

Значение признаков		Признак Y		Сумма
		Значение первое	Значение второе	
Признак X	Значение первое	$n_{11}$ <i>a</i>	$n_{12}$ <i>b</i>	$n_{1.}$ <i>a+b</i>
	Значение второе	$n_{21}$ <i>c</i>	$n_{22}$ <i>d</i>	$n_{2.}$ <i>c+d</i>
Сумма		$n_{.1}$ <i>a+c</i>	$n_{.2}$ <i>b+d</i>	$n$ <i>a+b+c+d</i>

## Коэффициент корреляции "φ"

$$\varphi_{\text{набл}} = \frac{n_{12} \cdot n_{21} - n_{11} \cdot n_{22}}{\sqrt{n_{.1} \cdot n_{.2} \cdot n_{1.} \cdot n_{2.}}}$$

$$\varphi_{\text{набл}} = \frac{b \cdot c - a \cdot d}{\sqrt{(a + c) \cdot (b + d) \cdot (a + b) \cdot (c + d)}}$$

$$\varphi_{\text{набл}} = \frac{\rho(x, y) - \rho(x) \cdot \rho(y)}{\sqrt{\rho(x) \cdot (1 - \rho(x)) \cdot \rho(y) \cdot (1 - \rho(y))}}$$

# Коэффициент корреляции

" $\varphi$ "

При **сравнении** двух переменных, измеренных **в дихотомической шкале**, мерой корреляционной связи служит так называемый **" $\varphi$ " коэффициент**, или **«коэффициент ассоциации»**.

# Коэффициент корреляции " $\varphi$ "

Коэффициент определяет **степень тесноты** двух дихотомистически измеренных признаков.

Величина коэффициента корреляции лежит в интервале **+1** до **-1**.

Коэффициент может быть **положительным** и **отрицательным**, характеризуюя **направленность связи** между двумя признаками, измеренными дихотомистически.

# Коэффициент корреляции

" $\varphi$ "

В общем виде формула вычисления коэффициента корреляции имеет вид:

$$\varphi_{\text{набл}} = \frac{\rho(x, y) - \rho(x) \cdot \rho(y)}{\sqrt{\rho(x) \cdot (1 - \rho(x)) \cdot \rho(y) \cdot (1 - \rho(y))}}$$

# Коэффициент корреляции

" $\varphi$ "

$$\varphi_{\text{набл}} = \frac{\rho(x, y) - \rho(x) \cdot \rho(y)}{\sqrt{\rho(x) \cdot (1 - \rho(x)) \cdot \rho(y) \cdot (1 - \rho(y))}}$$

$\rho(x)$  – **частота** или доля признака  $X$

**Способ вычисления:** количество единиц  $n_x$  в переменной  $X$

разделить на общее число элементов этой переменной  $n$ .

$$\rho(x) = \frac{n_x}{n}$$

# Коэффициент корреляции

" $\varphi$ "

$$\varphi_{\text{набл}} = \frac{\rho(x, y) - \rho(x) \cdot \rho(y)}{\sqrt{\rho(x) \cdot (1 - \rho(x)) \cdot \rho(y) \cdot (1 - \rho(y))}}$$

$\rho(y)$  – **частота** или доля признака  $Y$

**Способ вычисления:** количество единиц  $n_y$  в переменной  $Y$  разделить на общее число элементов этой переменной  $n$ .

$$\rho(y) = \frac{n_y}{n}$$

# Коэффициент корреляции

" $\varphi$ "

$$\varphi_{\text{набл}} = \frac{\rho(x, y) - \rho(x) \cdot \rho(y)}{\sqrt{\rho(x) \cdot (1 - \rho(x)) \cdot \rho(y) \cdot (1 - \rho(y))}}$$

$\rho(x, y)$  – **частота** или доля признака, имеющая 1 **одновременно** как по  $X$ , так и по  $Y$ .

