

## 第 6 講義—回帰直線と相関係数

戸瀬 信之

ITOSE PROJECT

経済数学, 2019 年 05 月 21 日 at HC

# モデル

- 2変量のデータ

モデル

$$y = ax + b$$

でデータに fit するものを考える。

$x_1$	$y_1$
$x_2$	$y_2$
$\vdots$	$\vdots$
$x_n$	$y_n$

- (誤差) = (実測値) - (理論値)

$$\epsilon = y - (ax + b)$$

- (1)  $\bar{\epsilon} = 0$

$$\bar{\epsilon} = \bar{y} - a\bar{x} - b = 0$$

- (2)  $V(\epsilon)$  が最小

## 最小分散

- (2)  $V(\epsilon)$  が最小

$$\epsilon_i = y_i - ax_i - b = y_i - ax_i - (\bar{y} - a\bar{x}) = y_i - \bar{y} - a(x_i - \bar{x})$$

から (このとき  $\bar{\epsilon} = 0$  に注意して)

$$\vec{\epsilon} = \frac{1}{\sqrt{n}}(y_i - \bar{y} - a(x_i - \bar{x})) = \vec{y} - a\vec{x}$$

- 

$$\begin{aligned} V(\epsilon) &= \|\vec{y} - a\vec{x}\|^2 \\ &= a^2\|\vec{x}\|^2 - 2a(\vec{x}, \vec{y}) + \|\vec{y}\|^2 \\ &= \|\vec{x}\|^2 \left( a - \frac{(\vec{x}, \vec{y})}{\|\vec{x}\|^2} \right)^2 + \|\vec{y}\|^2 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2} \end{aligned}$$

## 回帰直線

- 回帰直線

$$a = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\|^2} = \frac{C_{xy}}{V(x)}, \quad b = \bar{y} - a\bar{x}$$

- 回帰直線に対して

$$\begin{aligned} V(\epsilon) &= \|\vec{y}\|^2 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2} \\ &= \|\vec{y}\|^2 \left( 1 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2 \|\vec{y}\|^2} \right) = V(y)(1 - \rho_{xy}^2) \end{aligned}$$

- $\rho_{xy} \rightarrow \pm 1$  のとき  $V(\epsilon) \rightarrow 0$

# 最小二乗法



$$S(a, b) := \frac{1}{n} \sum_{i=1}^n (y_i - ax_i - b)^2$$

を最小にする  $a$  と  $b$  を求めても、同じ回帰直線を得る。