

Пример решения по правовой статистике

ЗАДАНИЕ.

Определите зависимость между воспитанием в семье и детской преступностью:

Воспитанием в семье	Подвергались приводу	Не подвергались приводу	Всего
Занимаются	50	150	200
Не занимаются	100	50	150
Итого	150	200	350

Для оценки связи между приведенными признаками определите коэффициенты ассоциации и контингенции и сделайте выводы.

РЕШЕНИЕ.

В случае альтернативных признаков (как у нас) составляется таблица сопряженности вида:

Таблица.

Сопряженность признаков

1-й признак	2-й признак		Всего
	да	нет	
Да	a	b	$a + b$
Нет	c	d	$c + d$
Итого	$a + c$	$b + d$	$n = a + b + c + d$

Для характеристики тесноты связи между признаками используются коэффициент ассоциации Юла и коэффициент контингенции Пирсона.

Коэффициент ассоциации Юла (K_J в соответствии с приведенной таблицей рассчитывают по формуле:

$$K_a = \frac{ad - bc}{ad + bc}.$$

Коэффициент ассоциации K_a может принимать значения от -1 до +1. В случаях когда $K_a = \pm 1$, теснота связи между признаками считается наиболее сильной, причем так же, как и для коэффициента корреляции, положительный или отрицательный знак K_a свидетельствует о прямой или соответственно обратной зависимости значений признаков.

Рассчитаем в нашем случае

$$K_a = \frac{50 * 50 - 100 * 150}{50 * 50 + 100 * 150} = -0.714.$$

Коэффициент контингенции Пирсона K_k рассчитывают по формуле:

$$K_k = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}.$$

Коэффициент контингенции также изменяется от -1 до +1, но его значения всегда (за исключением граничных случаев $K_k = \pm 1$)

несколько меньше значений коэффициента ассоциации. Эта характеристика имеет тот же смысл, что и K_a .

Подставляем:

$$K_k = \frac{50 * 50 - 100 * 150}{\sqrt{200 * 250 * 150 * 200}} \approx -0.323$$

Для качественной оценки силы связи при использовании коэффициента ассоциации Юла и коэффициента контингенции Пирсона можно руководствоваться шкалой Чеддока.

Таблица.

Шкала Чеддока

Коэффициент	0,1–0,3	0,3–0,5	0,5–0,7	0,7–0,9	0,9–0,99
Характеристика зависимости	Слабая	Умеренная	Заметная	Высокая	Весьма высокая

Таким образом, согласно коэффициенту ассоциации связь между признаками – высокая, а согласно коэффициенту контингенции – умеренная.