X : 母数 N , P の 2 項変数
parameter　binomial variable

$$P(X=i) = {}_N C_i \, p^i q^{N-i}$$

$(q = 1 - p)$

θ : 定数

$$f(\theta) = E(e^{\theta X}) = \sum_{i=0}^{N} e^{\theta i} P(X=i)$$

X のモーメント母関数.

$$= \sum_{i=0}^{N} e^{\theta i} \, {}_N C_i \, p^i q^{N-i}$$

$$= \sum_{i=0}^{N} {}_N C_i \, (e^{\theta} p)^i q^{N-i}$$

$$= (e^{\theta} p + q)^N$$

$$f(\theta) := E(e^{\theta X}) = \sum_{i=0}^{N} e^{\theta i} P(X=i)$$

$$= (e^{\theta} p + q)^N$$

$$f'(\theta) = \sum_{i=0}^{N} i\, e^{\theta i} P(X=i)$$

$$= N (e^{\theta} p + q)^{N-1} \cdot p e^{\theta}$$

$$f''(\theta) = \sum_{i=0}^{N} i^2 e^{\theta i} P(X=i)$$

$$= N(N-1)(e^{\theta} p + q)^{N-2} \cdot p^2 e^{2\theta}$$

$$+ N (e^{\theta} p + q)^{N-1} \cdot p e^{\theta}$$

zero

↓

$$f'(\theta) = \sum_{i=0}^{N} i\, P(X=i) = Np.$$

$$\leadsto \quad E(X) = Np.$$

$$f''(0) = N(N-1)p^2 + Np.$$

$$= \sum_{i=0}^{N} i^2 P(X=i) = E(X^2)$$

2次のモーメント.
(原点の回りの)

$$f^{(k)}(0) = E(X^k)$$

→ k次のモーメント.

モーメント法

現実 ⟷ 確率モデル.

sampling 　　確率変数. X

$x_1, \cdots, x_n$ 　　パラメータ. $\theta_1, \cdots, \theta_k$

$$\frac{1}{n}\sum_{i=1}^{n} x_i^r = E(X^r)$$

$(r = 1, \cdots, k)$

右辺は θ の関数.

X : $\mathbb{Z}_+$ 値の確率変数.

$$f(\theta) = E(e^{\theta X})$$

モーメント母関数

$$= \sum_{i=0}^{+\infty} e^{\theta i} P(X = i)$$

$g_k(\theta)$ (a, b)上微分可能.

1) $\sum_{k=0}^{+\infty} |g_k(\theta)| < +\infty$

2) $\sum_{k=0}^{+\infty} |g_k'(\theta)| < +\infty$

$\Rightarrow$ (1) から) $g(\theta) = \sum_{k=0}^{+\infty} g_k(\theta)$ 収束

$$g'(\theta) = \sum_{k=0}^{+\infty} g_k'(\theta)$$

統計の話

$x_1, \cdots, x_n \longleftrightarrow X$ 変量

$\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n) = \mu$ 平均

$V(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$ 分散

---

x	x_1	x_2	$\cdots$	x_n
y	y_1	y_2	$\cdots$	y_n

$\longleftrightarrow (X, Y)$

$C_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

共分散.

$\rho_{xy} = \dfrac{C_{xy}}{\sqrt{V(x)}\sqrt{V(y)}}$ 相関係数.

$$\vec{x} = \frac{1}{\sqrt{n}}\begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}, \quad \vec{y} = \frac{1}{\sqrt{n}}\begin{pmatrix} y_1 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{pmatrix}$$

$$\|\vec{x}\|^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = V(x)$$

$$(\vec{x}, \vec{y}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = C_{xy}.$$

今井先 2組 のポイント.

$\vec{a}, \vec{b} \in \mathbb{R}^N$ コーシー・シュヴァルツの不等式:

$$|(\vec{a}, \vec{b})| \leq \|\vec{a}\| \cdot \|\vec{b}\|.$$

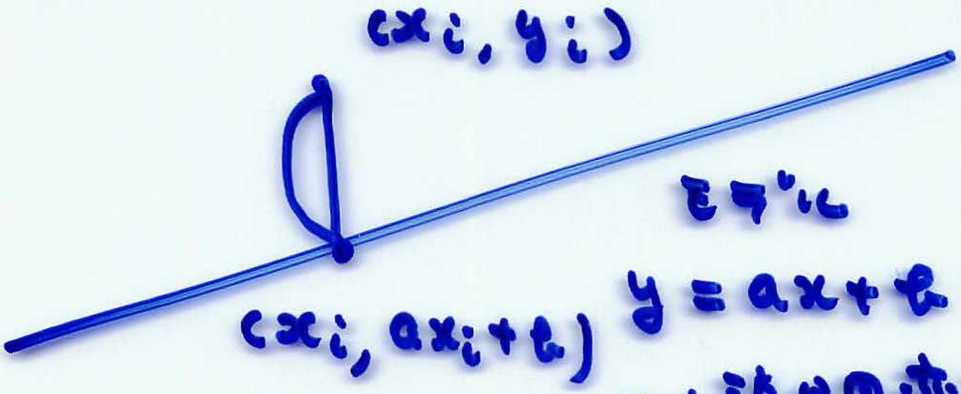
$N=2, 3$ $(\vec{a}, \vec{b}) = \|\vec{a}\| \|\vec{b}\| \cos\theta$

θ はなす角.

$$\rho_{xy} = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\| \|\vec{y}\|}$$

$$-1 \leq \rho_{xy} \leq 1$$

$|\rho_{xy}| \sim 1$ データは直線状に分布.



モデル

$y = ax + b$

x : 説明変数

y : 目的 —

x : 従属 対し

$ax_i + b$: 理論値

y_i : 実測値.

$\varepsilon_i = y_i - (ax_i + b)$ 誤差.

$$R = \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i^2$$ 最小にする a, b をとる.

$$= \frac{1}{N}\sum_{i=1}^{N} (y_i - a x_i - b)^2$$

$$\frac{\partial R}{\partial b} = \frac{1}{N}\sum_{i=1}^{N} 2(y_i - a x_i - b) \times (-1) = 0$$

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

$$R = \frac{1}{N}\sum_{i=1}^{N} \{(y_i - a x_i) - (\bar{y} - a\bar{x})\}^2$$

$$= \frac{1}{N}\sum_{i=1}^{N} \{(y_i - \bar{y}) - a(x_i - \bar{x})\}^2$$

$$= \|\vec{y} - a\vec{x}\|^2$$

$$\|\vec{a} \pm \vec{b}\|^2 = \|\vec{a}\|^2 \pm 2(\vec{a}, \vec{b}) + \|\vec{b}\|^2$$

$$= a^2\|\vec{x}\|^2 - 2a(\vec{x}, \vec{y}) + \|\vec{y}\|^2$$

$$= \|\vec{x}\|^2 \left(a - \frac{(\vec{x}, \vec{y})}{\|\vec{x}\|^2}\right)^2 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2} + \|\vec{y}\|^2$$

$$a = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\|^2} = \frac{c_{xy}}{V(x)}$$

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

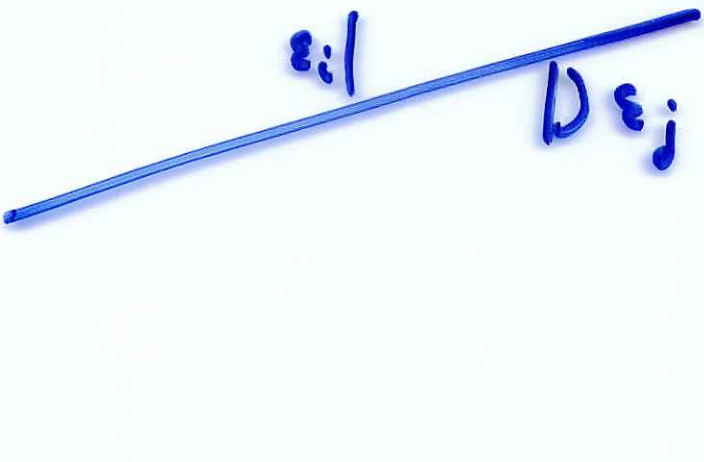
$\longrightarrow$ $y = ax + b$ 回帰直線.

$$R = \sum_{i=1}^{n} \varepsilon_i^2 = \|\vec{y}\|^2 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2}$$

$$= \|\vec{y}\|^2 \left(1 - \frac{(\vec{x}, \vec{y})^2}{\|\vec{x}\|^2 \cdot \|\vec{y}\|^2}\right)$$

$$= V(y)\left(1 - \rho_{xy}^2\right)$$

$|\rho_{xy}| \to 1$ a b z $R \to 0$



不 幾何変数 X を考える　$p+q=1$

$$P(X=i) = p^i q$$

に対して $\varphi(\theta) = E(e^{\theta X})$ を求めよ.

$e^{\theta X}$

$$= \sum_{i=0}^{+\infty} e^{\theta i} p^i q = \cdots$$