

研究ノート

2 標本正規分布の平均に順序があるときの 最尤推定量の期待値と ブートストラップ推定について

安 楽 和 夫

Expectation of the Order Restricted MLE of the Means
in a Two-Sample Normal Model and the Bootstrap Estimation

Kazuo Anraku

1 はじめに

2つの正規分布母集団の平均母数を θ_1, θ_2 とする。 θ_1, θ_2 は推定すべき未知母数である。それぞれの分布の分散はともに既知とする。2つの母集団から得られたデータから、それぞれの標本平均が求まるが、これらは θ_1, θ_2 についての尤度関数を最大にする最尤推定量 (maximum likelihood estimator; MLE) である。その期待値はそれぞれの平均母数 θ_1, θ_2 になる。この場合、 θ_1, θ_2 には制約条件は何も仮定されていないが、本稿では順序制約

$$\theta_1 \leq \theta_2 \quad (1)$$

を仮定し、平均母数 θ_1, θ_2 の順序制約下での最尤推定量 (order restricted MLE; ORMLE) ならびに関連する統計量の期待値について考察する。

2つの母集団の母数比較は2標本問題などと呼ばれるが、これは一般に、複数の母集団分布を規定する母数についての推測を行う多標本問題 (あるいは k 標本問題) の特殊な場合である。多標本でのORMLEについては、検定統計量や許容性との関係などから多くの研究があるが、推定量の期待値そのものに

ついて記述した文献は見当たらないようである。本稿では、2標本の場合の ORMLE の期待値の明示的な式を与える。またブートストラップ法による推定について考察する。

2 順序制約下での最尤推定量の期待値とブートストラップ法

2.1 順序制約下での最尤推定量の期待値

平均 0, 分散 1 の正規分布を標準正規分布という。標準正規分布の密度関数を $\phi(x)$ とすると、 $\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ である。また分布関数を $\Phi(x)$ と表す。すなわち $\Phi(x) = \int_{-\infty}^x \phi(t) dt$ である。平均 μ , 分散 τ^2 ($\tau > 0$) の正規分布の密度関数は、 $\phi((x-\mu)/\tau)/\tau$ で表され、その分布関数は $\Phi((x-\mu)/\tau)$ となる。

ここでは 2 つの 1 次元正規母集団を考える。 $X_1, \dots, X_m; Y_1, \dots, Y_n$ はそれぞれ独立に $N(\theta_1, \lambda_1^2)$, $N(\theta_2, \lambda_2^2)$ に従うものとする (以下このような場合には、 $X_1, \dots, X_m \sim N(\theta_1, \lambda_1^2)$ と表記する)。ここで、分散 λ_1^2, λ_2^2 ($\lambda_1 > 0, \lambda_2 > 0$) はそれぞれ既知である。対数尤度関数を $l(\theta_1, \theta_2)$ とすると、 $l(\theta_1, \theta_2)$ は次のように表せる:

$$l(\theta_1, \theta_2) = c_1 - \frac{1}{2} \sum_{i=1}^m \frac{(x_i - \theta_1)^2}{\lambda_1^2} - \frac{1}{2} \sum_{j=1}^n \frac{(y_j - \theta_2)^2}{\lambda_2^2} \quad (2)$$

$$= c_2 - \frac{1}{2} \left\{ \frac{(\theta_1 - \bar{x})^2}{\sigma_1^2} + \frac{(\theta_2 - \bar{y})^2}{\sigma_2^2} \right\}. \quad (3)$$

ここで、 $\bar{x} = m^{-1} \sum_{i=1}^m x_i$, $\bar{y} = n^{-1} \sum_{j=1}^n y_j$, $\sigma_1^2 = \lambda_1^2/m$, $\sigma_2^2 = \lambda_2^2/n$,

$$c_1 = -\frac{m+n}{2} \log(2\pi) - \frac{m}{2} \log \lambda_1^2 - \frac{n}{2} \log \lambda_2^2, \quad (4)$$

$$c_2 = c_1 - \frac{1}{2} \sum_{i=1}^m \frac{(x_i - \bar{x})^2}{\lambda_1^2} - \frac{1}{2} \sum_{j=1}^n \frac{(y_j - \bar{y})^2}{\lambda_2^2},$$

