University of Norway
Tromsø, Norway

For correspondence: *Grjibovski Andrej*; e-mail: Andrej.Grijbovski@gmail.com

Conflict of interest. The authors declare no conflict of interest.

Funding. The study had no sponsorship.

Received: 12.07.2017.

In this article, we have presented basic principles of the non-parametric criteria using for comparing two and three paired groups. Practical examples are presented with step-by-step algorithm on how to use SPSS and STATA software for Wilcoxon signed rank test and Friedman test with subsequent interpretation of the outputs.

Key words: statistical analysis, paired samples Wilcoxon signed rank test, Friedman test, SPSS, STATA.