信州大学

2016年 経済学部 第1問

$\boxed{1}$ 2つの変量 x, y のデータが，n 個の x, y の値の組として

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

のように与えられているとする．このとき，以下の問いに答えよ．

(1) x, y の平均値をそれぞれ $\overline{x}$, $\overline{y}$ とするとき，変量 x と y の共分散 s_{xy} は

$$s_{xy} = \frac{1}{n}\left(\sum_{k=1}^{n} x_k y_k\right) - \overline{x}\,\overline{y}$$

であることを示せ．

(2) これらのデータの間には，$y_k = ax_k + b$ $(k = 1,\ 2,\ \cdots,\ n)$ という関係があるとする．ただし，a, b は実数で，$a \neq 0$ である．変量 x の標準偏差 s_x は 0 でないとする．このとき，x と y の相関係数を求めよ．

(1) 共分散の定義より

$$s_{xy} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y})$$

展開する.

$$= \frac{1}{n}\sum_{k=1}^{n} x_k y_k - \overline{y}\cdot\underbrace{\frac{1}{n}\sum_{k=1}^{n} x_k}_{=\overline{x}} - \overline{x}\cdot\underbrace{\frac{1}{n}\sum_{k=1}^{n} y_k}_{=\overline{y}} + \frac{1}{n}\underbrace{\sum_{k=1}^{n}\overline{x}\,\overline{y}}_{=n\overline{x}\,\overline{y}}$$

$$= \frac{1}{n}\left(\sum_{k=1}^{n} x_k y_k\right) - \overline{y}\,\overline{x} - \overline{x}\,\overline{y} + \overline{x}\,\overline{y}$$

$$= \frac{1}{n}\left(\sum_{k=1}^{n} x_k y_k\right) - \overline{x}\,\overline{y} \quad ▨$$

(2) $s_x^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$ …① と $\overline{y} = a\overline{x} + b$ …② より

$$s_y^2 = \frac{1}{n}\sum_{k=1}^{n}(y_k - \overline{y})^2$$

$$= \frac{1}{n}\sum_{k=1}^{n}\{ax_k + b - (a\overline{x} + b)\}^2 \quad (\because ②より)$$

$$= a^2\cdot\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$$

$$= a^2 s_x^2 \quad (\because ①より)$$

また，$s_{xy} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y})$

$$= \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})\{ax_k + b - (a\overline{x} + b)\}$$

$$= a\cdot\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2$$

$$= a\cdot s_x^2$$

よって，$r = \dfrac{s_{xy}}{\sqrt{s_x^2}\sqrt{s_y^2}} = \dfrac{a\,s_x^2}{|a|\,s_x^2} = \dfrac{a}{|a|}$

$$\therefore\ r = \begin{cases} 1 & (a > 0 \text{ のとき}) \\ -1 & (a < 0 \text{ のとき}) \end{cases}$$

//