

9

相関係数（どのくらい正の/負の相関関係があるか）

2つの変量 x, y についてのデータの値が

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

と n 個あるとする。

x の偏差と y の偏差の積の平均値 を考えてみると

$$S_{xy} = \frac{1}{n} \{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})\}$$

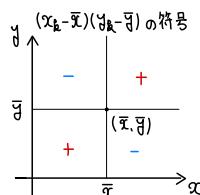
この値を 共分散 という。

共分散の正負は相関関係の正負の目安になり、

相関関係の強弱を

$$r = \frac{S_{xy}}{\sqrt{S_x S_y}} \quad (-1 \leq r \leq 1)$$

で表し、この値を 相関係数 という。



共分散	$S_{xy} = \frac{1}{n} \{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})\}$
相関係数	$r = \frac{S_{xy}}{\sqrt{S_x S_y}} \quad (-1 \leq r \leq 1)$

x	181	167	173	169	165
y	75	59	63	67	61
$x - \bar{x}$	10	-4	2	-2	-6
$(x - \bar{x})^2$	100	16	4	4	36
$y - \bar{y}$	10	-6	-2	2	-4
$(y - \bar{y})^2$	100	36	4	4	16
$(x - \bar{x})(y - \bar{y})$	100	24	-4	-4	24

$$S_x^2 = \frac{1}{5} (100 + 16 + 4 + 4 + 36) = 32$$

$$S_y^2 = \frac{1}{5} (100 + 36 + 4 + 4 + 16) = 32$$

$$S_{xy} = \frac{1}{5} (100 + 24 + (-4) + (-4) + 24) = 28$$

偏差の和 0

よって

$$r = \frac{S_{xy}}{\sqrt{S_x S_y}} = \frac{28}{\sqrt{32} \sqrt{32}} = \frac{28}{32} = 0.875 \quad \leftarrow \text{正の相関関係がある}$$

