

Понятия и формули:

- Вероятностен модел на проста линейна регресия:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, i = 1, \dots, n, \epsilon_i \sim N(0, \sigma^2).$$

- Оценки на регресионните коефициенти по метода на най-малките квадрати:

$$\begin{aligned}\hat{\beta}_1 &= \frac{S_{xy}}{S_{xx}}, \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x}, \text{ където} \\ S_{xx} &= \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}, \\ S_{xy} &= \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n} \\ S_{yy} &= \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n}.\end{aligned}$$

- Коефициент на корелация: $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$.
- Коефициент на определяне (coefficient of determination): $r^2 = \frac{S_{yy} - SSE}{S_{yy}}$, където $SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = S_{yy} - \hat{\beta}_1 S_{xy} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$.

Зад.1 Дадени са 5 точки със следните координати:

x	-2	-1	0	1	2
y	1	1	3	5	5

- а) намерете линейен модел по метода на най-малките квадрати
б) начертайте на графика получената в а) права и петте точки
Отговор: а) $\hat{y} = 3.0 + 1.2x$

Зад.2 Дадени са точки със следните координати:

x	1	2	3	4	5	6
y	5.6	4.6	4.5	3.7	3.2	2.7

- а) намерете линейен модел по метода на най-малките квадрати
б) начертайте на графика получената в а) права и шестте точки
Отговор: а) $\hat{y} = 6.0 - 0.557x$

Зад.3 Пресметнете коефициента на корелация за резултатите от тест за придобиване на степен, представени в следната таблица:

Студент	Тест по математика	Краен тест за придобиване на степен
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

Отговор: $r = 0.84$

Зад.4 Дадени са 5 точки със следните координати:

x	-2	-1	0	1	2
y	2	2	3	4	4

- а) намерете линеен модел по метода на най-малките квадрати
 б) начертайте на графика получената в а) права и петте точки. Базирайки се на чертежа, какъв ще бъде знакът на коефициента на корелация?
 в) пресметнете r и r^2
 Отговор: а) $\hat{y} = 3.0 + 6x$ б) положителен в) $r = 0.9487; r^2 = 0.9$

Зад.5 Дадени са следните данни:

x	1	2	3	4	5	6
y	7	5	5	3	2	0

- а) намерете линеен модел по метода на най-малките квадрати;
 б) начертайте на графика получената в а) права и шестте точки;
 в) пресметнете r и интерпретирайте;
 г) с какъв процент сумата от квадратите от отклоненията се редуцира, когато за предсказване на y се използва $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$, а не $\bar{y}$.
 Отговор: а) $\hat{y} = 8.267 - 1.314x$ в) -0.982 г) 96.5%