

УДК 519.25

СРАВНЕНИЕ КОЛИЧЕСТВЕННЫХ ДАННЫХ В ДВУХ ПАРНЫХ ВЫБОРКАХ С ИСПОЛЬЗОВАНИЕМ ПАКЕТА СТАТИСТИЧЕСКИХ ПРОГРАММ STATA

© 2014 г. А. В. Субботина, *А. М. Гржибовский

Университет г. Тромсё, Норвегия

*Норвежский институт общественного здравоохранения,
г. Осло, Норвегия

В предыдущих выпусках Практикума мы представляли методы, применяемые для сравнения количественных данных двух и более независимых выборок, подчеркивая, что независимость наблюдений друг от друга является одним из условий применения этих методов. В настоящей статье рассмотрим методы сравнения двух парных групп данных с помощью пакета статистических программ STATA [4]. Пользователям SPSS рекомендуем обратиться к нашей более ранней публикации [1].

Под парными данными мы понимаем значения связанных, то есть не являющихся независимыми, выборок. Парность изучаемых выборок определяется дизайном исследования. Примерами парных выборок могут служить 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 являться, к примеру, результаты измерения одного и того же показателя у одной и той же группы пациентов: определение уровня холестерина в крови в нескольких лабораториях или измерение артериального давления двумя исследователями. Парными также считаются данные, полученные для двух групп в результате исследования типа «случай — контроль» с использованием метода подобранных пар (matched case-control design). В данном случае «контроли» не могут считаться полностью независимыми от «случаев», так как их намеренно подбирали, исходя из характеристик «случаев», например пола, возраста, района проживания.

Парный t-критерий Стьюдента используется для того, чтобы определить, существует ли значимая разница между средними значениями одного и того же признака в парных группах. Оба измерения проводятся для каждой единицы выборки, а критерий основан на парных разностях между этими двумя значениями. Критерий также принимает во внимание вариацию в значениях внутри каждой выборки, в результате чего рассчитывается одно значение, известное как значение t Стьюдента.

Зная значение t, мы можем оценить вероятность того, что различие в наших двух выборках не будет отличаться от разницы в выборках, взятых из одной и той же генеральной совокупности, на основании значения достигнутого уровня значимости (величина p). Если разница значимая (в большинстве случаев определяемая как $p < 0,05$), то делается вывод о том, что с выбранным уровнем доверительной вероятности (обычно 95 %) наши две выборки происходят из разных генеральных совокупностей, а значит, между ними существуют статистически значимые различия.

В качестве примера рассмотрим изменение артериального давления у группы, состоящей из 33 пациентов, до и после начала применения гипотензивного препарата X. Значения данных переменных представлены переменными ad_t0 ad_t1 соответственно.

В статье рассматриваются основные принципы сравнения количественных данных в двух парных группах. Приводится пример проверки нормальности распределения данных, а также расчета парного критерия Стьюдента вручную и с помощью пакета статистических программ STATA. Авторы также рассматривают условия применения парного критерия Стьюдента и основные правила представления результатов в публикациях. Материал дает общие представления о применении парного критерия Стьюдента и не заменяет изучения специализированной литературы.
Ключевые слова: парные выборки, нормальное распределение, t-критерий, STATA

Если мы примем первое значение какого-то признака у данного индивида за y_1 , а второе за y_2 , то разность между двумя наблюдениями для каждого пациента составит $d = y_1 - y_2$. Для того чтобы рассчитать разность между парными наблюдениями, создадим новую переменную `diff0_1` в STATA, воспользовавшись командой:

```
. generate diff0_1 = ad_t0 - ad_t1
```

Общие принципы работы пакета статистических программ STATA рассматривались в одном из предыдущих выпусков Практикума [3].

Затем рассчитываем среднее значение разности между двумя наблюдениями и стандартную ошибку для разности.

