# セミパラメトリック法の漸近理論

白石高章
横浜市立大学大学院総合理学研究科数理科学

## 1　1標本モデルにおける漸近理論

### 1.1　モデルの設定

$(X_1, \cdots, X_n)$ を連続分布関数 $F(\frac{x-\mu}{\sigma})$ をもつ母集団からの大きさ $n$ の無作為標本とする. さらに, $F(x)$ の密度関数 $f(x) \equiv F'(x)$ は $f(-x) = f(x)$ を満たす 0 について対称な関数とし, 一般性を失うことなく $\int_{-\infty}^{\infty} x^2 dF(x) = 1$ と仮定する. すなわち $X_1, \cdots, X_n$ は互いに独立で各 $X_i$ は $\mu$ について対称な同一の連続分布関数 $F(\frac{x-\mu}{\sigma})$ をもつ. $\mu$ と $\sigma^2$ は, それぞれ $X_i$ の平均と分散であるが未知パラメータとする.

### 1.2　漸近線形性

$$W(\Delta, \omega) \equiv \sum_{i=1}^{n} \{\Psi(\frac{X_i - \mu - \Delta/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}) - \Psi(\frac{X_i - \mu}{\rho})\}/\sqrt{n} + d(\Psi)\Delta/\sigma + e(\Psi)\omega,$$

$$W'(\Delta, \omega) \equiv \sum_{i=1}^{n} \{\Psi(\frac{X_i - \mu - \Delta/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}) - \Psi(\frac{X_i - \mu}{\rho})\}/\sqrt{n} + d(\Psi)\Delta/\sigma,$$

ただし,

$$d(\Psi) \equiv -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho) f'(x) dx, \quad e(\Psi) \equiv -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho)\{1 + \frac{x f'(x)}{f(x)}\} f(x) dx$$

**定理 1.1**　正則条件の下で, $\forall C_1,\ C_2,\ \epsilon > 0$ に対して,

$$\lim_{n\to\infty} P\{\sup_{|\Delta|<C_1, |\omega|<C_2} |W(\Delta, \omega)| > \epsilon\} = 0. \qquad \Box$$

$\Psi(\cdot)$ が奇関数ならば, $e = 0$ より
**系 1.2**　$\Psi(\cdot)$ が奇関数ならば, 定理 1.1 の条件の下で, $\forall C_1,\ C_2,\ \epsilon > 0$ に対して,

$$\lim_{n\to\infty} P\{\sup_{|\Delta|<C_1, |\omega|<C_2} |W'(\Delta, \omega)| > \epsilon\} = 0. \qquad \Box$$

を得る.

**系 1.3**　$\Psi(\cdot)$ が奇関数ならば, 定理 1.1 の条件の下で, $\forall C_1,\ C_2,\ C_3,\ \epsilon > 0$ に対して,

$$\lim_{n\to\infty} P\{\sup_{|\Delta_1|<C_1,|\Delta_2|<C_2,|\omega|<C_3} |W'(\Delta_1+\Delta_2,\omega)| > \epsilon\} = 0. \qquad \square$$

検定統計量と推定量を求めるために使われる関数 $\psi(\cdot)$ を

$$\psi(x) \equiv \max(\min(x,b),-b) = \begin{cases} -b & (x < -b) \\ x & (-b \le x \le b) \\ b & (x > b) \end{cases}$$

で定義し, $b$ は, $F(x)$ が正規分布の $\varepsilon$ 近傍
$U_\varepsilon \equiv \{F(x) = (1-\varepsilon)\Phi(x) + \varepsilon H(x) : \Phi(x)$ は標準正規分布の分布関数, $H(x)$ は
$h(x) = H'(x)$ とするときすべての $x$ に対して $h(-x) = h(x)$ を満たすある分布関数 $\}$
の中にあるときの近似的に最良な手法を与えるように決められる. $b$ と $\varepsilon$ の関係は, $\varphi(x)$ と $\Phi(x)$ をそれぞれ N(0,1) の密度関数と分布関数として,

$$\frac{2\varphi(b)}{b} - 2\Phi(-b) = \frac{\varepsilon}{1-\varepsilon} \tag{1.1}$$

である.