**Способ вычисления:** количество единиц  $n_{xy}$  в переменных  $X, Y$  разделить на общее число элементов этой переменной  $n$ .

$$\rho(y) = \frac{n_{xy}}{n}$$

# Пример коэффициент корреляции " $\varphi$ "

**Пример** Влияет ли семейное положение на успешность учебы студентов-мужчин?

Для решения этой задачи психолог выясняет у 12 студентов мужчин, во-первых, женат он или холост, соответственно проставляя каждому **1 – женат** или **0 – холост**, и, во-вторых, насколько успешно тот учиться: **успешной** учебе проставляется **код 0**, при наличии академических задолженностей код **0**, при наличии академических **задолженностей** проставляется **код 1**.



# Пример коэффициент корреляции " $\varphi$ "

Данные опроса представлены в таблице

№	X – семейное положение 0- холост, 1- женат	Y – успешность обучения 0- успешно, 1- неуспешно
1	0	0
2	1	1
3	0	1
4	0	0
5	1	1
6	1	0
7	0	0
8	1	1
9	0	0
10	0	1
11	0	0
12	1	1

# Пример коэффициент корреляции "φ"

$$\rho(x) = \frac{n_x}{n} = \frac{5}{12} = 0,4167$$

№	X – семейное положение 0- холост, 1- женат	Y – успешность обучения 0- успешно, 1- неуспешно
1	0	0
2	1	1
3	0	1
4	0	0
5	1	1
6	1	0
7	0	0
8	1	1
9	0	0
10	0	1
11	0	0
12	1	1

# Пример коэффициент корреляции "φ"

$$\rho(y) = \frac{n_y}{n} = \frac{6}{12} = 0,5$$

№	X – семейное положение 0- холост, 1- женат	Y -успешность обучения 0- успешно, 1- неуспешно
1	0	0
2	1	1
3	0	1
4	0	0
5	1	1
6	1	0
7	0	0
8	1	1
9	0	0
10	0	1
11	0	0
12	1	1

# Пример коэффициент корреляции "φ"

$$\rho(x, y) = \frac{n_{xy}}{n} = \frac{4}{12} = 0,3333$$

№	X – семейное положение 0- холост, 1- женат	Y -успешность обучения 0- успешно, 1- неуспешно
1	0	0
2	1	1
3	0	1
4	0	0
5	1	1
6	1	0
7	0	0
8	1	1
9	0	0
10	0	1
11	0	0
12	1	1

# Пример коэффициент корреляции "φ"

Коэффициент корреляции «ассоциации» равен:

$$\varphi_{\text{набл}} = \frac{0,3333 - 0,4167 \cdot 0,5}{\sqrt{0,4167 \cdot (1 - 0,4167) \cdot 0,5 \cdot (1 - 0,5)}} = 0,507$$

# Второй способ вычисления коэффициента корреляции

Коэффициент корреляции можно вычислить используя **четырёхпольную** (корреляционную) **таблицу**. Каждую клетку таблицы обозначают соответствующими буквами

Значение признаков		Признак Y		Сумма
		Значение первое	Значение второе	
Признак X	Значение первое	$n_{11}$ <i>a</i>	$n_{12}$ <i>b</i>	$n_{1.}$ <i>a+b</i>
	Значение второе	$n_{21}$ <i>c</i>	$n_{22}$ <i>d</i>	$n_{2.}$ <i>c+d</i>
Сумма		$n_{.1}$ <i>a+c</i>	$n_{.2}$ <i>b+d</i>	$n$ <i>a+b+c+d</i>

# Второй способ вычисления коэффициента корр "φ"ции

$$\varphi_{\text{набл}} = \frac{n_{12} \cdot n_{21} - n_{11} \cdot n_{22}}{\sqrt{n_{.1} \cdot n_{.2} \cdot n_{1.} \cdot n_{2.}}}$$

