1　統計

1.1　相関係数 (correlation coefficient)

分散　n個の資料$x_1, x_2, \ldots, x_n$の分散σ_x^2は

$$\sigma_x^2 = \frac{1}{n} \sum (x_i - \bar{x})$$

と定義する。偏差の 2 乗和。

共分散　2変量(x, y)をもつn個の資料$(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$における共分散$\sigma_{xy}$は

$$\sigma_{xy} = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$$
$$(\text{ただし、}\bar{x} = \frac{1}{n} \sum x_i, \bar{y} = \frac{1}{n} \sum y_i)$$

と定義する。xの偏差とyの偏差の積の平均。

相関係数の定義　2変量(x, y)を持つn個の資料を$(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$とする。
このとき、相関係数 r は

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

と定義される。xとyの共分散を標準偏差の積で割る。
また、分子のnとxとyの標準偏差をかけることで生まれるnが打ち消しあうので以下の定義も可能

$$\begin{aligned} r &= \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum (y_i - \bar{y})^2}} \\ &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \end{aligned}$$

相関係数の絶対値が 1 以下の証明　相関係数rが$-1 \leq r \leq 1$となることを証明する。

まず

$$x_i - \bar{x} = a_i \qquad y_i - \bar{y} = b_i \qquad a_i b_i = c_i$$
$$\sum {a_i}^2 = A \qquad \sum {b_i}^2 = B \qquad \sum a_i b_i = C$$

とおくと

$$
\begin{aligned}
&-1 \leq r \leq 1 \\
&-1 \leq \frac{C}{\sqrt{A}\sqrt{B}} \leq 1 \\
&-c \leq x \leq c \leftrightarrow x^2 \leq c^2 \text{なので} \\
&\quad (\frac{C}{\sqrt{A}\sqrt{B}})^2 \leq 1^2 \\
&\quad \frac{C^2}{AB} \leq 1 \\
&\quad C^2 \leq AB \qquad (AB \geq 0) \\
&\quad C^2 - AB \leq 0 \\
&\quad AB - C^2 \geq 0
\end{aligned}
$$

のように変形できるのでこれを証明する

ここで以下の関数を考える

$$
\begin{aligned}
f(x) &= \sum (a_i x - b_i)^2 \\
&= \sum ({a_i}^2 x^2 - 2a_i b_i x + {b_i}^2) \\
&= \sum {a_i}^2 x^2 - 2(\sum a_i b_i) + \sum {b_i}^2 \\
&= Ax^2 - 2Cx + B
\end{aligned}
$$

これを平方完成すると

$$
\begin{aligned}
&= A(x^2 - \frac{2C}{A}x) + B \\
&= A\{(x - \frac{C}{A})^2\} - \frac{C^2}{A} + B \\
&= A\{(x - \frac{C}{A})^2\} + \frac{AB - C^2}{A}
\end{aligned}
$$

f(x) は 2 乗の和なので、f(x) ≥0

$x = \frac{C}{A}$のとき$f(x)$は最小値$\frac{AB - C^2}{A}$をとるので

$$
\begin{aligned}
&\frac{AB - C^2}{A} \geq 0 \\
&AB - C^2 \geq 0 \qquad (A \geq 0)
\end{aligned}
$$

となり、証明できた。