Среднее значение разности между первым и вторым наблюдениями рассчитывается как

$$\bar{y} = \frac{\sum d}{n} = \bar{y}_1 - \bar{y}_2.$$

Среднее значение разности представляет собой оценочное значение средней разности для генеральной совокупности

$$\mu_d = \mu_1 - \mu_2.$$

Для расчета стандартного отклонения для разности из каждого значения разностей между парами наблюдений отнимаем среднее \bar{y} , сумму квадратов полученных значений делим на количество пар наблюдений n и извлекаем квадратный корень из полученного числа.

$$SD = \sqrt{\frac{\sum [(d)_i - \bar{y}]^2}{n}},$$

где SD — стандартное отклонение разности в наблюдениях, n — размер выборки разностей (в данном случае, число пациентов).

То же самое несложно рассчитать в STATA, используя команду

```
. summarize ad_t0 ad_t1 diff0_1.
```

В результате мы получаем описательную статистику для значений переменных `ad_t1`, `ad_t2` и разности наблюдений (табл. 1):

Таблица 1
Описательная статистика для значений переменных `ad_t1`, `ad_t2` и разности наблюдений

Variable	Obs	Mean	Std. Dev.	Min	Max
<code>ad_t0</code>	33	147.8215	8.79961	125.46	164.68
<code>ad_t1</code>	33	145.9612	4.225814	137.22	154.93
<code>diff0_1</code>	33	1.860303	10.46136	-22.69	25.98

Среднее значение разностей будет равняться 1,86, стандартное отклонение 10,46. В силу того, что мы рассматриваем разности, для определения стандартной ошибки средней арифметической мы можем использовать ту же логику, какую использовали для одиночной выборки:

$$SE_{\bar{x}} = \frac{SD}{\sqrt{n}}.$$

Нулевой гипотезой для парного критерия Стьюдента будет являться утверждение об отсутствии различий между средними в парных выборках, то есть разность (`diff0_1`) равна нулю. Соответственно альтернативной гипотезой будет являться предположение о том, что средняя разность между парными наблюдениями значимо при выбранном уровне значимости отличается от нуля:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

В нашем примере, нулевая и альтернативная гипотеза будут звучать следующим образом:

H_0 — между средними значениями артериального давления при измерении до и после применения препарата различий нет.

H_1 — между средними значениями артериального давления при измерении до и после применения препарата различия есть.

Иными словами, данный тест предназначен для того, чтобы дать ответ на вопрос «Является ли наблюдаемая разность между парными наблюдениями достаточной большой, чтобы отклонить нулевую гипотезу об отсутствии различий?».

Основным условием применения парного t -критерия Стьюдента является нормальное распределение разности между индивидуальными значениями парных выборок. Необходимо обратить особое внимание на тот факт, что данный критерий может применяться при тестировании наблюдений, которые сами происходят из выборок, не соответствующих нормальному распределению, если значения разниц в наблюдениях нормально распределены.

Для проверки этого условия мы можем использовать графический метод, то есть визуально оценить распределение разниц и сравнить с нормальным распределением. Для этого можно построить гистограмму (рис. 1) и / или квантильную диаграмму (рис. 2).

```
. histogram diff0_1, bin(11) normal
```

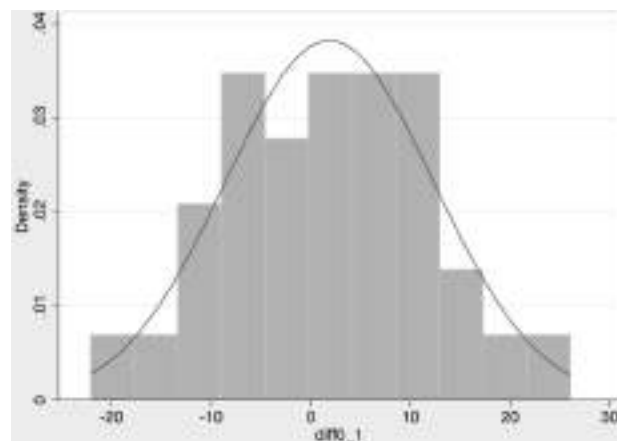


Рис. 1. Гистограмма для оценки распределения переменной `diff0_1`

Судя по гистограмме, распределение разностей в нашем примере может подчиняться закону нормального распределения, однако график не очень убедителен.

