確率・統計 B
標本分布・推定

若木宏文

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

2014.12.17

table of contents

順序統計量 (1/5)

定義 6.12

$X_1, \ldots, X_n$ 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 $X_{(1)} \leq \cdots \leq X_{(n)}$ とするとき

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

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．

例．$n = 4$，X_1, X_2, X_3, X_4 の実現値が

$X_1 = 3, X_2 = 4, X_3 = 1, X_4 = 5$

$\Rightarrow \quad X_{(1)} = 1, X_{(2)} = 3, X_{(3)} = 4, X_4 = 5$

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 (3/5)

命題 1 (最大値・最小値の分布)

$F(x)$: 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) $X_{(n)} = \max\{X_1, \ldots, X_n\}$ の分布関数と確率密度関数

$$G_n(y) = \mathsf{P}(X_{(n)} \le y) = \{F(y)\}^n,$$
$$g_n(y) = \frac{d}{dy} G_n(y) = nf(y)\{F(y)\}^{n-1}$$

(2) $X_{(1)} = \min\{X_1, \ldots, X_n\}$ の分布関数と確率密度関数

$$G_1(y) = \mathsf{P}(X_{(1)} \le y) = 1 - \{1 - F(y)\}^n,$$
$$g_1(y) = \frac{d}{dy} G_n(y) = nf(y)\{1 - F(y)\}^{n-1}$$

(証明は板書で)

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

(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

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

(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

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

7. 推定

点推定 (1/7)

$\boldsymbol{X} = (X_1, \ldots, X_n)$: 標本変量 (確率変数)
$\mathcal{P} = \{\mathsf{P}_\theta \,|\, \theta \in \Theta\}$: 統計モデル

$$\boldsymbol{X} \sim \mathsf{P} \in \mathcal{P}$$

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

$\mathcal{X} = \{\boldsymbol{x} = (x_1, \ldots, x_n \,|\, f(\boldsymbol{x};\theta)) > 0\}$

パラメータ $\gamma = \gamma(\theta)$ の**点推定** :
各 $\boldsymbol{x} \in \mathcal{X}$ に γ の推定値 $\hat{\gamma} = \hat{\gamma}(\boldsymbol{x})$ を対応させる方法

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

点推定 (2/7)

例 7.1

$X_1, \ldots, X_{10} \overset{i.i.d.}{\sim} B(1, \theta),\ \theta \in (0, 1)$

(1) $\hat{\theta}_1 = (X_1 + \cdots + X_{10})/10$

(2) $\hat{\theta}_2 = 1/2$

(3) $\hat{\theta}_3 = X_1$

例 7.2

繰り返し測定モデル $X_i = \mu + \varepsilon_i,\ (i = 1, \ldots, n)$

$\varepsilon_1, \ldots, \varepsilon_n \overset{i.i.d.}{\sim} \varepsilon$, $\mathsf{E}(\varepsilon) = 0,\ \mathsf{Var}(\varepsilon) = \sigma^2$

$\theta = (\mu, \sigma^2)$

(1) $\gamma(\theta) = \mu : \quad \hat{\mu} = \bar{X}_n = \dfrac{1}{n} \sum_{i=1}^{n} X_i$

(2) $\gamma(\theta) = \sigma^2 : \quad \hat{\sigma}^2 = S_n = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$

点推定 (3/7)

良さの基準

(1) 偏り: $\mathrm{Bias}(\gamma) = \mathsf{E}(\hat{\gamma}) - \gamma.$

(2) 分散: $\mathsf{Var}(\hat{\gamma}) = \mathsf{E}[\{\hat{\gamma} - \mathsf{E}(\hat{\gamma})\}^2].$

(3) 平均 2 乗誤差: $\mathrm{MSE}(\hat{\gamma}) = \mathsf{E}\{(\hat{\gamma} - \gamma)^2\}.$

(4) 集中確率: $\mathsf{P}(|\hat{\gamma} - \gamma| \leq a), \quad a > 0.$

(5) 一致性: $n \to \infty$ のとき, $\quad \hat{\gamma} \xrightarrow{p} \gamma.$

平均 2 乗誤差, 分散, 偏りの 2 乗の間の関係：

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

点推定 (4/7)

定義 7.1

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

(1) ${}^\forall \theta \in \Theta$ に対して, $\mathrm{E}(\hat{\gamma}) = \gamma$ となるとき, $\hat{\gamma}$ は不偏推定量であるという.

