

УДК 519.25

СРАВНЕНИЕ ОДНОЙ И ДВУХ НЕСВЯЗАННЫХ ВЫБОРОК С ПОМОЩЬЮ ПАКЕТА СТАТИСТИЧЕСКИХ ПРОГРАММ STATA: ПАРАМЕТРИЧЕСКИЕ КРИТЕРИИ

© 2014 г. **О. А. Харькова**, ***А. М. Гржибовский**

Северный государственный медицинский университет, г. Архангельск

*Норвежский институт общественного здравоохранения, г. Осло, Норвегия

Предыдущие статьи практикума информировали о том, какими бывают эмпирические данные, как нужно представлять количественные и качественные данные, а также в общем о пакете прикладных статистических программ STATA. Начиная с этой статьи, мы будем говорить об аналитической статистике, то есть будем проверять статистические гипотезы. Если исследовательскую гипотезу мы формулируем на этапе планирования исследования, то статистическую — при обработке собранных данных.

Статистическая гипотеза бывает альтернативная (H_1) и нулевая (H_0) [5]. Нулевая статистическая гипотеза предполагает, что две генеральные средние (μ_1 и μ_2), то есть средние арифметические для двух генеральных совокупностей, не отличаются друг от друга. Альтернативная гипотеза соответственно заключается в том, что генеральные средние различаются. Для того чтобы принять решение об отклонении нулевой статистической гипотезы, ориентируются на достигнутый уровень статистической значимости (p). Общепринятым, но не единственным критическим уровнем значимости, с которым сравнивают достигнутый уровень значимости, является 0,05. Это говорит о том, что исследователь допускает 5 % вероятность ошибки 1 типа, то есть нахождения различий там, где их на самом деле нет. Следовательно, если достигнутый уровень значимости меньше 0,05 ($p < 0,05$), то H_0 отклоняется, и наоборот, если достигнутый уровень значимости больше 0,05 ($p > 0,05$), то H_0 принимается.

Статистические критерии можно разделить на односторонние/двусторонние и параметрические/непараметрические.

Односторонние статистические критерии — это критерии, в которых альтернативная гипотеза определяет направление интересующего исследователя влияния вмешательства ($\mu_1 < \mu_2$ или $\mu_1 > \mu_2$), а двусторонние — когда направление не определено, а важно наличие различий между средними ($\mu_1 \neq \mu_2$) [4]. В научно-исследовательской практике намного чаще используются двусторонние статистические критерии.

В данной статье мы рассмотрим параметрические критерии, которые применяются в случае сравнения средних арифметических для одной или двух несвязанных групп.

Параметрические критерии называются параметрическими, так как оперируют параметрами распределения — средней арифметической и стандартным отклонением. Однако эти параметры корректно описывают только нормальное (Гауссово) распределение, в связи с чем проверка условия на нормальность распределения является обязательной.

Одновыборочный критерий Стьюдента применяется в случае, когда среднее арифметическое выборочной совокупности сравнивается со средним значением для некоторой генеральной совокупности или некоторой эталонной величиной.

Представлены теоретические основы применения одновыборочного и двухвыборочного критериев Стьюдента, практические рекомендации для применения вышеуказанных критериев с использованием пакета прикладных статистических программ STATA. В работе использованы практические примеры расчетов с помощью формул, а также алгоритм действий для использования пакета STATA с последующей интерпретацией результатов.

Ключевые слова: одновыборочный критерий Стьюдента, двухвыборочный критерий Стьюдента, параметрические критерии, средние арифметические

Условия применения: количественный признак; нормальное распределение признака в генеральной совокупности, из которой отобрана выборка; независимость наблюдений.

Алгоритм расчета одновыборочного критерия Стьюдента

$$t = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n},$$

где \bar{x} – среднее значение; μ_0 – эталонная величина; s – стандартное отклонение; n – число наблюдений.

Далее по таблице t-распределения с учетом количества степеней свободы ($df = n - 1$) сравнивается значение эмпирического уровня значимости с критическим значением.