Значение признаков		Признак Y		Сумма
		Значение первое	Значение второе	
Признак X	Значение первое	$n_{11}$ <i>a</i>	$n_{12}$ <i>b</i>	$n_{1.}$ <i>a+b</i>
	Значение второе	$n_{21}$ <i>c</i>	$n_{22}$ <i>d</i>	$n_{2.}$ <i>c+d</i>
Сумма		$n_{.1}$ <i>a+c</i>	$n_{.2}$ <i>b+d</i>	$n$ <i>a+b+c+d</i>

# Второй способ вычисления коэффициента корреляции

$$\varphi_{\text{набл}} = \frac{b \cdot c - a \cdot d}{\sqrt{(a + c) \cdot (b + d) \cdot (a + b) \cdot (c + d)}}$$

Значение признаков		Признак Y		Сумма
		Значение первое	Значение второе	
Признак X	Значение первое	$n_{11}$ <i>a</i>	$n_{12}$ <i>b</i>	$n_{1.}$ <i>a+b</i>
	Значение второе	$n_{21}$ <i>c</i>	$n_{22}$ <i>d</i>	$n_{2.}$ <i>c+d</i>
Сумма		$n_{.1}$ <i>a+c</i>	$n_{.2}$ <i>b+d</i>	$n$ <i>a+b+c+d</i>



# Второй способ вычисления коэффициента корр "φ"ции

Представим данные задачи о влиянии семейного положения на успешность учебы студентов-мужчин в виде четырехпольной таблицы

Значение признаков		Признак Y		Сумма
		Холостые	Женатые	
Признак X	Плохо учится	2 <i>a</i>	4 <i>b</i>	6 <i>a+b</i>
	Учится хорошо	5 <i>c</i>	1 <i>d</i>	6 <i>c+d</i>
Сумма		7 <i>a+c</i>	5 <i>b+d</i>	12 <i>a+b+c+d</i>

# Второй способ вычисления коэффициента корр "φ"ции

$$\varphi_{\text{набл}} = \frac{b \cdot c - a \cdot d}{\sqrt{(a+c) \cdot (b+d) \cdot (a+b) \cdot (c+d)}} = \frac{20 - 2}{\sqrt{7 \cdot 5 \cdot 6 \cdot 6}} = \frac{18}{\sqrt{1260}} = 0,507$$

Значение признаков		Признак Y		Сумма
		Холостые	Женатые	
Признак X	Плохо учится	2 <i>a</i>	4 <i>b</i>	6 <i>a+b</i>
	Учится хорошо	5 <i>c</i>	1 <i>d</i>	6 <i>c+d</i>
Сумма		7 <i>a+c</i>	5 <i>b+d</i>	12 <i>a+b+c+d</i>

# Условия применения коэффициента корреляции " $\varphi$ "

1. Сравнимые признаки должны быть измерены в дихотомической шкале.
2. Число варьирующих признаков в сравниваемых переменных  $X$  и  $Y$  должно быть одинаковым.

# Лабораторная работа 6

Вычислить коэффициент ассоциации зависимости выбора профессии от пола.

Данные представлены в таблице.

Значение признаков		Пол человека	
		Мужчина	Женщина
Профессии	Психолог	10	75
	Медик	13	72

# Указания к лабораторной работе 6

Обозначим буквами **a**, **b**, **c**, **d** клетки четырехпольной таблицы.

Значение признаков		Пол человека	
		Мужчина	Женщина
Профессии	Психолог	10 <i>a</i>	75 <i>b</i>
	Медик	13 <i>c</i>	72 <i>d</i>

# Указания к лабораторной работе 6

Заполните четырехпольную таблицу

Значение признаков		Пол человека		Сумма
		Мужчины	Женщины	
Профессии	Психолог	10 <i>a</i>	75 <i>b</i>	<i>a+b</i>
	Медик	13 <i>c</i>	72 <i>d</i>	<i>c+d</i>
Сумма		<i>a+c</i>	<i>b+d</i>	<i>a+b+c+d</i>

# Литература

