

# 確率・統計 B

## 標本分布・推定

若木宏文

wakaki@math.sci.hiroshima-u.ac.jp

2013.12.4

# table of contents

正規母集団からの統計量の分布

順序統計量

点推定

最小 2 乗法と単回帰モデル

一様最小分散不偏推定量

## 復習

## 正規母集団からの統計量の分布 (1/9)

### 定理 6.8

$X_1, \dots, X_n \stackrel{i.i.d.}{\sim} N(\mu, \sigma^2)$  とする. このとき, 次が成り立つ.

- (1)  $\bar{X}_n \sim N(\mu, \sigma^2/n)$ .
- (2)  $(n-1)S_n^2/\sigma^2 \sim \chi_{n-1}^2$ .
- (3)  $\bar{X}_n$  と  $S_n$  は互いに独立である.

## 順序統計量 (1/5)

### 定義 6.12

$X_1, \dots, X_n$  を大きさ  $n$  のランダム標本とする. これを大きさの順に並べて  $X_{(1)} \leq \dots \leq X_{(n)}$  とするとき

$$X_{(1)}, \dots, X_{(n)}$$

をこの標本の順序統計量といい,  $X_{(i)}$  を第  $i$  順序統計量とよぶ.

## 順序統計量 (2/5)

### ランダム標本の特性量

#### (1) 標本メディアン (中央値)

$$\tilde{X} = \begin{cases} X_{(k)} & n = 2k - 1 \\ (X_{(k)} + X_{(k+1)})/2 & n = 2k \end{cases}$$

#### (2) 標本範囲

$$R = X_{(n)} - X_{(1)}$$

#### (3) 標本中点

$$(X_{(1)} + X_{(n)})/2$$

#### (4) 100 $\alpha$ %調整平均

$$\frac{1}{n - 2k} \sum_{i=k+1}^{n-k} X_{(i)}, \quad k = [n\alpha]$$

( $[a]$  は  $a$  を超えない最大の整数, ガウス記号と呼ばれる)

## 順序統計量 (4/5)

### 定理 6.11 (順序統計量の分布)

$F(x)$  : 母集団分布の分布関数

$f(x)$  : 母集団分布の確率密度関数 (連続型の場合)

$$(1) \quad P(X_{(i)} \leq y) = \sum_{k=i}^n \frac{n!}{k!(n-k)!} \{F(y)\}^k \{1 - F(y)\}^{n-k}.$$

(2)  $g_i(y)$  :  $X_{(i)}$  の確率密度関数

$$g_i(y) = \frac{n!}{(i-1)!(n-i)!} \{F(y)\}^{i-1} \{1 - F(y)\}^{n-i} f(y).$$

## 順序統計量 (5/5)

### 定理 6.11 (順序統計量の分布 (つづき))

(3)  $i < j, y_i \leq y_j$  のとき

$$\begin{aligned}
 & P(X_{(i)} \leq y_i, X_{(j)} \leq y_j) \\
 &= \sum_{k=j}^n \sum_{l=i}^k \frac{n! \{F(y_i)\}^l \{F(y_j) - F(y_i)\}^{k-l} \{1 - F(y_j)\}^{n-k}}{l!(k-l)!(n-k)!}.
 \end{aligned}$$

(4)  $g_{ij}(y_i, y_j) : (X_{(i)}, X_{(j)})$  の同時確率密度関数  
 $y_i < y_j$  のとき

$$\begin{aligned}
 & g_{ij}(y_i, y_j) \\
 &= \frac{n! F(y_i)^{i-1} \{F(y_j) - F(y_i)\}^{j-i-1} \{1 - F(y_i)\}^{n-j} f(y_i) f(y_j)}{(i-1)!(j-i-1)!(n-j)!}
 \end{aligned}$$



## 点推定 (1/7)

$$\mathbf{X} = (X_1, \dots, X_n) \sim \mathbf{P} \in \{\mathbf{P}_\theta \mid \theta \in \Theta\}$$

$f(\mathbf{x}; \theta)$  :  $\mathbf{P}_\theta$  の確率密度関数 (確率関数)

パラメータ  $\gamma = \gamma(\theta)$  の点推定 :

$$\hat{\gamma} : \{\mathbf{x} = (x_1, \dots, x_n \mid f(\mathbf{x}; \theta)) > 0\} \rightarrow \gamma(\Theta)$$