Квантильная диаграмма, в свою очередь, достаточно убедительно демонстрирует, что переменная diff0_1 подчиняется закону нормального распределения, по крайней мере в середине распределения.

```
. qnorm diff0_1
```

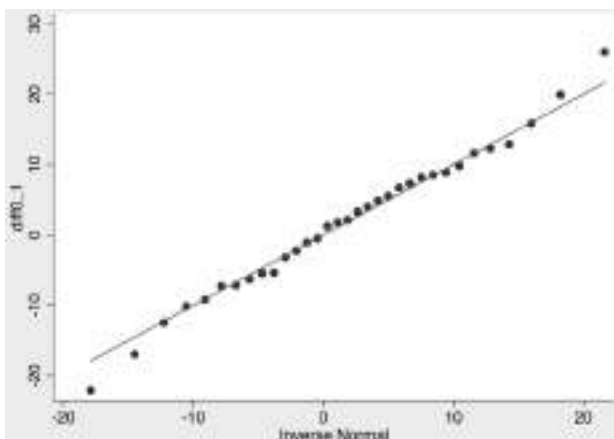


Рис. 2. Квантильная диаграмма для оценки распределения переменной diff0_1

Также мы можем провести формальный тест на соответствие распределения нормальному (табл. 2). Воспользуемся командой

```
. swilk diff0_1
```

Таблица 2

Результат проверки нормальности распределения с помощью критерия Шапиро – Уилка

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
diff0_1	33	0.99522	0.163	-3.771	0.99992

Нулевая гипотеза для критерия Шапиро – Уилка, с помощью которого мы будем проверять данные на соответствие нормальному распределению, состоит в том, что наблюдения принадлежат совокупности с нормальным распределением. Значение $p = 0,99$, поэтому мы не можем отклонить нулевую гипотезу о том, что наши наблюдения распределены нормально, значит, условие нормальности разностей выполняется. Более детально о проверке нормальности распределения количественных данных можно прочитать в [2].

Альтернативным тестом для сравнения парных данных, для проведения которого нормальное распределение разностей не требуется, является критерий Вилкоксона, который мы рассмотрим в одном из последующих выпусков, однако этот тест обладает меньшей мощностью, чем парный t-критерий Стьюдента.

Статистика t для парного критерия Стьюдента рассчитывается как

$$t = \frac{\bar{y} - \Delta}{\frac{SD}{\sqrt{n}}}$$

где \bar{y} – средняя разница между парами наблюдений, Δ – предполагаемая разница (0 при тестировании на равенство средних), SD – стандартное отклонение разниц, n – размер выборки. Количество степеней свободы в данном тесте $n - 1$.

Подставляя значения, в нашем примере мы получаем:

$$t = \frac{1,86 - 0}{\frac{10,46}{\sqrt{33}}} = 1,02$$

В STATA парный t-критерий Стьюдента можно рассчитать двумя способами (табл. 3 и 4).

```
. ttest ad_t0 = ad_t1
```

Таблица 3

Результат применения парного критерия Стьюдента с помощью команды «ttest ad_t0 = ad_t1»

Paired t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
ad_t0	33	147.8215	1.531816	8.79961	144.7613 150.9417
ad_t1	33	145.9612	.7354885	4.225814	144.4631 147.4593
diff	33	1.860303	1.821889	18.46126	-1.949135 5.569741

mean(diff) = mean(ad_t0 - ad_t1) t = 1.0215
 Ho: mean(diff) = 0 degrees of freedom = 32
 Ha: mean(diff) > 0 Pr(T > |t|) = 0.3147 Pr(T < t) = 0.1573

