

# Критерий $\chi^2$ Пирсона

# Назначение критерия

1. для сопоставления эмпирического распределения признака с теоретическим (равномерным, нормальным или каким-то иным);
2. для сопоставления двух эмпирических распределений одного и того же признака.

# Критерий отвечает на вопрос

- С одинаковой ли частотой встречаются разные значения признака в эмпирическом и теоретическом распределениях или в двух эмпирических распределениях?

# Ограничения

1.  $n \geq 30$ ;
2. Теоретическая частота для каждой ячейки таблицы не должна быть меньше 5:  $f_T \geq 5$ .

$f_T = n/k$ , где  $k$  – количество ячеек в таблице (видов, типов, уровней, оценок и т.д.)

Значение критерия при сопоставлении  
2х выборок или замеров

$$\chi^2 = \sum_{j=1}^k \left( \frac{(f_{\text{э}1j} - f_{m1j})^2}{f_{m1j}} + \frac{(f_{\text{э}2j} - f_{m2j})^2}{f_{m2j}} \right)$$

# Пример расчетов

- Различается ли распределение видов профессиональной идентичности у студентов 1 и 4 курсов
- Гипотезы:
- $H_1$ : Распределение видов профессиональной идентичности у студентов 1 и 4 курсов статистически различается
- $H_0$ : Распределение видов профессиональной идентичности у студентов 1 и 4 курсов статистически не различается

# Решение

	1 курс	4 курс					
	f1	f2	f1+f2	f <sub>T1</sub>	f <sub>T2</sub>	(f1-f <sub>T1</sub> ) <sup>2</sup> /f <sub>T1</sub>	(f2-f <sub>T2</sub> ) <sup>2</sup> /f <sub>T2</sub>
Диффузная	4	1	5	2,9	2,1	0,42	0,58
Мораторий на идентичность	12	8	20	11,6	8,4	0,01	0,02
Достигнутая позитивная идент	14	17	31	17,98	13,02	0,88	1,22
Псевдоидентичность	20	10	30	17,4	12,6	0,39	0,54
n	50	36				1,70	2,35
	$n_1/(n_1+n_2)=$	0,581395			$\chi:2=$	4,05	

# Принятие гипотезы

- Число степеней свободы  $v = k - 1 = 4 - 1 = 3$
- $\chi^2_{0,01} = 11,345$
- $\chi^2_{0,05} = 7,815$
- $\chi^2_{эм} = 4,05$
- Принимается гипотеза  $H_0$

v	ρ	
	0,05	0,01
1	3,841	6,635
2	5,991	9,210
3	7,815	11,345
4	9,488	13,277
5	11,070	15,086
6	12,592	16,812
7	14,067	18,475