$\hat{\gamma}(\mathbf{x})$  : 推定値,  $\hat{\gamma}(\mathbf{X})$  : 推定量

## 点推定 (3/7)

### 良さの基準

- (1) 偏り:  $\text{Bias}(\hat{\gamma}) = E(\hat{\gamma}) - \gamma.$
- (2) 分散:  $\text{Var}(\hat{\gamma}) = E\{[\hat{\gamma} - E(\hat{\gamma})]^2\}.$
- (3) 平均 2 乗誤差:  $\text{MSE}(\hat{\gamma}) = E\{(\hat{\gamma} - \gamma)^2\}.$
- (4) 集中確率:  $P(|\hat{\gamma} - \gamma| \leq a), \quad a > 0.$
- (5) 一貫性:  $n \rightarrow \infty$  のとき,  $\hat{\gamma} \xrightarrow{p} \gamma.$

平均 2 乗誤差, 分散, 偏りの 2 乗の関係 :

$$\text{MSE}(\hat{\gamma}) = E\{[\hat{\gamma} - E(\hat{\gamma}) + E(\hat{\gamma}) - \gamma]^2\} = \text{Var}(\hat{\gamma}) + \text{Bias}(\hat{\gamma})^2.$$

## 点推定 (4/7)

### 定義 7.1

$\hat{\gamma} = \hat{\gamma}(\mathbf{X})$  : パラメータ  $\gamma = \gamma(\theta)$  の推定量

- (1)  $\forall \theta \in \Theta$  に対して,  $E(\hat{\gamma}) = \gamma$  となるとき,  $\hat{\gamma}$  は**不偏推定量**であるという.
- (2)  $\gamma$  の不偏推定量が存在するとき,  $\gamma$  は**推定可能**であるという.
- (3)  $\gamma$  の不偏推定量  $\hat{\gamma}$  が, 任意の不偏推定量  $\tilde{\gamma}$  に対して

$$\forall \theta \in \Theta \text{ に対して, } \text{Var}_{\theta}(\hat{\gamma}) \leq \text{Var}_{\theta}(\tilde{\gamma})$$

を満たすとき,  $\hat{\gamma}$  は  $\gamma$  の**一様最小分散不偏推定量** (UMVUE: Uniformly Minimum Variance Unbiased Estimator) であるという.

## 点推定 (6/7)

### 例 7.1

$X_1, \dots, X_n \stackrel{i.i.d.}{\sim} X, E(X) = \mu, \text{Var}(X) = \sigma^2$

#### (1) $\mu$ の推定

$$\hat{\mu}_{\mathbf{c}} = c_1 X_1 + \dots + c_n X_n$$

$$E(\hat{\mu}_{\mathbf{c}}) = c_1 \mu + \dots + c_n \mu = (c_1 + \dots + c_n) \mu$$

$\hat{\mu}_{\mathbf{c}}$  が不偏であるための必要十分条件は  $c_1 + \dots + c_n = 1$

$$\text{Var}(\hat{\mu}_{\mathbf{c}}) = (c_1^2 + \dots + c_n^2) \sigma^2.$$

条件  $c_1 + \dots + c_n = 1$  のもとで  $c_1^2 + \dots + c_n^2$  が最小となるのは

$$c_1 = \dots = c_n = \frac{1}{n}$$

## 点推定 (7/7)

### 例 7.4 (続き)

#### (2) $\sigma^2$ の推定

標本分散  $S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$  は  $\sigma^2$  の不偏推定量

$$E \left\{ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right\} = \frac{n-1}{n} \sigma^2$$

$n^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$  は  $\sigma^2$  の不偏推定量ではない。

# 最小 2 乗法と単回帰モデル (1/7)

## 最小 2 乗法 モデル

$$y_i = g_i(\theta_1, \dots, \theta_p) + \varepsilon_i, \quad i = 1, \dots, n$$

$g_i$  : 既知関数

$\theta = (\theta_1, \dots, \theta_p)'$  : パラメータ

$\varepsilon_1, \dots, \varepsilon_n$  : 誤差, 互いに独立で, 平均 0, 分散  $\sigma^2$

$$\text{残差平方和 : } \sum_{i=1}^n \{y_i - g_i(\theta_1, \dots, \theta_p)\}^2$$

を最小にするような  $\theta = \hat{\theta}$  を求める方法を**最小 2 乗法**,  $\hat{\theta}$  を**最小 2 乗推定量**という

# 最小 2 乗法と単回帰モデル (2/7)

## 単回帰モデル

$(y_1, x_1), \dots, (y_n, x_n)$  : データ

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, \dots, n$$

を単回帰モデルと呼ぶ.