Значение критерия Стьюдента $t = 1,02$ совпадает с ранее рассчитанным по формуле значением. Для оценки уровня значимости двустороннего теста мы смотрим на значение

$$Pr(|T| > |t|) = 0.3147$$

Уровень значимости для двустороннего теста $p = 0,31$. Таким образом, мы не можем отвергнуть нулевую гипотезу об отсутствии различий между средними значениями артериального давления до лечения и через неделю после начала приема препарата. Тот же результат мы можем получить с помощью критерия Стьюдента для одной выборки для ранее рассчитанной переменной diff0_1:

```
. ttest diff0_1=0
```

Таблица 4

Результат применения парного критерия Стьюдента с помощью команды «ttest diff0_1=0»

One-sample t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
diff0_1	33	1.860303	1.821889	18.46126	-1.949135 5.569741

mean = mean(diff0_1) t = 1.0215
 Ho: mean = 0 degrees of freedom = 32
 Ha: mean < 0 Pr(T > |t|) = 0.3147 Pr(T < t) = 0.1573

Представляя результаты применения парного критерия Стьюдента, рекомендуется указывать значение критерия, количество степеней свободы и достигнутый уровень значимости. Рассмотренный выше пример можно представить следующим образом: средние значения артериального давления до лечения 147,8 мм рт. ст. (SD = 8,8) и через неделю после начала приема препарата X 146,0 мм рт. ст. (SD = 4,2) статистически не различались ($t_{32} = 1,02$, $p = 0,31$). Подробнее об особенностях представления результатов статистической обработки данных можно прочитать в [3].

Список литературы

1. Гржибовский А. М. Одномерный анализ повторных измерений // Экология человека. 2008. № 4. С. 51–60.
2. Субботина А. В., Гржибовский А. М. Описательная статистика и проверка нормальности распределения количественных данных // Экология человека. 2014. № 2. С. 51–57.
3. Унгурияну Т. Н., Гржибовский А. М. Краткие рекомендации по описанию, статистическому анализу и представлению данных в научных публикациях // Экология человека. 2011. № 5. С. 55–60.
4. Унгурияну Т. Н., Гржибовский А. М. Программное обеспечение для статистической обработки данных STATA: введение // Экология человека. 2014. № 1. С. 60–63.

References

1. Grjibovski A. M. Univariate analysis of repeated measurements. *Ekologiya cheloveka* [Human Ecology] 2008, 4, pp. 51-60. [in Russian]
2. Unguryanu T. N., Grjibovski A. M. Brief recommendations on description, analysis and presentation of data in scientific papers. *Ekologiya cheloveka* [Human Ecology] 2011, 5, pp. 51-57. [in Russian]
3. Unguryanu T. N., Grjibovski A. M. Introduction to STATA

- statistical software for data analysis. *Ekologiya cheloveka* [Human Ecology] 2011, 5, pp. 55-60. [in Russian]

4. Subbotina A. V., Grjibovski A. M. Descriptive statistics and normality testing for quantitative data. *Ekologiya cheloveka* [Human Ecology] 2014, 2, pp. 60-63. [in Russian]

COMPARISON OF QUANTITATIVE DATA IN TWO PAIRED SAMPLES USING STATA SOFTWARE

A. V. Subbotina, *A. M. Grjibovski

University of Tromsø, Tromsø, Norway

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

In this paper we have presented basic principles of analysis of quantitative data in two paired samples. Examples of normality testing, calculating paired t-test both manually and using the STATA software have been given. We have also considered assumptions for using paired t-test as well as the main principles of presentation of the results in scientific publications. The article has given only basic information on the use of t-test in research and it does not substitute reading specialized literature.

Keywords: paired samples, normal distribution, t-test, STATA

Контактная информация:

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

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

Тел.: +47 22048319, +47 45268913; e-mail: andrej.grjibovski@gmail.com