## 1.3　M 検定

$Y_i \equiv X_i - \mu_0$ で定義したものとする. 尺度母数 $\sigma$ を
$\hat{\sigma}_n \equiv \frac{1}{0.6745}(|Y_1|,\cdots,|Y_n|$ の中央値) で推定する. $\hat{\sigma}_n$ は $X_1,\cdots,X_n$ が同一の正規分布 $N(\mu_0,\sigma^2)$ に従っている場合の $\sigma$ の一致推定量である. $\hat{\sigma}_n$ は異常値に関する頑健性をもっている.

$$T_M \equiv \sum_{i=1}^n \psi(\frac{Y_i}{\hat{\sigma}_n}) = \sum_{i=1}^n \mathrm{sign}(Y_i)\cdot\psi(\frac{|Y_i|}{\hat{\sigma}_n})$$

とおくと, 帰無仮説 $H_0 : \mu = \mu_0$ の下で $Y_i$ は分布関数 $F(\frac{x}{\sigma})$ をもつ.

$$Z_M \equiv \frac{T_M}{\sqrt{\sum_{i=1}^n \{\psi(\frac{|Y_i|}{\hat{\sigma}_n})\}^2}}$$

とおく.

**定理 1.4**　$n \to \infty$ として, $H_0$ の下で $Z_M \xrightarrow{\mathcal{L}} N(0,1)$, すなわち, $Z_M$ は標準正規分布に分布収束する.

**証明**　$\hat{\sigma}_n$ は $\frac{\sigma F^{-1}(0.75)}{\Phi^{-1}(0.75)}$ の一致推定量で 前節の $\rho$ を $\rho \equiv \frac{\sigma F^{-1}(0.75)}{\Phi^{-1}(0.75)}$ とおく. このとき, 系 1.2 より

$$\sum_{i=1}^n \psi(\frac{Y_i}{\hat{\sigma}_n})/\sqrt{n} \approx \sum_{i=1}^n \psi(\frac{Y_i}{\rho})/\sqrt{n} \tag{1.2}$$

ただし, $A_n \approx B_n$ は $A_n - B_n \xrightarrow{P} 0$ を意味する.

また, $\Psi=\psi^2$ で定理 1.1 を適用することにより,

$$\sum_{i=1}^{n}\{\psi(\frac{Y_i}{\hat{\sigma}_n})\}^2/n \approx \sum_{i=1}^{n}\{\psi(\frac{Y_i}{\rho})\}^2/n. \tag{1.3}$$

大数の法則, 中心極限定理, 著書の系 3.11.1 より, 結果を得る. □

## 1.4 M 推定量

$$T_M(\mu) \equiv \sum_{i=1}^{n}\psi(\frac{X_i-\mu}{\breve{\sigma}_n})=0$$

の解 $\breve{\mu}_n$ を $\mu$ の点推定量で M 推定量と呼ばれている, ただし,

$$\breve{\sigma}_n \equiv \frac{1}{\Phi^{-1}(0.75)}(|X_1-\mathrm{med}(X)|,\cdots,|X_n-\mathrm{med}(X)| \text{ の中央値}),$$

$$\mathrm{med}(X) \equiv (X_1,\cdots,X_n \text{ の中央値}).$$

$\breve{\sigma}_n$ は $\rho$ の一致推定量である.

系 1.2 より

$$0=\sum_{i=1}^{n}\psi(\frac{X_i-\breve{\mu}_n}{\breve{\sigma}_n})/\sqrt{n} \approx \sum_{i=1}^{n}\psi(\frac{X_i-\mu}{\rho})/\sqrt{n}-\sqrt{n}d(\psi)(\breve{\mu}_n-\mu)/\sigma.$$

$$\sqrt{n}(\breve{\mu}_n-\mu) \approx (\sigma/d(\psi))\sum_{i=1}^{n}\psi(\frac{X_i-\mu}{\rho})/\sqrt{n} \xrightarrow{\mathcal{L}} N(0,c(\psi,f)\sigma^2/d^2(\psi)) \tag{1.4}$$

ただし, $c(\psi,f)\equiv\int_{-\infty}^{\infty}\psi^2(\sigma x/\rho)f(x)dx$.

