

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ ФЕДЕРАЛЬНОЕ
ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ «ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ» (ФГБОУ ВО «ВГУ»)

Факультет прикладной математики, информатики и механики

Отчет

на тему:

«Критерий t Стьюдента для зависимых выборок»

Выполнил: студент 3 к. 7 гр. ПМИ

Ноздрин Станислав Сергеевич

Проверил: преподаватель

Титова Ирина Александровна

Воронеж – 2020

Критерий Стьюдента t относится к одному из наиболее давно разработанных и широко используемых методов статистики. Чаще всего он применяется для проверки нулевой гипотезы о равенстве средних значений двух совокупностей, хотя существует также и одновыборочная модификация этого метода.

Начать, пожалуй, стоит с математических допущений, на которых основан критерий Стьюдента. Основных таких допущений, как известно, два:

- Сравниваемые выборки должны происходить из нормально распределенных совокупностей;
- Дисперсии сравниваемых генеральных совокупностей должны быть равны (для независимых выборок, хотя в принципе существует вариант теста в случае неравных дисперсий)

Зачастую требования о «нормальности» являются слишком сильными для выборочных данных (особенно если речь идет о «больших» данных), что приводит к многочисленным спорам о применимости критерия Стьюдента в той или иной ситуации.

На практике возможным компромиссом стало использование данного критерия не сколько в случае безоговорочности «нормальности» выборок, а столько в случае несущественных отклонений от нормальности, к ним относятся:

- Отсутствие выбросов в данных
- Отсутствие явной асимметрии гистограммы
- Отсутствие «смеси» распределений (например, бимодальности гистограммы)

Отметим, что если существуют серьезные отклонения от нормальности и исследователь не может позволить себе воспользоваться критерием Стьюдента, то стоит прибегнуть к непараметрической статистике: к тесту МакНемара, или к тесту Знаков, или же к тесту Вилкоксона (все – для случая парных выборок).

Зависимыми, или парными, являются две выборки, содержащие результаты измерений какого-либо количественного признака, выполненных на одних и тех же объектах. Во многих исследованиях какой-то определенный отклик измеряется у одних и тех же объектов до и после экспериментального воздействия. При такой схеме эксперимента исследователь более точно оценивает эффект воздействия именно потому, что прослеживает его у одних и тех же объектов.

t-критерий для двух зависимых (парных) выборок применяется, например, для оценки состояния больного до и после лечения. Нулевая гипотеза также гласит об отсутствии различий (среднее значение разности наблюдений в двух группах равно нулю).

$$H_0: \bar{Z} = 0, \quad z_i = x_{1i} - x_{2i}$$

$$H_1: \bar{Z} \neq 0$$

$$t = \frac{\bar{Z}}{s_z / \sqrt{n}} \sim t_{n-1},$$

$$s_z = \sqrt{\frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{n-1}}$$

где

Интерпретация результатов теста будет следующей: допустив, что нулевая гипотеза верна, мы можем рассчитать, насколько велика *вероятность* получить *t*-критерий, равный или превышающий то реальное значение, которое мы рассчитали по имеющимся выборочным данным. Если эта вероятность оказывается меньше, чем заранее принятый уровень значимости (например, $P < 0.05$), мы вправе отклонить проверяемую нулевую гипотезу.

Разберем как работать с данным критерием на конкретном примере.

Работать будем в пакете Statistica 12.

Итак, рассмотрим таблицу:

	1	2	3
	Номер участника	До изменения цвета	После изменения цвета
1	1	39,64	37,53
2	2	44,26	42,02
3	3	33,84	31,99
4	4	53,88	39,48
5	5	38,94	43,09
6	6	48,77	36
7	7	31,56	47,98
8	8	35,24	46,95
9	9	47,77	36,11
10	10	41,75	40,97
11	11	39,37	44
12	12	46,2	45,48
13	13	47,05	47,36
14	14	43,13	41
15	15	44,61	37,57
16	16	38,7	36,27
17	17	37,47	48,12
18	18	44,66	40,5
19	19	40,83	42,92
20	20	54,3	38,24
21	21	54,33	38,72
22	22	51,37	38,97

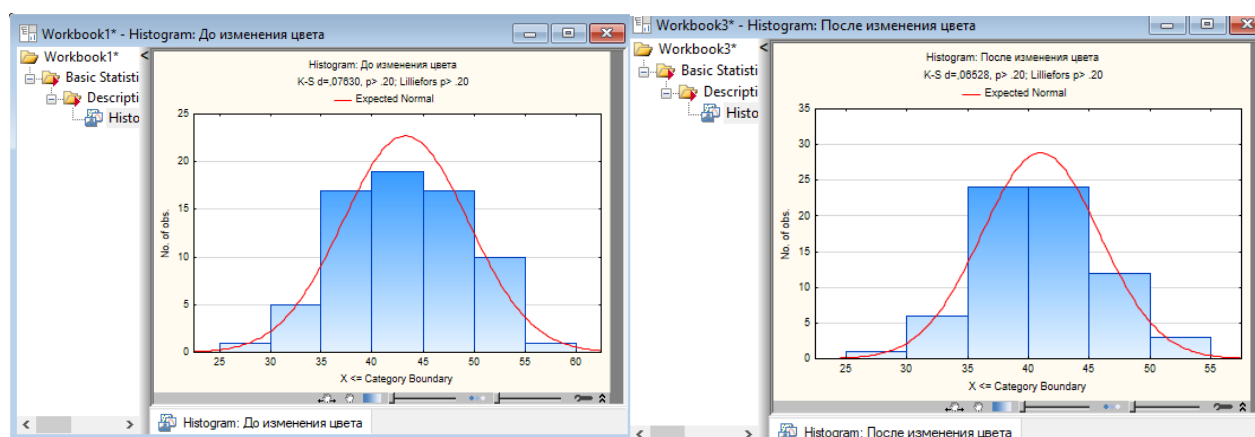
Смысл : перед нами данные исследования, проводившегося в 1935 году, в котором изучалась скорость чтения названия цветов в зависимости от того, шрифт какого цвета используется(то, как влияет фон на наше восприятие).