仮定

(1)  $(\alpha, \beta)$  : 未知パラメータ

(2)  $x_1, \dots, x_n$  : 既知定数

(3)  $\varepsilon_1, \dots, \varepsilon_n$  : 互いに独立な確率変数

(4)  $E(\varepsilon_i) = 0, \text{Var}(\varepsilon_i) = \sigma^2, i = 1, \dots, n$

$y_i$  を目的変数 (応答変数),  $x_i$  を説明変数,  $\beta$  を回帰係数 と呼ぶ

## 最小 2 乗法と単回帰モデル (3/7)

### 残差平方和

$$Q(\alpha, \beta) = \sum_{i=1}^n \{y_i - \alpha - \beta x_i\}^2$$

$(\alpha, \beta)$  の最小 2 乗推定量は

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \quad \hat{\beta} = \frac{s_{xy}}{s_x^2},$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$s_{xy} = \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}), \quad s_x^2 = s_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2,$$

(証明は板書で)



## 最小 2 乗法と単回帰モデル (4/7)

### 定理 7.1

$$(1) E(\hat{\alpha}) = \alpha, E(\hat{\beta}) = \beta$$

$$(2) \text{Var}(\hat{\alpha}) = \sigma^2 \left( \frac{1}{n} + \frac{\bar{x}^2}{s_{xx}} \right), \quad \text{Var}(\hat{\beta}) = \frac{\sigma^2}{s_{xx}}, \quad \text{Cov}(\hat{\alpha}, \hat{\beta}) = -\frac{\sigma^2 \bar{x}}{s_{xx}}$$

### 証明 (1)

$$E(\hat{\alpha}) = E(\bar{y}) - E(\hat{\beta})\bar{x}, \quad E(\hat{\beta}) = \frac{1}{s_x^2} E(s_{xy})$$

$$E(\bar{y}) = \frac{1}{n} \sum_{i=1}^n E(y_i) = \frac{1}{n} \sum_{i=1}^n (\alpha + \beta x_i) = \alpha + \beta \bar{x}$$

$$\begin{aligned} E(s_{xy}) &= \sum_{i=1}^n E(y_i - \bar{y})(x_i - \bar{x}) = \sum_{i=1}^n \{(\alpha + \beta x_i) - (\alpha + \beta \bar{x})\}(x_i - \bar{x}) \\ &= \beta \sum_{i=1}^n (x_i - \bar{x})^2 = \beta s_{xx} \end{aligned}$$

## 最小 2 乗法と単回帰モデル (5/7)

証明 (2)

$$\text{Var}(\bar{y}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(y_i) = \frac{1}{n} \sigma^2$$

$$\text{Cov}(\bar{y}, s_{xy}) = \frac{1}{n} \sum_{i=1}^n \text{Cov}(\bar{y}, y_i - \bar{y})(x_i - \bar{x}) = 0 \quad \left( \text{Cov}(y_i, \bar{y}) = \frac{1}{n} \sigma^2 \right)$$

$$\begin{aligned} \text{Var}(s_{xy}) &= \sum_{i=1}^n \text{Var}(y_i - \bar{y})(x_i - \bar{x})^2 \\ &\quad + \sum_{i \neq j} \text{Cov}(y_i - \bar{y}, y_j - \bar{y})(x_i - \bar{x})(x_j - \bar{x}) \\ &= \frac{n-1}{n} \sigma^2 s_{xx} + \sum_{i \neq j} \left( -\frac{\sigma^2}{n} \right) (x_i - \bar{x})(x_j - \bar{x}) = \sigma^2 s_{xx} \end{aligned}$$

$$\left( \sum_{i=1}^n \sum_{j: j \neq i} (x_i - \bar{x})(x_j - \bar{x}) \right) = \sum_{i=1}^n (x_i - \bar{x}) \{ -(x_i - \bar{x}) \}$$