## 1.5 区間推定

$$\breve{\eta}_n \equiv \{T_M(\breve{\mu}_n-\Delta/\sqrt{n})-T_M(\breve{\mu}_n+\Delta/\sqrt{n})\}/(2\sqrt{n}\Delta)$$

とおくと, 系 1.3 より,

$$\breve{\eta}_n \xrightarrow{P} d(\psi)/\sigma \tag{1.5}$$

さらに,

$$\breve{c}_n(\psi,f) \equiv \frac{1}{n}\sum_{i=1}^{n}\psi^2(\frac{X_i-\breve{\mu}_n}{\breve{\sigma}_n})$$

とおけば, $\Psi=\psi^2$ で定理 1.1 を適用することにより,

$$\sum_{i=1}^{n}\psi^2(\frac{X_i-\breve{\mu}_n}{\breve{\sigma}_n})/\sqrt{n} \approx \sum_{i=1}^{n}\psi^2(\frac{X_i-\mu}{\rho})/\sqrt{n}-\sqrt{n}d(\psi^2)(\breve{\mu}_n-\mu)/\sigma-\sqrt{n}e(\psi^2)(\log\breve{\sigma}_n-\log\rho)$$

により,

$$\breve{c}_n(\psi, f) \approx \sum_{i=1}^{n} \psi^2(\frac{X_i - \mu}{\rho})/n \xrightarrow{P} c(\psi, f). \tag{1.6}$$

(1.4)-(1.6) より

**補題 1.5**

$$\frac{\sqrt{n}\breve{\eta}_n}{\sqrt{\breve{c}_n(\psi, f)}}(\breve{\mu}_n - \mu) \xrightarrow{\mathcal{L}} N(0, 1).$$

**定理 1.6**

$$(\breve{\mu}_n - \frac{\sqrt{\breve{c}_n(\psi, f)}z_{\alpha/2}}{\sqrt{n}\breve{\eta}_n},\ \breve{\mu}_n + \frac{\sqrt{\breve{c}_n(\psi, f)}z_{\alpha/2}}{\sqrt{n}\breve{\eta}_n})$$

は $1 - \alpha$ 漸近信頼区間である． □

# 2 2標本モデルにおける漸近理論

## 2.1 モデルの設定

$(X_1, \cdots, X_{n_1})$ を連続分布関数 $F(\frac{x-\mu_1}{\sigma})$ をもつ母集団からの大きさ $n_1$ の無作為標本, $(Y_1, \cdots, Y_{n_2})$ を連続分布関数 $F(\frac{x-\mu_2}{\sigma})$ をもつ母集団からの大きさ $n_2$ の無作為標本とする．すなわち, $X_1, \cdots, X_{n_1}, Y_1, \cdots, Y_{n_2}$ は互いに独立で, 各 $X_i$ は同一の分布関数 $F(\frac{x-\mu_1}{\sigma})$ をもち, 各 $Y_j$ は同一の分布関数 $F(\frac{x-\mu_2}{\sigma})$ をもつとする．さらに, 一般性を失うことなく $\int_{-\infty}^{\infty} x dF(x) = 0$, $\int_{-\infty}^{\infty} x^2 dF(x) = 1$ を仮定する．このとき,

$$E(X_i) = \mu_1, \quad E(Y_j) = \mu_2, \quad V(X_i) = V(Y_j) = \sigma^2$$

が成り立ち, $\mu_1$, $\mu_2$ はそれぞれ $X_i$ と $Y_j$ の平均で, $\sigma^2$ は共通の分散となる．これらは未知パラメータとする．