В ходе каждого теста суммировалось время, которое было затрачено на чтение слов(названия цветов).

Вопрос: смогли ли участники абстрагироваться от цвета букв? В первом случае цвет в каждом слове менялся(2 столбец), в другом – всегда черный(3 столбец). Это пример парной выборки.

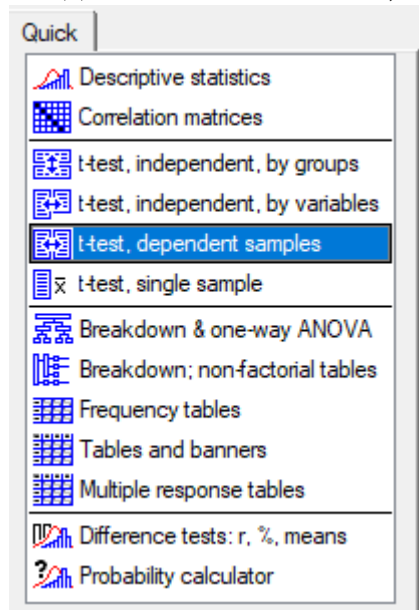
Выясним, существенны ли различия.

Прежде чем применить t-тест Стьюдента, убедимся, что наши выборки не имеют существенных отклонений от нормальности, посмотрим на гистограммы:

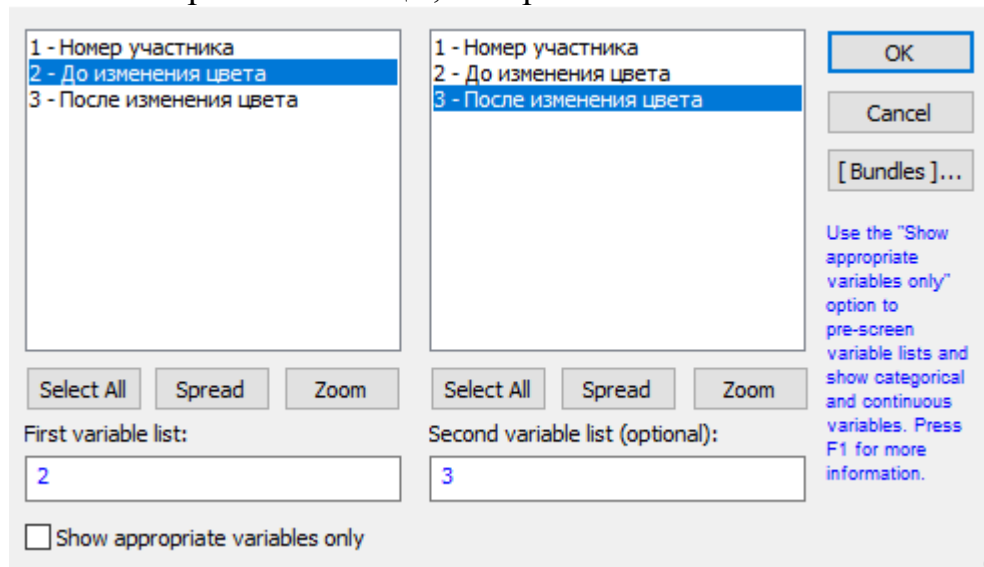


Заметим, что данные выдерживают наши требования, значит, мы можем применить критерий Стьюдента.

Зайдем в Basic Statistics, выберем t-test для зависимых(парных) выборок.



Затем выберем те столбцы, которые и являются объектом нашего исследования.



Посмотрим результат:

T-test for Dependent Samples (Spreadsheet1.sta)										
Marked differences are significant at $p < .05000$										
Variable	Mean	Std.Dv.	N	Diff.	Std.Dv. Diff.	t	df	p	Confidence -95,000%	Confidence +95,000%
До изменения цвета	43,30014	6,149710								
После изменения цвета	41,00014	4,840233	70	2,300000	7,810196	2,463854	69	0,016242	0,437725	4,162275

Заметим, что уровень значимости(p-value) меньше чем 0.05, и мы не принимаем основную гипотезу о том, что средние наших выборок не отличаются.

Однако если мы в качестве порога для уровня значимости возьмем 0.01, то в таком случае основная гипотеза будет не отвергнута.

Вопрос лишь в том, насколько сильно мы боимся ошибки 1 рода.

Больше моих работ об анализе данных, машинном обучении здесь:

