t критерий Стьюдента

Назначение критерия

- Критерий применяется в случае, когда стоит задача сравнить средние показатели двух распределений.

Ограничения критерия

- Распределения должны быть нормальными.
- $n \geq 2$

Эмпирическое значение критерия вычисляется по формуле:

- $$t = \frac{|X_{cp1} - X_{cp2}|}{\sqrt{m_1^2 + m_2^2}}$$
- m - ошибка среднего,
- Хср - средние арифметическое значение,
- n – объем соответствующей выборки.

Среднее арифметическое значение

- Это сумма всех значений выборки, поделенная на количество этих значений.
- Обозначается Хср и вычисляется по формуле:
- xi- i-е значение выборки;
- i – порядковый номер значения в выборке;
- n – объем выборки.

$$X = \frac{\sum_{i=1}^{n} x_i}{n}$$

Пример нахождения среднего арифметического значения

$$6,\ 7,\ 7,\ 7,\ 8,\ 8,\ 9,\ 9,\ 9,\ 10,\ 10,\ 10,\ 10,\ 10,\ 11,\ 11,\ 12,\ 12,\ 12,\ 14$$

$$X = \frac{6+7+7+7+8+8+9+9+9+10+10+10}{20} +$$

$$+ \frac{10+10+11+11+12+12+12+14}{20} = 9.1$$

формула для случая, когда значения повторяются.

- y_i – i-я варианта выборки,
- f_i – частота i-й варианты,
- k – количество вариант.

$$X = \frac{\sum_{i=1}^{k} f_i \cdot y_i}{n}$$

Так, для нашего случая мы получаем:

$$X = \frac{6+3\cdot 7+2\cdot 8+3\cdot 9+5\cdot 10+2\cdot 11+3\cdot 12+14}{20} = 9.1$$

Дисперсия

- мера разброса данных вокруг среднего арифметического значения. Обозначается S^2_x и вычисляется по формуле:

$$S^2{}_x = \frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2}{n-1}$$

Стандартное отклонение

- данная мера тесно связана с дисперсией, так как является квадратным корнем из нее. Обозначается σ и вычисляется по формуле:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \overline{X}\right)^2}{n-1}}$$

Пример расчета по t- критерию

- Различаются ли по средним значениям
- студенты с низким и высоким учебным
- стрессом по ситуативной тревожности,
- измеренной по методике Спилберга.

низкий	высокий
30	38
52	36
33	36
43	37
48	52
30	48
47	52
34	45
43	51
47	45
36	43
38	30
	48

Найдем значения для 1 выборки

x	f	x*f	f*(x-Xcp)^2
30	2	60	204,02
33	1	33	50,41
34	1	34	37,21
36	1	36	16,81
38	1	38	4,41
43	2	86	16,82
47	2	94	95,22
48	1	48	62,41
52	1	52	141,61
	n1=12	∑xf=481	∑f*(x-Xcp)^2=628,92
	Xcp1=481/12=40,1		Sx1=628,92/11=57,17
			m1^2=57,17/12=4,76

Найдем значения для 2 выборки

x	f	x*f	f*(x-Xcp)^2	
30	1	30	174,24	
36	2	72	103,68	
37	1	37	38,44	
38	1	38	27,04	
43	1	43	0,04	
45	2	90	6,48	
48	2	96	46,08	
51	1	51	60,84	
52	2	104	154,88	
	n2=13	∑xf=561	∑f*(x-Xcp)^2=611,72	
	Xcp2=561/13=43,2		Sx2=611,72/12=50,97	
			m2^2=50,97/13=3,92	

Значение критерия

- $$t = \frac{|40{,}1-43{,}2|}{\sqrt{4{,}76+3{,}92}} = 1{,}05$$

Гипотезы

- H1: Студенты с высоким учебным стрессом по среднему значению ситуативной тревожности значимо превосходят студентов с низким учебным стрессом
- H0: Студенты с высоким и низким учебным стрессом по среднему значению ситуативной тревожности статистически не различаются

Принятие гипотезы

- Число степеней свободы
- f=n1+n2-2
- f=12+13-2=23
- t0,01=2,807
- t0,05=2,069
- tэ=1,05<t0,05, следовательно
- Принимается H0

f	0,05	0,01	f	0,05	0,01
1	12,71	63,66	21	2,080	2,831
2	4,303	9,925	22	2,074	2,819
3	3,082	5,841	23	2,069	2,807
4	2,776	4,604	24	2,064	2,797
5	2,571	4,032	25	2,060	2,787
6	2,447	3,707	26	2,056	2,779
7	2,365	3,449	27	2,052	2,771
8	2,306	3,355	28	2,048	2,763
9	2,262	3,250	29	2,045	2,756
10	2,228	3,169	30	2,042	2,750
11	2,201	3,106	40	2,021	2,704
12	2,178	3,055	50	2,009	2,678
13	2,160	3,012	60	2,000	2,660
14	2,145	3,977	80	1,990	2,639
15	2,131	2,947	100	1,984	2,626
16	2,120	2,921	200	1,972	2,601
17	2,110	2,899	500	1,965	2,586
18	2,101	2,878			
19	2,093	2,861			
20	2,086	2,845			