## 2.2 漸近線形性

$n \equiv n_1 + n_2$ とする．

$$\begin{aligned} W_1(\Delta_1 + \Delta_2, \omega) \quad &\equiv \quad \frac{\sqrt{n}}{n_1}\sum_{i=1}^{n_1}\{\Psi(\frac{X_i - \mu_1 - (\Delta_1 + \Delta_2)/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}) - \Psi(\frac{X_i - \mu_1}{\rho})\} \\ &\quad +d(\Psi)(\Delta_1 + \Delta_2)/\sigma + e(\Psi)\omega, \end{aligned}$$

$$\begin{aligned} W_2(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega) \quad &\equiv \quad \frac{\sqrt{n}}{n_2}\sum_{j=1}^{n_2}\{\Psi(\frac{Y_j - \mu_2 - (\Delta_1 - \frac{n_1}{n_2}\Delta_2)/\sqrt{n}}{\rho e^{\omega/\sqrt{n}}}) - \Psi(\frac{Y_j - \mu_2}{\rho})\} \\ &\quad +d(\Psi)(\Delta_1 - \frac{n_1}{n_2}\Delta_2)/\sigma + e(\Psi)\omega, \end{aligned}$$

$$W(\Delta_1, \Delta_2, \omega) \equiv W_1(\Delta_1 + \Delta_2, \omega) - W_2(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega)$$

ただし,

$$d(\Psi) \equiv -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho) f'(x) dx, \quad e(\Psi) \equiv -\int_{-\infty}^{\infty} \Psi(\sigma x/\rho)\{1 + \frac{x f'(x)}{f(x)}\} f(x) dx$$

$$0 < \lim_{n\to\infty} n_1/n = \lambda < 1$$

と仮定する.

**補題 2.1** 正則条件の下で, $\forall C_1,\ C_2,\ C_3,\ \epsilon > 0$ に対して,

$$\lim_{n\to\infty} P\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} |W_1(\Delta_1 + \Delta_2, \omega)| > \epsilon\} = 0,$$

$$\lim_{n\to\infty} P\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} |W_2(\Delta_1 - \frac{n_1}{n_2}\Delta_2, \omega)| > \epsilon\} = 0. \qquad \square$$

**定理 2.2** 正則条件の下で, $\forall C_1,\ C_2,\ C_3,\ \epsilon > 0$ に対して,

$$\lim_{n\to\infty} P\{\sup_{|\Delta_1|<C_1, |\Delta_2|<C_2, |\omega|<C_3} |W(\Delta_1, \Delta_2, \omega)| > \epsilon\} = 0. \qquad \square$$

### 2.3 M 検定

$$\tilde{\mu} = \frac{1}{n}(n_1\bar{X} + n_2\bar{Y})$$

とおき, 尺度母数 $\sigma$ を

$$\hat{\sigma}_n = \frac{\sqrt{\pi}}{\sqrt{2}\cdot n}\left(\sum_{i=1}^{n_1} |X_i - \tilde{\mu}| + \sum_{j=1}^{n_2} |Y_j - \tilde{\mu}|\right)$$

で推定する. $\hat{\sigma}_n$ は $X_1, \cdots, X_{n_1}, Y_1, \cdots, Y_{n_2}$ が同一の正規分布 $N(\mu, \sigma^2)$ に従っている場合の $\sigma$ の一致推定量である. 関数 $\psi(\cdot)$ を $\psi(x) = \max(\min(x, b), -b)$ で定義する. このとき

$$T_M \equiv \frac{1}{n_1}\sum_{i=1}^{n_1} \psi(\frac{X_i - \tilde{\mu}}{\hat{\sigma}_n}) - \frac{1}{n_2}\sum_{j=1}^{n_2} \psi(\frac{Y_j - \tilde{\mu}}{\hat{\sigma}_n})$$

で定義すれば, $H_0 : \mu_1 = \mu_2$ の下で $X_1, \cdots, X_{n_1}, Y_1, \cdots, Y_{n_2}$ は同一の連続分布に従う.

$$\bar{\psi}.(\boldsymbol{X}, \boldsymbol{Y}) \equiv \frac{1}{n}\{\sum_{i=1}^{n_1} \psi(\frac{X_i - \tilde{\mu}}{\hat{\sigma}_n}) + \sum_{j=1}^{n_2} \psi(\frac{Y_j - \tilde{\mu}}{\hat{\sigma}_n})\} \tag{2.1}$$

とおき,

$$Z_M \equiv \frac{\sqrt{n_1 n_2 (n-1)} T_M}{\sqrt{n[\sum_{i=1}^{n_1}\{\psi(\frac{X_i - \tilde{\mu}}{\hat{\sigma}_n}) - \bar{\psi}.(\boldsymbol{X}, \boldsymbol{Y})\}^2 + \sum_{j=1}^{n_2}\{\psi(\frac{Y_j - \tilde{\mu}}{\hat{\sigma}_n}) - \bar{\psi}.(\boldsymbol{X}, \boldsymbol{Y})\}^2]}}$$

とおく.