Рассмотрим пример. В одном из неблагоприятных районов города Л. у подростков изучали возраст инициации табакокурения (табл. 1). Можно ли утверждать, что в данном районе возраст начала употребления табака меньше, чем по городу в целом? Средний возраст начала табакокурения в городе Л. составил 14,0 лет.

Таблица 1

Возраст начала табакокурения, годы																	
№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Возраст, x	14	13	13	16	16	10	15	12	14	13	13	16	15	17	13	10	15

1. Рассчитаем среднее арифметическое

$$\bar{x} = \frac{1}{n} \sum x = 13,8235$$

2. Рассчитаем стандартное отклонение

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} = 2,0073$$

3. $\mu_0 = 14,0$ $n = 17$

4. Подставляем в формулу

$$t = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n} = -0,3624$$

(далее учитывается значение t-критерия без учета знака)

5. $df = n - 1 = 16$

6. По табл. 2 сравниваем значение эмпирическое с критическим (табличным) [3, 6]

Таблица 2

Критические значения t-распределения						
df	p-уровень					
	0,2	0,1	0,05	0,02	0,01	0,001
1	3,08	6,31	12,71	31,82	63,66	318,3
2	1,89	2,92	4,30	6,96	9,92	22,33
3	1,64	2,35	3,18	4,54	5,84	10,21
4	1,53	2,13	2,78	3,75	4,60	7,17
...						
15	1,34	1,75	2,13	2,60	2,95	3,73
16	1,34	1,75	2,12	2,58	2,92	3,69

7. $t_{кр \text{ при } p = 0,05} = 2,12$ $t_{эмп} = 0,3624$. Следовательно, принимается нулевая гипотеза ($p > 0,05$), так как $t_{кр} > t_{эмп}$.

Рассмотрим пример использования одновыборочного критерия Стьюдента в программе STATA [8, 9]. Возьмем тот же самый пример. В одном из неблагоприятных районов города Л. у подростков изучали возраст инициации табакокурения (переменная age) (рис. 1).

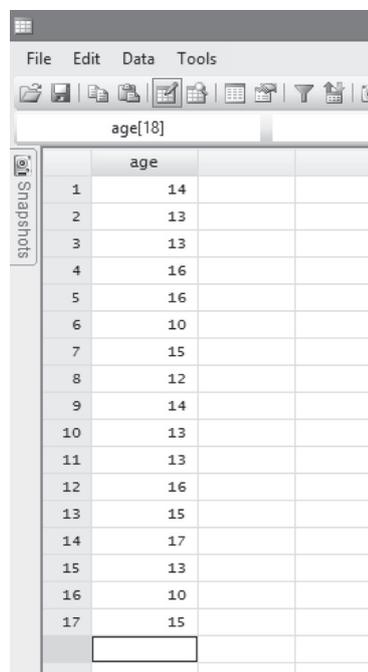


Рис. 1. Окно STATA Data Editor для примера использования одновыборочного критерия Стьюдента

Прежде чем приступить к расчету параметрического критерия, проверим условие нормального распределения количественного признака age с помощью статистического критерия Shapiro-Wilk (рис. 2).

```

swilk age
      Shapiro-wilk w test for normal data
+-----+-----+-----+-----+-----+
| variable | obs | w      | v      | z      | Prob>z |
+-----+-----+-----+-----+-----+
| age      | 17  | 0.93444 | 1.385  | 0.649  | 0.25801 |
+-----+-----+-----+-----+-----+
    
```

Рис. 2. Результаты проверки распределения переменной age

Полученный уровень статистической значимости ($p = 0,258$) позволяет нам принять нулевую гипотезу, свидетельствующую о том, что количественный признак переменной age имеет нормальное распределение.

Для применения одновыборочного критерия Стьюдента в STATA следует зайти в диалоговое окно One-sample mean-comparison test, которое открывается при помощи меню Statistics > Summaries, tables, and tests > Classical test of hypotheses > One-sample mean-comparison test (рис. 3).

Появляется диалоговое окно под названием ttest-Mean-comparison test. Во вкладке Main в окошке Variable name выбирается изучаемая переменная age (возраст); в окошке Hypothesized mean вводим эталонную величину и нажимаем Ok (рис. 4). Результаты сравнения выборочной совокупности с эталонной величиной представлены на рис. 5.

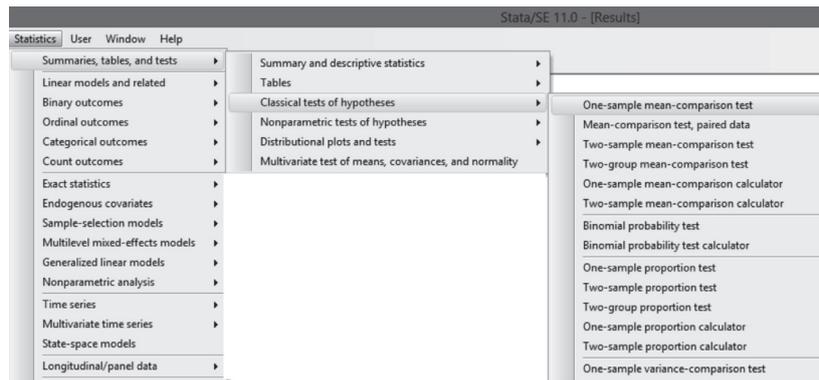


Рис. 3. Алгоритм поиска одновыборочного критерия Стьюдента в программе STATA

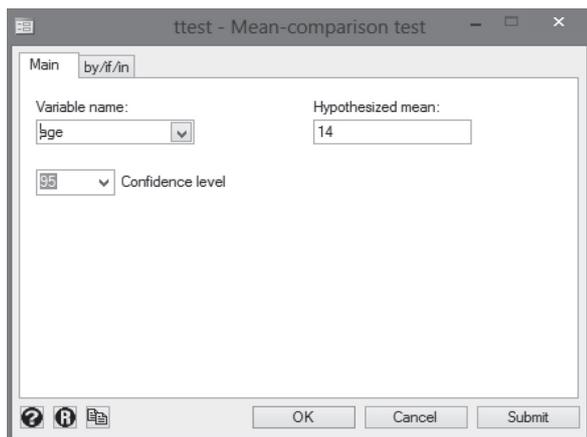


Рис. 4. Диалоговое окно One-sample mean-comparison test

```
. ttest age = 14
one-sample t test
variable | obs   Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
age      |  17   13.82353   .4868513    2.007339    12.79145   14.85561

mean = mean(age)          t = -0.3625
Ho: mean = 14             degrees of freedom = 16
Ha: mean < 14            Pr(T < t) = 0.3609
Ha: mean != 14           Pr(|T| > |t|) = 0.7217
Ha: mean > 14            Pr(T > t) = 0.6391
```

Рис. 5. Результат сравнения возраста инициации табакокурения в одном из неблагоприятных районов города Л. со средним возрастом начала употребления табака в городе Л. (с помощью одновыборочного критерия Стьюдента)

В таблице рис. 5 представлено количество наблюдений в группе (obs), среднее арифметическое (mean), ошибка среднего (std. err.), стандартное отклонение (std. dev.), 95 % доверительный интервал (95 % conf. interval). Под таблицей дается значение t-критерия ($t = -0,3625$), количество степеней свободы ($df = 16$), а также значения p-уровней для одностороннего статистического критерия и двустороннего. Значение уровня статистической значимости для двустороннего критерия составило 0,7217, что не позволяет отклонить нулевую гипотезу. Следовательно, возраст начала табакокурения в неблагоприятном районе города Л. не отличается от среднего возраста инициации употребления табака в городе Л. в целом.

Двухвыборочный критерий Стьюдента (или критерий Стьюдента для несвязанных выборок, или непарный критерий Стьюдента) применяется тогда, когда необходимо сравнить средние величины в двух независимых группах [1, 2, 7].

Условия применения: количественный признак; две независимые группы; независимые наблюдения в каждой из групп; нормальное распределение изучаемого признака в генеральных совокупностях, из которых отобраны группы; равенство дисперсий изучаемого признака в генеральных совокупностях, из которых отобраны группы.

Две независимые группы – это две разные группы, которые сравниваются по одному или нескольким признакам, причем одно и то же наблюдение не может быть отнесено к более чем одной группе; например, группа «мальчиков» и группа «девочек» сравниваются по массе тела при рождении или группа «поликлиника № 1» и группа «поликлиника № 2» сравниваются по количеству пациентов за неделю.

Алгоритм расчета критерия Стьюдента для несвязанных групп

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{06} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

где \bar{X}_1 – среднее значение признака в первой группе; \bar{X}_2 – среднее значение признака во второй группе; s_{06} – объединенное стандартное отклонение; n_1 – число наблюдений в первой группе; n_2 – число наблюдений во второй группе.

$$s_{06} = \sqrt{\frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}}$$

где n_1 – число наблюдений в первой группе; n_2 – число наблюдений во второй группе; s_1 – стандартное отклонение признака в первой группе; s_2 – стандартное отклонение признака во второй группе

$$df = (n_1 - 1) + (n_2 - 1),$$

где df – количество степеней свободы; n_1 – число наблюдений в первой группе; n_2 – число наблюдений во второй группе.

Далее по таблице t-распределения с учетом количества степеней свободы сравниваем эмпирический и критический уровни значимости [3, 6].

Рассмотрим пример. Изучался возраст возникновения язвенной болезни (ЯБ) у мужчин и женщин (табл. 3). Можно ли утверждать, что у мужчин ЯБ диагностируется в более молодом возрасте, чем у женщин?

Таблица 3

Возраст возникновения язвенной болезни, годы									
№	1	2	3	4	5	6	7	8	9
Женщины	23	28	19	30	27	25	20	35	
Мужчины	21	25	25	19	32	19	28	27	27

1. Рассчитаем среднее арифметическое и стандартное отклонение для каждой группы

$$\bar{X}_m = 24,78; s_m = 4,38; n_m = 8$$

$$\bar{X}_j = 24,78; s_j = 4,38; n_j = 8$$

2. Рассчитаем объединенное стандартное отклонение

$$s_{об} = \sqrt{\frac{(8 - 1) \cdot 4,38^2 + (9 - 1) \cdot 5,30^2}{8 + 9 - 2}} = 4,89$$

3. Рассчитаем значение t-критерия и количество степеней свободы

$$t = \frac{25,88 - 24,78}{4,89 \cdot \sqrt{\frac{1}{9} + \frac{1}{8}}} = 0,463$$

$$df = (8 - 1) + (9 - 1) = 15$$

4. По табл. 2 сравниваем значение эмпирическое с критическим (табличным)

5. $t_{кр}$ при $p = 0,05 = 2,13$; $t_{эмп} = 0,463$. Следовательно, принимается нулевая гипотеза ($p > 0,05$), так как $t_{кр} > t_{эмп}$.

Рассмотрим пример использования критерия Стьюдента для несвязанных выборок в программе STATA [8, 9]. Возьмем тот же самый пример. Изучался возраст возникновения язвенной болезни (переменная age) у мужчин (значение переменной gender – male) и женщин (значение переменной gender – female) (рис. 6).

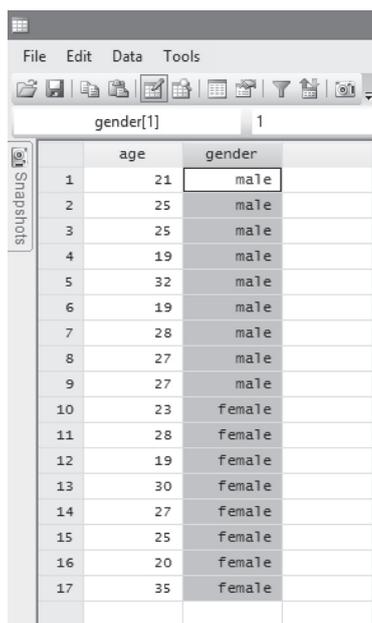


Рис. 6. Окно STATA Data Editor для примера использования критерия Стьюдента для несвязанных выборок

