以下の表は，2 つの変量 $x$, $y$ のデータである．このとき，$x$ と $y$ の相関係数を求めよ．

| $x$ | 8 | 10 | 1 | 3 | 2 | 6 |
|---|---|---|---|---|---|---|
| $y$ | 9 | 4 | 3 | 3 | 1 | 4 |

(24 青森公立大 1(3))

【答】 0.625

【解答】

データの個数を $n$，$x$ と $y$ の平均をそれぞれ $\overline{x}$, $\overline{y}$，標準偏差をそれぞれ $\sigma_x$，$\sigma_y$ とおく．また，$x$ と $y$ の共分散を $\sigma_{xy}$，相関係数を $r$ とおく．

$x$ と $y$ の平均は

$$\overline{x} = \frac{1}{6}\sum_{k=1}^{6} x_k = \frac{8+10+1+3+2+6}{6} = \frac{30}{6} = 5$$

$$\overline{y} = \frac{1}{6}\sum_{k=1}^{6} y_k = \frac{9+4+3+3+1+4}{6} = \frac{24}{6} = 4$$

であるから，$x$ と $y$ の標準偏差は

$$\begin{aligned}\sigma_x &= \sqrt{\frac{1}{6}\sum_{k=1}^{6}(x_k-\overline{x})^2}\\ &= \sqrt{\frac{1}{6}\{(8-5)^2+(10-5)^2+(1-5)^2+(3-5)^2+(2-5)^2+(6-5)^2\}}\\ &= \sqrt{\frac{1}{6}(9+25+16+4+9+1)} = \sqrt{\frac{64}{6}} = \frac{8}{\sqrt{6}}\end{aligned}$$

$$\begin{aligned}\sigma_y &= \sqrt{\frac{1}{6}\sum_{k=1}^{6}(y_k-\overline{y})^2}\\ &= \sqrt{\frac{1}{6}\{(9-4)^2+(4-4)^2+(3-4)^2+(3-4)^2+(1-4)^2+(4-4)^2\}}\\ &= \sqrt{\frac{1}{6}(25+0+1+1+9+0)} = \sqrt{\frac{36}{6}} = \sqrt{6}\end{aligned}$$

である．$x$ と $y$ の共分散 $\sigma_{xy}$ は

$$\begin{aligned}\sigma_{xy} &= \frac{1}{6}\sum_{k=1}^{6}(x_k-\overline{x})(y_k-\overline{y})\\ &= \frac{1}{6}\{(8-5)(9-4)+(10-5)(4-4)+(1-5)(3-4)+(3-5)(3-4)\\ &\qquad +(2-5)(1-4)+(6-5)(4-4)\}\\ &= \frac{1}{6}(15+0+4+2+9+0)\\ &= \frac{30}{6}\\ &= 5\end{aligned}$$

である．

よって，$x$ と $y$ の相関係数 $r$ は

$$r = \frac{\sigma_{xy}}{\sigma_x\sigma_y} = \frac{5}{\frac{8}{\sqrt{6}}\cdot\sqrt{6}} = \frac{5}{8} = \mathbf{0.625} \qquad \cdots\cdots(答)$$

である．

- 標準偏差は次のように計算してもよい．

$$\begin{aligned}
\sigma_x &= \sqrt{\frac{1}{6}\sum_{k=1}^{6} {x_k}^2 - \overline{x}^2} = \sqrt{\frac{1}{6}(64+100+1+9+4+36) - 5^2} \\
&= \sqrt{\frac{214}{6} - 25} = \sqrt{\frac{64}{6}} = \frac{8}{\sqrt{6}} \\
\sigma_y &= \sqrt{\frac{1}{6}\sum_{k=1}^{6} {y_k}^2 - \overline{y}^2} = \sqrt{\frac{1}{6}(81+16+9+9+1+16) - 4^2} \\
&= \sqrt{\frac{132}{6} - 16} = \sqrt{\frac{36}{6}} = \sqrt{6}
\end{aligned}$$

である．