**定理 2.3**　$n \to \infty$ として, $H_0$ の下で $Z_M \xrightarrow{\mathcal{L}} N(0,1)$, すなわち, $Z_M$ は標準正規分布に分布収束する.

**証明**　$\mu_1 = \mu_2 \equiv \mu$, $\rho \equiv \frac{\sqrt{\pi}}{\sqrt{2}} \int_{-\infty}^{\infty} |x| dF(x)$ とおくと, 定理 2.2 より,

$$\sqrt{n} T_M \approx \sqrt{n} \{ \frac{1}{n_1} \sum_{i=1}^{n_1} \psi(\frac{X_i - \mu}{\rho}) - \frac{1}{n_2} \sum_{j=1}^{n_2} \psi(\frac{Y_j - \mu}{\rho}) \}. \tag{2.2}$$

補題 2.1 より,

$$\begin{aligned} & \frac{1}{n} [\sum_{i=1}^{n_1} \{ \psi(\frac{X_i - \tilde{\mu}}{\hat{\sigma}_n}) - \bar{\psi}.(\boldsymbol{X}, \boldsymbol{Y}) \}^2 + \sum_{j=1}^{n_2} \{ \psi(\frac{Y_j - \tilde{\mu}}{\hat{\sigma}_n}) - \bar{\psi}.(\boldsymbol{X}, \boldsymbol{Y}) \}^2 ] \\ \approx & \frac{1}{n} [\sum_{i=1}^{n_1} \{ \psi(\frac{X_i - \mu}{\rho}) - \bar{\psi}_{\cdot 0}(\boldsymbol{X}, \boldsymbol{Y}) \}^2 + \sum_{j=1}^{n_2} \{ \psi(\frac{Y_j - \mu}{\rho}) - \bar{\psi}_{\cdot 0}(\boldsymbol{X}, \boldsymbol{Y}) \}^2 ], \end{aligned} \tag{2.3}$$

ただし,

$$\bar{\psi}_{\cdot 0}(\boldsymbol{X}, \boldsymbol{Y}) \equiv \frac{1}{n} \{ \sum_{i=1}^{n_1} \psi(\frac{X_i - \mu}{\rho}) + \sum_{j=1}^{n_2} \psi(\frac{Y_j - \mu}{\rho}) \}.$$

定理 1.4 と同様の議論により証明が行える.　□

## 2.4　M 推定量

$$\bar{X} \equiv \frac{1}{n_1} \sum_{i=1}^{n_1} X_i, \ \bar{Y} \equiv \frac{1}{n_2} \sum_{j=1}^{n_2} Y_j,$$

$$\tilde{\mu} \equiv \frac{1}{n} (n_1 \bar{X} + n_2 \bar{Y}),$$

$$\breve{\sigma}_n \equiv \frac{\sqrt{\pi}}{\sqrt{2} \cdot n} (\sum_{i=1}^{n_1} |X_i - \bar{X}| + \sum_{j=1}^{n_2} |Y_i - \bar{Y}|)$$

とおき,

$$T_M(\theta) \equiv \frac{1}{n_1} \sum_{i=1}^{n_1} \psi(\frac{X_i - \tilde{\mu} - \theta}{\breve{\sigma}_n}) - \frac{1}{n_2} \sum_{j=1}^{n_2} \psi(\frac{Y_j - \tilde{\mu} + (\frac{n_1}{n_2}) \cdot \theta}{\breve{\sigma}_n})$$

とおく. $T_M(\theta) = 0$ の解を $\breve{\theta}_n$ とし, $\breve{\delta}_n = (1 + \frac{n_1}{n_2}) \cdot \breve{\theta}_n$ を $\delta \equiv \mu_1 - \mu_2$ の点推定量とする.

