

РЕГРЕССИЯ ЛИНЕЙНАЯ ПАРНАЯ - причинная модель статистической связи линейной (см.) между двумя количественными переменными (см.) x и y , представленная уравнением $y = a + bx$, где x - переменная независимая (см.) (предиктор), y - переменная зависимая (см.) (см. также Анализ регрессионный). Коэффициент регрессии b и свободный член уравнения регрессии a вычисляются по формулам:

$$b = r s_y / s_x = \sum (x_i - \bar{x})(y_i - \bar{y}) / \sum (x_i - \bar{x})^2; a = \bar{y} - b\bar{x},$$

где r - коэффициент линейной корреляции Пирсона для переменных x и y ; s_x и s_y - стандартные отклонения (см.) для переменных x и y ; $\bar{x}, \bar{y}$ - средние арифметические (см.) для переменных x и y .

Существуют два подхода к интерпретации коэффициента регрессии b .

Согласно первому из них, b представляет собой величину, на которую изменяется предсказанное по модели значение $\hat{y}_i = a + bx_i$ при увеличении значения независимой переменной x на одну единицу измерения, согласно второй - величину, на которую в среднем изменяется значение переменной y_i при увеличении независимой переменной x на единицу. На диаграмме рассеяния (см.) коэффициент b представляет тангенс угла наклона линии регрессии $y = a + bx$ к оси абсцисс. Знак коэффициента регрессии совпадает со знаком коэффициента линейной корреляции: значение $b > 0$ свидетельствует о прямой линейной связи, значение $b < 0$ - об обратной. Если $b = 0$, линейная связь между переменными отсутствует (линия регрессии параллельна оси абсцисс).

Свободный член уравнения регрессии a интерпретируется, если для независимой переменной значение $x = 0$ имеет смысл. В этом случае $y = a$, если $x = 0$.

Качество (объясняющая способность) уравнения парной линейной регрессии оценивается с помощью коэффициента детерминации (см.).

О.В. Терещенко