(2) γ の不偏推定量が存在するとき, γ は推定可能であるという.

(3) γ の不偏推定量 $\hat{\gamma}$ が, 任意の不偏推定量 $\tilde{\gamma}$ に対して

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

を満たすとき, $\hat{\gamma}$ は γ 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 (5/7)

例 7.3 (例 7.1 の続き)

(1) $\mathsf{E}(\hat{\theta}_1) = \theta$, $\mathsf{Var}(\hat{\theta}_1) = \theta(1-\theta)/10$.

(2) $\mathsf{E}(\hat{\theta}_2) = 1/2$, $\mathsf{Var}(\hat{\theta}_2) = 0$.

(3) $\mathsf{E}(\hat{\theta}_3) = \theta$, $\mathsf{Var}(\hat{\theta}_3) = \theta(1-\theta)$.

点推定 (6/7)

例 7.4

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} X$, $\mathsf{E}(X) = \mu$, $\mathsf{Var}(X) = \sigma^2$

(1) μ の推定

$\hat{\mu}_{\boldsymbol{c}} = c_1 X_1 + \cdots + c_n X_n$

$$\mathsf{E}(\hat{\mu}_{\boldsymbol{c}}) = c_1\mu + \cdots + c_n\mu = (c_1 + \cdots + c_n)\mu$$

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

$$\mathsf{Var}(\hat{\mu}_{\boldsymbol{c}}) = (c_1^2 + \cdots + c_n^2)\sigma^2.$$

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

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

点推定 (7/7)

例 7.4 (続き)

(2) σ^2 の推定

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

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

$n^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$ は σ^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
$\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$: パラメータ
$\varepsilon_1, \dots, \epsilon_n$: 誤差, 互いに独立で, 平均 0, 分散 σ^2

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

を最小にするような $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ を求める方法を最小 2 乗法, $\hat{\boldsymbol{\theta}}$ 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

$(y_1, x_1), \ldots, (y_n, x_n)$: データ

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

を単回帰モデルと呼ぶ.

仮定

(1) (α, β) : 未知パラメータ

(2) $x_1, \ldots, x_n$: 既知定数

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

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

y_i を目的変数 (応答変数), x_i を説明変数, β 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

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

(α, β) 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) $\mathrm{E}(\hat{\alpha}) = \alpha$, $\mathrm{E}(\hat{\beta}) = \beta$

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

証明 (1)

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

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

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

最小 2 乗法と単回帰モデル (5/7)

証明 (2)

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

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

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

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

最小 2 乗法と単回帰モデル (6/7)

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

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

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

注 $y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i$ と表すと, $\beta_0 = \alpha + \beta\bar{x}, \quad \beta_1 = \beta$ となり, β_0, β_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 (ガウスーマルコフの定理)

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

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

とする．定数 ℓ_1, ℓ_2 に対して

(1) $\hat{\theta} = \ell_1 \hat{\alpha} + \ell_2 \hat{\beta}$ は, $\theta = \theta(\alpha, \beta) = \ell_1 \alpha + \ell_2 \beta$ の線形不偏推定量

(2) 任意の θ の線形不偏推定量 $\tilde{\theta}$ に対して

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

注．ここでの線形不偏推定量とは, $y_1, \ldots, y_n$ の一次結合
$a_1 y_1 + \cdots + a_n y_n$ で, θ 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)