$\breve{\sigma}_n$ は $\rho \equiv \frac{\sqrt{\pi}}{\sqrt{2}} \int_{-\infty}^{\infty} |x| dF(x)$ の一致推定量である.

定理 2.2 より

$$0 = \sqrt{n} T_M(\breve{\theta}_n) \approx \frac{\sqrt{n}}{n_1} \sum_{i=1}^{n_1} \{ \psi(\frac{X_i - \mu_1}{\rho}) - \bar{\psi} \} - \frac{\sqrt{n}}{n_2} \sum_{j=1}^{n_2} \{ \psi(\frac{Y_j - \mu_2}{\rho}) - \bar{\psi} \} - \sqrt{n} d(\psi)(\breve{\delta}_n - \delta)/\sigma$$

ただし, $\bar{\psi} \equiv \int_{-\infty}^{\infty} \psi(\frac{\sigma x}{\rho}) dF(x)$ とする.

$$\begin{aligned} \sqrt{n}(\breve{\delta}_n - \delta) \quad & \approx \quad (\sigma / d(\psi)) [\frac{\sqrt{n}}{n_1} \sum_{i=1}^{n_1} \{ \psi(\frac{X_i - \mu_1}{\rho}) - \bar{\psi} \} - \frac{\sqrt{n}}{n_2} \sum_{j=1}^{n_2} \{ \psi(\frac{Y_j - \mu_2}{\rho}) - \bar{\psi} \} ] \\ & \xrightarrow{\mathcal{L}} \quad N(0, c'(\psi, f) \sigma^2 / \{ \lambda(1 - \lambda) d^2(\psi) \}) \end{aligned} \tag{2.4}$$

ただし, $c'(\psi,f)\equiv\int_{-\infty}^{\infty}\{\psi(\sigma x/\rho)-\bar{\psi}\}^2f(x)dx.$

## 2.5　区間推定

$$\breve{\eta}_n\equiv\sqrt{n}\{T_M(\breve{\theta}_n-\Delta/\sqrt{n})-T_M(\breve{\theta}_n+\Delta/\sqrt{n})\}/\{2(1+\frac{n_1}{n_2})\Delta\}$$

とおくと, 定理 2.2 を使って, (1.5) と同様に

$$\breve{\eta}_n\xrightarrow{P}d(\psi)/\sigma \tag{2.5}$$

さらに,

$$\breve{c}_n(\psi,f)\equiv\frac{1}{n}[\sum_{i=1}^{n_1}\{\psi(\frac{X_i-\bar{X}}{\breve{\sigma}_n})-\breve{\psi}.(X,Y)\}^2+\sum_{j=1}^{n_2}\{\psi(\frac{Y_j-\bar{Y}}{\breve{\sigma}_n})-\breve{\psi}.(X,Y)\}^2]$$

$$\breve{\psi}.(X,Y)\equiv\frac{1}{n}\{\sum_{i=1}^{n_1}\psi(\frac{X_i-\bar{X}}{\breve{\sigma}_n})+\sum_{j=1}^{n_2}\psi(\frac{Y_j-\bar{Y}}{\breve{\sigma}_n})\}$$

とおけば, (1.6) と同様に,

$$\breve{c}_n(\psi,f)\xrightarrow{P}c'(\psi,f). \tag{2.6}$$

(2.4)-(2.6) より

**補題 2.4**

$$\frac{\sqrt{n_1n_2}\breve{\eta}_n}{\sqrt{n\breve{c}_n(\psi,f)}}(\breve{\delta}_n-\delta)\xrightarrow{\mathcal{L}}N(0,1).\qquad\square$$

**定理 2.5**

$$(\breve{\delta}_n-\frac{\sqrt{n\breve{c}_n(\psi,f)}z_{\alpha/2}}{\sqrt{n_1n_2}\breve{\eta}_n},\ \breve{\delta}_n+\frac{\sqrt{n\breve{c}_n(\psi,f)}z_{\alpha/2}}{\sqrt{n_1n_2}\breve{\eta}_n})$$

は $1-\alpha$ 漸近信頼区間である.　　$\square$

著書と発展のページにリンク