## 最小 2 乗法と単回帰モデル (6/7)

$$\text{Var}(\hat{\beta}) = \frac{1}{s_{xx}^2} \text{Var}(s_{xy}) = \frac{\sigma^2}{s_{xx}}$$

$$\text{Var}(\hat{\alpha}) = \text{Var}(\bar{y}) + \bar{x}^2 \text{Var}(\hat{\beta}) = \sigma^2 \left( \frac{1}{n} + \frac{\bar{x}^2}{s_{xx}} \right)$$

$$\text{Cov}(\hat{\alpha}, \hat{\beta}) = -\bar{x} \text{Var}(\hat{\beta}) = -\frac{\sigma^2 \bar{x}}{s_{xx}}$$

**注**  $y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i$  と表すと,  $\beta_0 = \alpha + \beta\bar{x}$ ,  $\beta_1 = \beta$  となり,  $\beta_0, \beta_1$  の最小 2 乗推定量は

$$\hat{\beta}_0 = \bar{y}, \quad \hat{\beta}_1 = \hat{\beta}$$

## 最小 2 乗法と単回帰モデル (7/7)

### 定理 7.2 (ガウス–マルコフの定理)

単回帰モデル  $y_i = \alpha + \beta x_i + \varepsilon_i$  ( $i = 1, \dots, n$ )  
において

$$E(\varepsilon_i) = 0, \quad \text{Var}(\varepsilon_i) = \sigma^2, \quad \text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (i \neq j = 1, \dots, n)$$

とする. 定数  $l_1, l_2$  に対して

- (1)  $\hat{\theta} = l_1 \hat{\alpha} + l_2 \hat{\beta}$  は,  $\theta = \theta(\alpha, \beta) = l_1 \alpha + l_2 \beta$  の線形不偏推定量
- (2) 任意の  $\theta$  の線形不偏推定量  $\tilde{\theta}$  に対して

$$\text{Var}(\hat{\theta}) \leq \text{Var}(\tilde{\theta})$$

注. ここでの線形不偏推定量とは,  $y_1, \dots, y_n$  の一次結合  $a_1 y_1 + \dots + a_n y_n$  で,  $\theta$  の不偏推定量となるもののこと.

(証明は省略. テキスト 9 章を参照)

## 一様最小分散不偏推定量 (1/10)

定理 1 (クラメールーラオの不等式 (1次元パラメーターの場合))

$\mathbf{X} = (X_1, \dots, X_n) \sim P \in \{P_\theta \mid \theta \in \Theta\}$ ,  $\Theta \subset \mathbb{R}$

$f(\mathbf{x}; \theta)$ :  $P_\theta$  の確率密度関数 (確率関数)

(仮定) 各  $\theta_0 \in \Theta$  に対して  $\theta_0$  の近傍  $U$  と関数  $M(\mathbf{x})$  が存在して,  $\theta \in U$  ならば

$$\left| \frac{\partial f(\mathbf{x}; \theta)}{\partial \theta} \right| \leq M(\mathbf{x}), \text{ かつ } E_{\theta_0} \left[ \left\{ \frac{M(\mathbf{X})}{f(\mathbf{X}; \theta_0)} \right\}^2 \right] < \infty$$

$\theta$  に任意の不偏推定量  $\hat{\theta}$  に対して

$$\text{Var}(\hat{\theta}) \geq \frac{1}{J(\theta)}$$

## 一様最小分散不偏推定量 (2/10)

### 定理 1 (クラメールーラオの不等式 (つづき))

ただし,

$$E_{\theta}[g(\mathbf{X})] = \begin{cases} \iint g(\mathbf{x})f(\mathbf{x};\theta)dx_1 \cdots dx_n & \text{(連続型)} \\ \sum_{j=1}^{\infty} g(\mathbf{x}_j)f(\mathbf{x}_j;\theta) & \text{(離散型)} \end{cases},$$
$$J(\theta) = E_{\theta} \left[ \left\{ \frac{\partial}{\partial \theta} \log f(\mathbf{X}; \theta) \right\}^2 \right]$$

であり,  $J(\theta) > 0$  と仮定する.

