アフィン変換群に関する分散の共変推定量

平成 20 年 7 月 10 日

決定問題

$$
\begin{aligned}
&\boldsymbol{X} = (X_1, \cdots, X_n), \quad X_1, \cdots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2) \\
&D = \{\sigma^2 | \sigma \in (0, \infty)\} \quad (\sigma^2 \text{ の推定問題 }) \\
&L((\mu, \sigma), d) = \Big(\frac{d - \sigma^2}{\sigma^2}\Big)^2
\end{aligned}
$$

変換群

$$G = \{g_{a,b} | g_{a,b}(\boldsymbol{x}) = (\tfrac{x_1 - a}{b}, \cdots, \tfrac{x_n - a}{b}), a \in \mathbf{R}, b > 0\}$$

とすると、決定問題は不変で、母数空間上の変換、決定空間上の変換は、それぞれ、

$$\bar{g}_{a,b} = \Big(\frac{\mu - a}{b}, \frac{\sigma^2}{b^2}\Big), \quad g^*_{a,b}(d) = \frac{d}{b^2}$$

共変推定量

δ を σ^2 の共変推定量とすると

$$\delta\Big(\frac{X_1 - a}{b}, \cdots, \frac{X_n - a}{b}\Big) = \frac{1}{b^2}\delta(X_1, \cdots, X_n)$$

$$a = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, \quad b = B_n = \sqrt{\sum_{i=1}^n (X_i - \bar{X})}$$

ととると

$$\delta(X_1, \cdots, X_n) = B_n^2 \delta\Big(\frac{X_1 - \bar{X}}{B_n}, \cdots, \frac{X_n - \bar{X}}{B_n}\Big)$$

逆に、任意の n 変数関数 $\gamma(y_1, \cdots, y_n)$ を用いて

$$\delta(X_1, \cdots, X_n) = B_n^2 \gamma\Big(\frac{X_1 - \bar{X}}{B_n}, \cdots, \frac{X_n - \bar{X}}{B_n}\Big)$$

と定義すると、δ は共変推定量となる。

共変推定量の分布について

$$
\begin{aligned}
&(X_1 - \bar{X}, \cdots, X_n - \bar{X})' = \Pi_n \boldsymbol{X}, \\
&\quad \Pi_n = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}', \\
&\quad \mathbf{1}_n = (1, 1, \cdots, 1)' : n \times 1 \text{ ベクトル} \\
&\quad I_n : n \times n \text{ 単位行列 },
\end{aligned}
$$

と表されるが、Π_n の固有値は、1 ($(n-1)$ 重根) と 0 (単根) なので、固有値 1 に属する一次独立な $n-1$ 個の長さ 1 の固有ベクトルを横に並べてできる $(n \times (n-1))$-行列を H とすると、

$$\Pi_n = HH', \quad H'H = I_{n-1}$$

が成り立つ。($(\mu, \sigma) = (0, 1)$) のとき、$\boldsymbol{X} \sim N_n(\mathbf{0}, I_n)$ となるので、

$$\boldsymbol{Y} = H'\boldsymbol{X} \sim N_{n-1}(\mathbf{0}, I_{n-1})$$

となり、$B_n = \|\boldsymbol{Y}\|$ となるので、

$$
\begin{aligned}
&B_n^2 \sim \chi^2_{n-1} \\
&\quad pdf : g(x) = \frac{1}{2^{n-1}\Gamma[\frac{n-1}{2}]} e^{-x/2} \Big(\frac{x}{2}\Big)^{\frac{n-1}{2} - 1} \quad (x > 0) \\
&\quad Y_1 = B_n \cos\Theta_1, \\
&\quad Y_2 = B_n \sin\Theta_1 \cos\Theta_2, \\
&\quad Y_3 = B_n \sin\Theta_1 \sin\Theta_2 \cos\Theta_3, \\
&\qquad\qquad \vdots \\
&\quad Y_{n-2} = B_n \Big(\prod_{i=1}^{n-3} \sin\Theta_i\Big) \cos\Theta_{n-1}, \\
&\quad Y_{n-1} = B_n \Big(\prod_{i=1}^{n-3} \sin\Theta_i\Big) \sin\Theta_{n-1}, \\
&\quad 0 \le \Theta_i < \pi \ (i = 1, \cdots, n-3); 0 \le \Theta_{n-2} < 2\pi
\end{aligned}
$$

と曲座標変換すると、B_n と $(\Theta_1, \cdots, \Theta_{n-2})$ が独立であることがわかる。$\frac{X_i - \bar{X}}{B_n}$ $(i = 1, \cdots, n)$ は、$\Theta_1, \cdots, \Theta_{n-2}$ の関数なので、

$$B_n^2 \text{ と } C := \delta\Big(\frac{X_1 - \bar{X}}{B_n}, \cdots, \frac{X_n - \bar{X}}{B_n}\Big) \text{ は独立}$$

最小リスク共変推定量

リスク（期待損失）の不変性から

$$
\begin{aligned}
\mathrm{E}_{(\mu,\sigma)}[L((\mu, \sigma), B_n^2 C)] &= \mathrm{E}_{(0,1)}[L((0, 1), B_n^2 C)] \\
&= \mathrm{E}[(B_n^2 C - 1)^2] \\
&= \mathrm{E}[\mathrm{E}[B_n^4 C^2 - \mathrm{E}[B_n^2]]C + 1] \\
&= \mathrm{E}[(n-1)(n+1)C^2 - 2(n-1)C + 1]
\end{aligned}
$$

$$C \equiv \frac{1}{n+1}$$

ととると、期待損失は最小となる。