Прежде чем применять параметрический критерий Стьюдента для несвязанных выборок, нам необходимо проверить условия нормального распределения количественного признака (рис. 7) и равенства дисперсий (рис. 8).

```
-> gender = male
```

Shapiro-wilk w test for normal data					
variable	obs	w	v	z	Prob>z
age	9	0.93117	1.011	0.019	0.49260

```
-> gender = female
```

Shapiro-wilk w test for normal data					
variable	obs	w	v	z	Prob>z
age	8	0.97188	0.392	-1.355	0.91233

Рис. 7. Результат проверки условия нормального распределения переменной возраст (age) для каждой группы отдельно (группы мужчин = male и группы женщин = female)

Как в группе мужчин, так и в группе женщин переменная «возраст начала ЯБ» (age) распределена нормально ($p = 0,493$ и $p = 0,912$ соответственно).

Условие равенства дисперсий можно проверить по следующему алгоритму Statistics > Summaries, tables, and tests > Classical tests of hypotheses > Two-group variance-comparison test. Результаты свидетельствуют о том, что данное условие соблюдается, так как уровень статистической значимости составил 0,6016, что больше 0,05 (рис. 8).

Variance ratio test					
Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
male	9	24.77778	1.460382	4.381146	21.41013 28.14543
female	8	25.875	1.875	5.303301	21.44133 30.30867
combined	17	25.29412	1.143282	4.713872	22.87047 27.71777

ratio = sd(male) / sd(female) f = 0.6825
 Ho: ratio = 1 degrees of freedom = 8, 7
 Ha: ratio < 1 Pr(F < f) = 0.3008 Ha: ratio != 1 2*Pr(F < f) = 0.6016 Ha: ratio > 1 Pr(F > f) = 0.6992

Рис. 8. Результат проверки условия равенства дисперсий

Чтобы выбрать критерий Стьюдента для несвязанных выборок, необходимо зайти Statistics > Summaries, tables, and tests – Classical tests of hypotheses > Two-group mean-comparison test (см. рис. 3). Во вкладке Main в окошке Variable name выбирается изучаемая переменная age (возраст); в

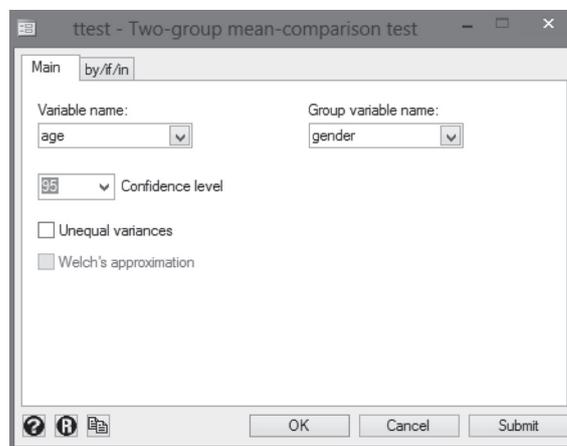


Рис. 9. Диалоговое окно критерия Стьюдента для несвязанных выборок

окошке Grouping variable name – группирующая переменная gender (пол) и Ok (рис. 9). В случае, когда дисперсии не равны, необходимо поставить галочку Unequal variances. Результаты сравнения групп представлены на рис. 10.

Group	obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
male	9	24.77778	1.460382	4.381146	21.41013	28.14543
female	8	25.875	1.875	5.303301	21.44133	30.30867
combined	17	25.29412	1.143282	4.713872	22.87047	27.71777
diff		-1.097222	2.348624		-6.103197	3.908752

diff = mean(male) - mean(female) t = -0.4672
 Ho: diff = 0 degrees of freedom = 15
 Ha: diff < 0 Pr(T < t) = 0.3235 Ha: diff != 0 Pr(|T| > |t|) = 0.6471 Ha: diff > 0 Pr(T > t) = 0.6765

Рис. 10. Результат сравнения возраста начала язвенной болезни (переменная age) у мужчин (переменная male) и женщин (переменная female)

В таблице рис. 10 представлено количество наблюдений в каждой группе (obs), средние арифметические (mean), ошибки среднего (std. err.), стандартные отклонения (std. dev.), 95 % доверительные интервалы (95 % conf. interval). Под таблицей дается значение t-критерия ($t = -0,4672$), количество степеней свободы ($df = 15$), а также достигнутые уровни значимости для односторонней и двусторонней статистических гипотез. Значение достигнутого уровня статистической значимости для двустороннего критерия составило 0,647, что не позволяет отклонить нулевую гипотезу. Следовательно, можно сделать вывод о том, что возраст возникновения ЯБ у мужчин и женщин не различается.