- 定理 1 の  $J(\theta)$  を **フィッシャー情報量** と呼ぶ.
- $\text{Var}(\hat{\theta}) = \frac{1}{J(\theta)}$  となる不偏推定量が存在するとき, それを **有効推定量** と呼ぶ

## 一様最小分散不偏推定量 (3/10)

### 命題 1 (シュワルツの不等式)

可測関数  $g, h$  に対して

$$\{E[g(\mathbf{X})h(\mathbf{X})]\}^2 \leq E[\{g(\mathbf{X})\}^2]E[\{h(\mathbf{X})\}^2]$$

### (証明のヒント)

任意の実数  $t$  に対して

$$\begin{aligned} E[\{tg(\mathbf{X}) + h(\mathbf{X})\}^2] \\ = t^2E[\{g(\mathbf{X})\}^2] + 2tE[g(\mathbf{X})h(\mathbf{X})] + E[\{h(\mathbf{X})\}^2] \geq 0 \end{aligned}$$

であることから、2 次方程式の判別式を考える。

# 一様最小分散不偏推定量 (4/10)

## 補題 1

定理 1 の仮定の下,  $E_{\theta}[\{g(\mathbf{X})\}^2] < \infty$  ならば

$$\frac{d}{d\theta} E_{\theta}[g(\mathbf{X})] = E_{\theta}\left[g(\mathbf{X}) \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta}\right]$$

**証明** (連続型の場合)  $\iint \sim dx_1 \cdots dx_n = \int \sim d\mathbf{x}$  と書く.

$$\begin{aligned} \left|g(\mathbf{x}) \frac{\partial f(\mathbf{x}; \theta)}{\partial \theta}\right| &\leq |g(\mathbf{x})| \left\{\frac{M(\mathbf{x})}{f(\mathbf{x}; \theta_0)}\right\} f(\mathbf{x}; \theta_0), \\ \int |g(\mathbf{x})| \left\{\frac{M(\mathbf{x})}{f(\mathbf{x}; \theta_0)}\right\} f(\mathbf{x}; \theta_0) d\mathbf{x} &= E_{\theta_0} \left[ |g(\mathbf{X})| \left\{\frac{M(\mathbf{X})}{f(\mathbf{X}; \theta_0)}\right\} \right] \\ &\leq \left\{ E_{\theta}[\{g(\mathbf{X})\}^2] E_{\theta_0} \left[ \left\{\frac{M(\mathbf{X})}{f(\mathbf{X}; \theta_0)}\right\}^2 \right] \right\}^{1/2} < \infty \end{aligned}$$

(シュワルツの不等式)



# 一様最小分散不偏推定量 (5/10)

優収束定理より

$$\begin{aligned}\frac{d}{d\theta} \mathbb{E}_\theta[g(\mathbf{X})] &= \frac{d}{d\theta} \int g(\mathbf{x}) f(\mathbf{x}; \theta) d\mathbf{x} = \int g(\mathbf{x}) \frac{\partial f(\mathbf{x}; \theta)}{\partial \theta} d\mathbf{x} \\ &= \int g(\mathbf{x}) \frac{\partial \log f(\mathbf{x}; \theta)}{\partial \theta} f(\mathbf{x}; \theta) d\mathbf{x} = \mathbb{E}_\theta \left[ g(\mathbf{X}) \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta} \right]\end{aligned}$$

離散型の場合には、積分を級数で置き換えればよい。

## 一様最小分散不偏推定量 (6/10)

### 定理 1 の証明

補題 1 で  $g(\mathbf{x}) \equiv 1$  ととると

$$0 = \mathbf{E}_\theta \left[ \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta} \right] \quad (1)$$

$\hat{\theta} = \hat{\theta}(\mathbf{X})$  を  $\theta$  の不偏推定量として,  $g(\mathbf{x}) = \hat{\theta}(\mathbf{x})$  ととると

$$1 = \mathbf{E}_\theta \left[ \hat{\theta} \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta} \right] \quad (2)$$

(2) - (1)  $\times \theta$

$$1 = \mathbf{E}_\theta \left[ (\hat{\theta} - \theta) \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta} \right]$$

$$1^2 \leq \mathbf{E}_\theta [(\hat{\theta} - \theta)^2] \mathbf{E}_\theta \left[ \left\{ \frac{\partial \log f(\mathbf{X}; \theta)}{\partial \theta} \right\}^2 \right] = \text{Var}(\hat{\theta}) J(\theta)$$