である。なお、変数をデータと見なす場合は小文字で、確率変数と見なすときは大文字で表記するのが一般的である。本稿では両方の記法を適宜用いるが、同一視してもよい。

順序制約が無い場合の θ_1, θ_2 の最尤推定量 (MLE) を $\hat{\theta}_1, \hat{\theta}_2$ で表すと、上の (3) 式から $\hat{\theta}_1 = \bar{x}$, $\hat{\theta}_2 = \bar{y}$ となる。一方、制約 (1) を仮定した場合の、 θ_1, θ_2 の

順序制約下での最尤推定量 (ORMLE) を $\tilde{\theta}_1, \tilde{\theta}_2$ で表すと, $\hat{\theta}_1 (= \bar{x}) \leq \hat{\theta}_2 (= \bar{y})$ のとき, $\tilde{\theta}_i = \hat{\theta}_i (i = 1, 2)$ であるが, $\hat{\theta}_1 (= \bar{x}) \geq \hat{\theta}_2 (= \bar{y})$ のとき, MLE $\tilde{\theta}_1, \tilde{\theta}_2$ は, θ_1 と θ_2 の座標平面で領域の境界 $\theta_1 = \theta_2 = \theta$ 上に同一値として得られる. したがって $\tilde{\theta}_1 (= \tilde{\theta}_2)$ は $dl(\theta, \theta)/d\theta = 0$ より次で与えられる:

$$\tilde{\theta}_i = \begin{cases} \hat{\theta}_i & \text{if } \hat{\theta}_1 \leq \hat{\theta}_2 \\ \frac{\sigma_2^2 \hat{\theta}_1 + \sigma_1^2 \hat{\theta}_2}{\sigma_1^2 + \sigma_2^2} & \text{if } \hat{\theta}_1 > \hat{\theta}_2 \end{cases} \quad (i = 1, 2). \quad (5)$$

これから $\tilde{\theta}_1$ の期待値は次により求められる:

$$E[\tilde{\theta}_1] = \iint_{x \leq y} x \frac{1}{\sigma_1} \phi\left(\frac{x - \theta_1}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{y - \theta_2}{\sigma_2}\right) dx dy \\ + \iint_{x > y} \frac{\sigma_2^2 x + \sigma_1^2 y}{\sigma_1^2 + \sigma_2^2} \frac{1}{\sigma_1} \phi\left(\frac{x - \theta_1}{\sigma_1}\right) \frac{1}{\sigma_2} \phi\left(\frac{y - \theta_2}{\sigma_2}\right) dx dy$$

ここで直交行列を用いた変換

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sigma_1 / \sqrt{\sigma_1^2 + \sigma_2^2} & -\sigma_2 / \sqrt{\sigma_1^2 + \sigma_2^2} \\ \sigma_2 / \sqrt{\sigma_1^2 + \sigma_2^2} & \sigma_1 / \sqrt{\sigma_1^2 + \sigma_2^2} \end{pmatrix} \begin{pmatrix} (x - \theta_1) / \sigma_1 \\ (y - \theta_2) / \sigma_2 \end{pmatrix}$$

を行うと, 積分領域は次の領域に変換される:

$$\{(x, y) | x \leq y\} \Rightarrow \{(u, v) | u \leq (\theta_2 - \theta_1) / \sqrt{\sigma_1^2 + \sigma_2^2}\}$$

ここで

$$a = (\theta_2 - \theta_1) / \sqrt{\sigma_1^2 + \sigma_2^2} \quad (6)$$

とおくと, 上の変換により

$$E[\tilde{\theta}_1] = \iint_{u \leq a} \left(\frac{\sigma_1^2 u + \sigma_1 \sigma_2 v}{\sqrt{\sigma_1^2 + \sigma_2^2}} + \theta_1 \right) \phi(u) \phi(v) du dv \\ + \iint_{u > a} \left(\frac{\sigma_1 \sigma_2 v}{\sqrt{\sigma_1^2 + \sigma_2^2}} + \frac{\sigma_2^2 \theta_1 + \sigma_1^2 \theta_2}{\sigma_1^2 + \sigma_2^2} \right) \phi(u) \phi(v) du dv$$

