

データの分析 (2) データの変換・相関係数

<データの変換>

① x_k を a 倍にすると

$$\bar{x} = \frac{1}{n} (\underbrace{x_1 + x_2 + \dots + x_n}_{\text{カッコ内のデータが } a \text{ 倍、その和}}) \quad \Rightarrow \quad y = ax \Rightarrow \bar{y} = a\bar{x}$$

カッコ内のデータが a 倍、その和

$$s^2 = \frac{1}{n} \{ \underbrace{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}_{\text{カッコ内が } a \text{ 倍され、その2乗の和}} \} \Rightarrow y = ax \Rightarrow s_y^2 = a^2 s_x^2$$

カッコ内が a 倍され、その2乗の和

② x_k に b を足すと

$$\bar{x} = \frac{1}{n} (\underbrace{x_1 + x_2 + \dots + x_n}_{\text{カッコ内のデータが } b \text{ 増加し、その和}}) \quad \Rightarrow \quad y = x + b \Rightarrow \bar{y} = \bar{x} + b$$

カッコ内のデータが b 増加し、その和

$$s^2 = \frac{1}{n} \{ \underbrace{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}_{\text{カッコ内が不変、その2乗の和}} \} \Rightarrow y = x + b \Rightarrow s_y^2 = s_x^2$$

カッコ内が不変、その2乗の和

まとめ

① データを a 倍すると

平均 a 倍 分散 a^2 倍 標準偏差 $|a|$ 倍

② データに b を足すと

平均 $+b$ 分散 不変 標準偏差 不変

<例題> あるクラスの生徒を対象に 100 点満点の試験を行ったところ、平均値は 60 点、分散は 36 であった。得点調整のため、生徒全員の得点を 0.5 倍して、更に 50 点を加えたとき、得点調整後の平均値、分散、標準偏差を求めよ。

① 変量 x のデータが次のように与えられている。

672, 693, 644, 665, 630, 644

いま、 $c=7$, $x_0=630$, $u = \frac{x-x_0}{c}$ として新たな変量 u を作る。

(1) 変量 u のデータの平均値と標準偏差を求めよ。

(2) 変量 x のデータの平均値と標準偏差を求めよ。

<相関係数>

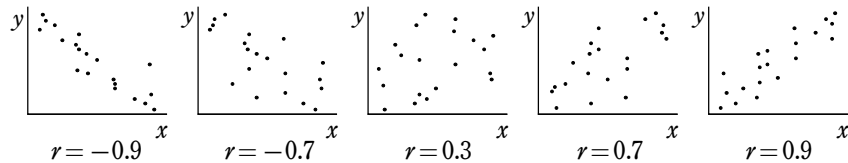
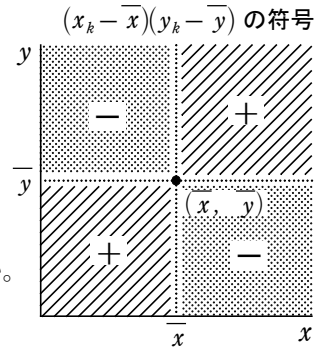
相関係数

$$r = \frac{s_{xy}}{s_x s_y} \leftarrow \frac{(x \text{ と } y \text{ の共分散})}{(x \text{ の標準偏差}) \times (y \text{ の標準偏差})}$$

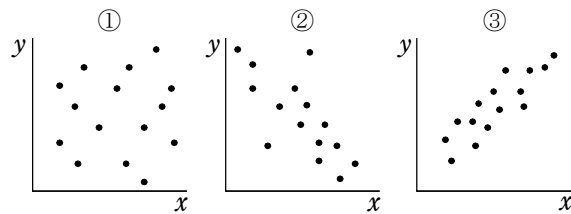
$$= \frac{\frac{1}{n} \{(x_1 - \bar{x})(y_1 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})\}}{\sqrt{\frac{1}{n} \{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2\}} \sqrt{\frac{1}{n} \{(y_1 - \bar{y})^2 + \dots + (y_n - \bar{y})^2\}}}$$

相関係数 r について

- [1] $-1 \leq r \leq 1$
- [2] $r=1$ のとき、右上がり
- [3] $r=-1$ のとき、右下がり
- [4] r の値が 0 に近いとき、相関関係はない。
 r の値が 1 に近いことは正の相関関係が強い。
 r の値が -1 に近いことは負の相関関係が強い。



- [2] 右の①, ②, ③は、ある2つの変数 x, y のデータについての散布図である。データ①, ②, ③の x と y の相関係数は、



0.87, 0.04, -0.71

のいずれかである。各データの相関係数を答えよ。

- [3] 次の表は、10人の生徒に10点満点の国語と数学の小テストを行った結果である。

生徒の番号	1	2	3	4	5	6	7	8	9	10
国語	6	4	7	5	8	6	2	9	8	5
数学	7	10	2	4	3	4	4	5	4	7

国語と数学の得点の間には、どのような相関関係があると考えられるか。相関係数 r を計算して答えよ。ただし、 $\sqrt{5} = 2.24$ とし、小数第3位を四捨五入せよ。