変量 x, y の相関係数について考える．

(1) 変量 x, y の値の組

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_{25},\ y_{25})$$

をデータ Z とする．図 1 は Z の散布図であり，Z から図 1 の点 A の表す値の組を除いた 24 個の値の組からなるデータを Z' とする．図 2 は Z' の散布図である．

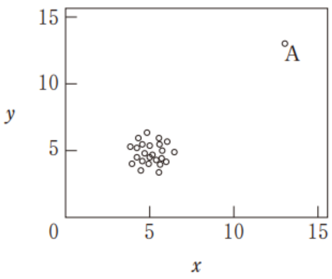

図 1　データ Z の散布図

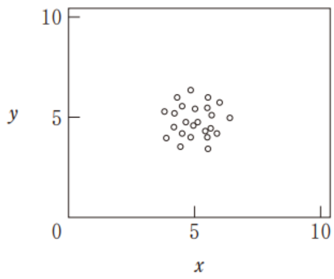

図 2　データ Z' の散布図

データ Z の x と y の相関係数を計算すると 0.8 となる．データ Z' の x と y の相関係数はおよそ $\boxed{\text{ク}}$ である．

$\boxed{\text{ク}}$ については，最も適当なものを，次の⓪〜⑤のうちから一つ選べ．

⓪ -0.8	① -0.6	② 0.0	③ 0.5	④ 0.7	⑤ 0.9

(2) 変量 x, y の値の組

$$(-1,\ -1),\ (-1,\ 1),\ (1,\ -1),\ (1,\ 1)$$

をデータ W とする．データ W の x と y の相関係数は 0 である．データ W に，新たに 1 個の値の組を加えたときの相関係数について調べる．なお，必要に応じて，後に示す表 1 の計算表を用いて考えてもよい．

a を実数とする．データ W に $(5a,\ 5a)$ を加えたデータを W' とする．W' の x の平均値 $\overline{x}$ は $\boxed{\text{ケ}}$，W' の x は y の共分散 s_{xy} は $\boxed{\text{コ}}$ となる．ただし，x と y の共分散とは，x の偏差と y の偏差の積の平均値である．

W' の x と y の標準偏差を，それぞれ s_x, s_y とする．積 $s_x s_y$ は $\boxed{\text{サ}}$ となる．また相関係数が 0.95 以上となるための必要十分条件は $s_{xy} \geqq 0.95 s_x s_y$ である．これより，相関係数が 0.95 以上となるような a の値の範囲は $\boxed{\text{シ}}$ である．

表 1　計算表

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{y})(x-\overline{y})$
-1	-1			
-1	1			
1	-1			
1	1			
$5a$	$5a$			

$\boxed{\text{ケ}}$ の解答群

⓪ 0　① $5a$　② $5a+4$　③ a　④ $a+\dfrac{4}{5}$

$\boxed{\text{コ}}$ の解答群

⓪ $4a^2$　① $4a^2+\dfrac{4}{5}$　② $4a^2+\dfrac{4}{5}a$　③ $5a^2$　④ $20a^2$

$\boxed{\text{サ}}$ の解答群

⓪ $4a^2+\dfrac{16}{5}a+\dfrac{4}{5}$　① $4a^2+1$

② $4a^2+\dfrac{4}{5}a$　③ $2a^2+\dfrac{2}{5}$

$\boxed{\text{シ}}$ の解答群

⓪ $-\dfrac{\sqrt{95}}{4} \leqq a \leqq \dfrac{\sqrt{95}}{4}$　① $a \leqq -\dfrac{\sqrt{95}}{4},\ \dfrac{\sqrt{95}}{4} \leqq a$

② $-\dfrac{\sqrt{95}}{5} \leqq a \leqq \dfrac{\sqrt{95}}{5}$　③ $a \leqq -\dfrac{\sqrt{95}}{5},\ \dfrac{\sqrt{95}}{5} \leqq a$

④ $-\dfrac{2\sqrt{19}}{5} \leqq a \leqq \dfrac{2\sqrt{19}}{5}$　⑤ $a \leqq -\dfrac{2\sqrt{19}}{5},\ \dfrac{2\sqrt{19}}{5} \leqq a$

(23 共通テスト 追・再試験 I 4[2])

【答】

ク	ケ	コ	サ	シ
2	3	0	2	3

【解答】

(1) 相関係数の値は -1 と 1 の間にある．データ Z' の散布図ではデータが点 $(5,\ 5)$ に集中しているから，選択肢として当てはまる値は　**0.0**　である．　(②)　……(答)

(2) W' の x の平均値 $\overline{x}$ は

$$\overline{x}=\frac{-1-1+1+1+5a}{5}=\boldsymbol{a}\qquad(③)$$ ……(答)

である．同じく W' の x の平均値 $\overline{y}$ は

$$\overline{y}=\frac{-1+1-1+1+5a}{5}=a$$

であり，計算表は次のようになる．

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{y})(x-\overline{y})$
-1	-1	$-1-a$	$-1-a$	$(1+a)^2$
-1	1	$-1-a$	$1-a$	a^2-1
1	-1	$1-a$	$-1-a$	a^2-1
1	1	$1-a$	$1-a$	$(1-a)^2$
$5a$	$5a$	$4a$	$4a$	$16a^2$

W' の x と y の共分散 s_{xy} は

$$\begin{aligned}s_{xy}&=(x\text{ の偏差と }y\text{ の偏差の積の平均値})\\&=\frac{(1+a)^2+2(a^2-1)+(1-a)^2+16a^2}{5}\\&=\frac{20a^2}{5}\\&=\mathbf{4a^2}\qquad(\text{⓪})\end{aligned}$$

……(答)

となる．

W' の x の標準偏差 s_x は

$$s_x{}^2=\frac{2(-1-a)^2+2(1-a)^2+(4a)^2}{5}=\frac{20a^2+4}{5}=4a^2+\frac{4}{5}$$

同じく，W' の y の標準偏差 s_y は

$$s_y{}^2=4a^2+\frac{4}{5}$$

であり，積 s_xs_y は

$$s_xs_y=\sqrt{4a^2+\frac{4}{5}}\sqrt{4a^2+\frac{4}{5}}=\mathbf{4a^2+\frac{4}{5}}\qquad(\text{②})$$

……(答)

となる．

また，相関係数 $\dfrac{s_{xy}}{s_xs_y}$ が 0.95 以上となるための必要十分条件は

$$\frac{s_{xy}}{s_xs_y}\geqq 0.95\qquad\therefore\quad s_{xy}\geqq 0.95s_xs_y$$

であるから，a の値の範囲は

$$\begin{aligned}&4a^2\geqq 0.95\left(4a^2+\frac{4}{5}\right)\\&(1-0.95)a^2\geqq 0.95\times\frac{1}{5}\\&5a^2\geqq 19\end{aligned}$$

$$\therefore\quad \boldsymbol{a\leqq -\frac{\sqrt{95}}{5},\ \frac{\sqrt{95}}{5}\leqq a}\qquad(\text{③})$$

……(答)

である．