

10. Коэффициент корреляции

Для заданных выборок X и Y

$x_1 = 1,37$	$x_6 = -0,76$	$x_{11} = -0,56$	$x_{16} = 1,13$	$x_{21} = 1,55$	$x_{26} = 0,01$
$x_2 = 0,11$	$x_7 = -0,13$	$x_{12} = 1,28$	$x_{17} = -0,17$	$x_{22} = 0,29$	$x_{27} = -1,56$
$x_3 = 1,56$	$x_8 = -0,64$	$x_{13} = 1,16$	$x_{18} = 0,6$	$x_{23} = -2,16$	$x_{28} = 1,59$
$x_4 = -0,11$	$x_9 = -0,46$	$x_{14} = -0,3$	$x_{19} = -1,16$	$x_{24} = -0,77$	$x_{29} = -1,13$
$x_5 = 0,23$	$x_{10} = -0,88$	$x_{15} = -0,31$	$x_{20} = 2,65$	$x_{25} = 0,93$	$x_{30} = -1,74$

$y_1 = 0,08$	$y_6 = 0,89$	$y_{11} = 0,92$	$y_{16} = 0,47$	$y_{21} = 1,48$	$y_{26} = 0,59$
$y_2 = 0,64$	$y_7 = 1,44$	$y_{12} = 0,51$	$y_{17} = 0,11$	$y_{22} = 1,61$	$y_{27} = 0,34$
$y_3 = 1,59$	$y_8 = 0,72$	$y_{13} = 0,52$	$y_{18} = 1,27$	$y_{23} = 1,47$	$y_{28} = 0,17$
$y_4 = 1,75$	$y_9 = 1,26$	$y_{14} = 0,43$	$y_{19} = 1,33$	$y_{24} = 0,98$	$y_{29} = 0,20$
$y_5 = 0,74$	$y_{10} = 0,03$	$y_{15} = 1,06$	$y_{20} = 0,32$	$y_{25} = 0,54$	$y_{30} = 0,27$
					$y_{31} = 0,44$

- 1) вычислить выборочный коэффициент корреляции;
- 2) построить уравнения линейной регрессии Y на X и X на Y .

Решение

Объем второй выборки $n_2 = 31$ больше объема первой $n_1 = 30$, поэтому можно не учитывать последнее значение (0,44) во второй выборке, рассматривая только 30 первых значений.

- 1) выборочный коэффициент корреляции

$$\hat{r} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\tilde{S}_x \tilde{S}_y},$$

$$\overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \tilde{S}_x = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2}, \quad \tilde{S}_y = \sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2 - (\bar{y})^2}.$$

$$\hat{r} = \frac{\frac{1}{30} (1,37 \cdot 0,08 + 0,11 \cdot 0,64 + \dots + (-1,74) \cdot 0,27 - 0,05 \cdot 0,79)}{1,12 \cdot 0,51} \approx -0,08;$$

- 2) выборочное уравнение линейной регрессии Y на X имеет вид:

$$y = b_1 x + b_0, \quad \text{где } b_0 = \bar{y} - \hat{r} \frac{\tilde{S}_y}{\tilde{S}_x} \bar{x}, \quad b_1 = \hat{r} \frac{\tilde{S}_y}{\tilde{S}_x}.$$

$$b_0 = 0,79 - (-0,08) \cdot \frac{0,51}{1,12} \cdot 0,05 \approx 0,79, \quad b_1 = (-0,08) \cdot \frac{0,51}{1,12} \approx -0,03.$$

Таким образом, искомое уравнение:

$$y = -0,03x + 0,79 .$$

Выборочное уравнение линейной регрессии X на Y имеет вид:

$$x = b_1^* y + b_0^* ,$$

где $b_0^* = \bar{x} - \hat{r} \frac{\tilde{S}_X}{\tilde{S}_Y} \bar{y}$,

$$b_1^* = \hat{r} \frac{\tilde{S}_X}{\tilde{S}_Y} .$$

$$b_0^* = 0,05 - (-0,08) \cdot \frac{1,12}{0,51} \cdot 0,79 \approx 0,19, \quad b_1^* = (-0,08) \cdot \frac{1,12}{0,51} \approx -0,17 .$$

Таким образом, искомое уравнение:

$$x = -0,17 y + 0,19 .$$

На рисунке 2 изображены графики линейной регрессии.

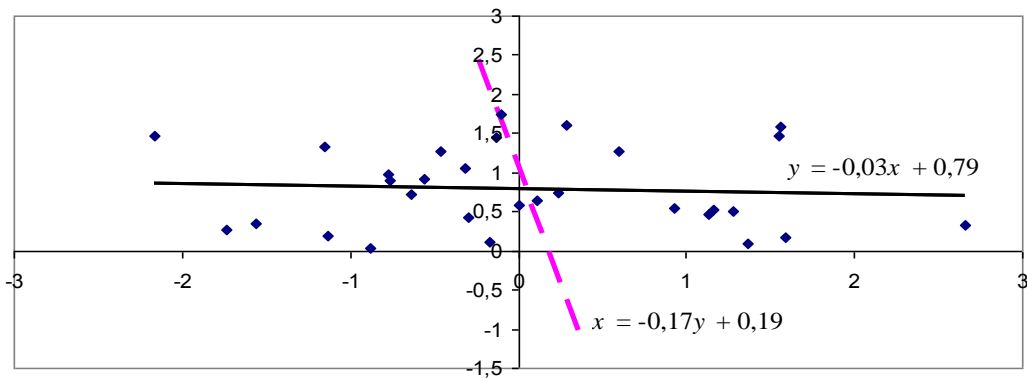


Рисунок 2