Представляя результаты применения критерия Стьюдента для несвязанных выборок, рекомендуется указывать значение t-критерия и абсолютную величину достигнутого уровня значимости (p). Пример такого представления в табл. 4.

Таблица 4

Особенности возраста начала язвенной болезни у мужчин и женщин, M (SD)

Признак	Пол		t	p
	Мужчины	Женщины		
Возраст начала язвенной болезни, годы	24,8 (4,38)	25,9 (5,30)	-0,467	0,647

В следующем выпуске мы рассмотрим применение непараметрических критериев для сравнения одной и двух независимых выборок.

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. Статистический анализ эмпирических исследований [Электронный ресурс]. URL: <http://statexpert.org> (дата обращения 06.01.2014).

4. Статистический словарь. Качественная клиническая практика [Электронный ресурс]. URL: <http://www.clininvest.ru/> (дата обращения 06.01.2014).

5. Титкова Л. С. Математические методы в психологии. Владивосток : Изд-во Дальневосточного университета, 2002. 85 с.

6. Электронный учебник по статистике StatSoft [Электронный ресурс]. URL: <http://statsoft.ru/> (дата обращения 20.02.2014).

7. Field A. Discovering statistics using SPSS. 2-nd edition. London : SAGE Publication, 2005. 781 p.

8. Hamilton L. C. Statistics with STATA: Updated for Version 10. USA, 2009. 400 p.

9. Stata: Release 12. Statistical Software. College Station, TX : StataCorp LP, 2011.

References

1. Glantz S. *Basic biostatistics* (translated from English into Russian). Moscow, Praktika, 1998. 459 p.
2. Grjibovski A. M. Analysis of quantitative data for two independent groups. *Ekologiya cheloveka* [Human Ecology]. 2008, 2, p. 54-61. [in Russian]
3. Statistical analysis of empirical research. Available at: <http://statexpert.org> (Accessed 6 January 2014).
4. Statistical dictionary. Quality clinical practice. Available at: <http://www.clininvest.ru/> (Accessed 6 January 2014).
5. Titkova L. S. *Mathematical methods in psychology*. Vladivostok, Far East University Publ., 2002, 85 p.
6. Electronic textbook on statistic StatSoft. Available at: <http://statsoft.ru/> (Accessed 20 February 2014).
7. Field A. Discovering statistics using SPSS. 2nd edition. SAGE Publication, London, 2005, 781 p.
8. Hamilton L. C. Statistics with STATA: Updated for Version 10. USA, 2009, 400 p.
9. Stata: Release 12. Statistical Software. College Station, TX, StataCorp LP, 2011.

ANALYSIS OF ONE AND TWO INDEPENDENT SAMPLES USING STATA SOFTWARE: PARAMETRIC TESTS

O. A. Kharkova, *A. M. Grjibovski

Northern State Medical University, Arkhangelsk, Russia
 *Norwegian Institute of Public Health, Oslo, Norway

In this paper we present theoretical principles of the use of one- and two-samples Student’s t-tests. Practical examples are presented with detailed explanations of the use of formulas as well as with step-by-step algorithm on how to use STATA software for comparisons of means in one- and two independent samples with subsequent interpretation of the outputs.

Keywords: one-sample t-test, t-test for two independent samples, parametric criteria, means

Контактная информация:

Гржибовский Андрей Мечиславович – профессор, доктор медицины, старший советник Норвежского института общественного здоровья
 Адрес: Nasjonalt folkehelseinstitutt, Pb 4404 Nydalen, 0403 Oslo, Norway
 Тел.: +47 22048319
 E-mail: angr@fhi.no