## 一様最小分散不偏推定量 (7/10)

### 定理 7.3 (クラメル-ラオの不等式)

$\mathbf{X} = (X_1, \dots, X_n) \sim P \in \{P_{\boldsymbol{\theta}} \mid \boldsymbol{\theta} \in \Theta\}$ ,  $\Theta \subset \mathbb{R}^p$

$f(\mathbf{x}; \boldsymbol{\theta})$ :  $P_{\boldsymbol{\theta}}$  の確率密度関数 (確率関数)

$\gamma(\boldsymbol{\theta})$ : 微分可能な関数

(仮定) 各  $\boldsymbol{\theta}_0$  に対して,  $\boldsymbol{\theta}_0$  の近傍  $U$  と関数  $M(\mathbf{x})$  が存在して,  
 $\boldsymbol{\theta} \in U$  ならば

$$\left\| \frac{\partial f(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| \leq M(\mathbf{x}), \text{ かつ, } E_{\boldsymbol{\theta}_0} \left[ \left\{ \frac{M(\mathbf{X})}{f(\mathbf{X}, \boldsymbol{\theta}_0)} \right\}^2 \right] < \infty$$

$\gamma(\boldsymbol{\theta})$  の任意の不偏推定量  $\hat{\gamma}(\mathbf{X})$  に対して

$$\text{Var}\{\hat{\gamma}(\mathbf{X})\} \geq \dot{\gamma}(\boldsymbol{\theta})' J(\boldsymbol{\theta})^{-1} \dot{\gamma}(\boldsymbol{\theta})$$

が成り立つ.

## 一様最小分散不偏推定量 (8/10)

定理 7.3 (クラメールーラオの不等式 (つづき))

ただし,  $\dot{\gamma}(\boldsymbol{\theta}) = \partial\gamma(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ ,

$$J(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} \left[ \left\{ \frac{\partial}{\partial\boldsymbol{\theta}} \log f(\mathbf{X}, \boldsymbol{\theta}) \right\} \left\{ \frac{\partial}{\partial\boldsymbol{\theta}} \log f(\mathbf{X}, \boldsymbol{\theta}) \right\}' \right]$$

であり,  $J(\boldsymbol{\theta})$  は正則であると仮定する.

$J(\boldsymbol{\theta})$  を **フィッシャー情報量行列** と呼ぶ. (証明は省略, テキスト 7.3 節参照)

## 一様最小分散不偏推定量 (9/10)

## 例 7.5 (単回帰モデル)

$$y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i, \quad i = 1, \dots, n$$

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_1 - \bar{x} \\ \vdots & \vdots \\ 1 & x_n - \bar{x} \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

誤差に正規性を仮定すると,  $\mathbf{y}$  の同時確率密度関数は

$$f(\mathbf{y}; \boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y} - X\boldsymbol{\beta})' (\mathbf{y} - X\boldsymbol{\beta}) \right\}$$

## 一様最小分散不偏推定量 (10/10)

## 例 7.5 (単回帰モデル (つづき))

$\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma^2)'$  のフィッシャー情報量行列は

$$J(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{\sigma^2} X'X & O \\ O & \frac{n}{2\sigma^4} \end{pmatrix}$$

$\boldsymbol{\ell} = (\ell_1, \ell_2)'$  とすると,  $\gamma(\boldsymbol{\beta}) = \boldsymbol{\ell}'\boldsymbol{\beta}$  の任意の不偏推定量に対して

$$\text{Var}\{\hat{\gamma}(\mathbf{y})\} \geq \sigma^2 \boldsymbol{\ell}'(X'X)^{-1} \boldsymbol{\ell}$$

一方,  $\boldsymbol{\beta}$  の最小 2 乗推定量は,  $\hat{\boldsymbol{\beta}} = (X'X)^{-1} X' \mathbf{y}$  と表わされ,

$$\text{Var}\{\boldsymbol{\ell}'\hat{\boldsymbol{\beta}}\} = \sigma^2 \boldsymbol{\ell}'(X'X)^{-1} \boldsymbol{\ell}.$$

$\boldsymbol{\ell}'\hat{\boldsymbol{\beta}}$  は,  $\boldsymbol{\ell}'\boldsymbol{\beta}$  の有効推定量であり, したがって一様最小分散不偏推定量でもある。