となる. ここで関係式 $\phi'(u) = -u\phi(u)$, $\int v\phi(v) dv = 0$, $\Phi(-x) = 1 - \Phi(x)$ を用いると $\tilde{\theta}_1$ の期待値は次のように表される:

$$E[\tilde{\theta}_1] = \theta_1 - \frac{\sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \{\phi(a) - a\Phi(-a)\} \quad (7)$$

同様に, $\bar{\theta}_2$ の期待値も次のように表される:

$$E[\bar{\theta}_2] = \theta_2 + \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \{ \phi(a) - a\Phi(-a) \} \quad (8)$$

特に, $\theta_1 = \theta_2 = \theta$ のとき,

$$E[\bar{\theta}_1] = \theta - \frac{\sigma_1^2}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}}, \quad E[\bar{\theta}_2] = \theta + \frac{\sigma_2^2}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \quad (9)$$

を得る. ここで $m = n$ とすると, $\theta_1 = \theta_2 = \theta$ のとき次が得られる:

$$E[\bar{\theta}_1 - \theta] = -\frac{\lambda_1^2}{\sqrt{2\pi n(\lambda_1^2 + \lambda_2^2)}}$$

一般に, 二つの数列 $\{a_n\}$, $\{b_n\}$ について, $|a_n/b_n|$ が有界であるとき $a_n = O(b_n)$ と表す. 上の結果は, $\theta_1 = \theta_2 = \theta$ のとき $E[\bar{\theta}_1 - \theta] = O(1/\sqrt{n})$ を意味している.

これは Chernoff(1954)の結果と符合する.

なお, $h(a) = \phi(a) - a\Phi(-a)$ とおくと,

$$h'(a) = -a\phi(a) - \Phi(-a) + a\phi(a) = -\Phi(-a) < 0.$$

したがって, $h(a)$ は a の単調減少関数であり, $a \geq 0$ より, $h(a)$ は $a = 0$ で最大となり, $h(0) = 1/\sqrt{2\pi}$ である.

3 ブートストラップ法による推定

式(7)や(8)からわかるように, $E[\bar{\theta}_1]$ や $E[\bar{\theta}_2]$ は θ_1 , θ_2 に依存するので, これらの値は未知である. そこでこれを推定するためにブートストラップ法の利用を考える. ブートストラップ法は Efron(1979) によりその名前とともに提唱された方法であり, 広範に利用される統計手法である. ブートストラップ法には, 分布形を想定しないノンパラメトリック・ブートストラップ法と具体的な分布族として仮定するパラメトリック・ブートストラップ法とがある. 本稿では後者を採用する.

パラメトリック・ブートストラップ法は, 仮定した分布の未知母数の値にデータから得られた推定値を代入することで得られる分布を利用する. この分布から, 疑似乱数によりデータを人工的に生成し, これを利用して推定を行う

方法である。ここでは、 $x_1, x_2, \dots, x_m \sim N(\theta_1, \lambda_1^2)$ とする。未知母数である θ_1 を、 $\hat{\theta}_1 = \bar{x} = m^{-1} \sum_{i=1}^m x_i$ で置き換え、分布 $N(\hat{\theta}_1, \lambda_1^2)$ から正規乱数を生成する。 θ_2 についても同じである。

一般に、通常のブートストラップ法では元のデータと同数の n 個を生成し、これを 1 回分とする。生成されたデータ（あるいは確率変数）を x^* （あるいは X^* ）等で表すこととし、第 i 回目に生成されたデータを $x_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{in}^*)$ とする。またある指標の推定関数 $h(x)$ を用いて $h(x_i^*)$ により推定するとする。ブートストラップ法ではこれを B 回繰り返し、その平均

$$\frac{1}{B} \sum_{i=1}^B h(x_i^*)$$

を推定値とする。この B を大きな値に設定することで、よりよい推定値が得られるとされる。具体的な B の値を設定しなければならないが一般には $B \geq 1000$ が望ましいだろう。

パラメトリック・ブートストラップ法による推定量の期待値について考える。生成される確率変数を

$$X_1^*, X_2^*, \dots, X_{m_B}^* \left(\sim N(\hat{\theta}_1, \lambda_1^2) \right); Y_1^*, Y_2^*, \dots, Y_{n_B}^* \left(\sim N(\hat{\theta}_2, \lambda_2^2) \right)$$

とする。ブートストラップ法では母数とデータの間を模して

$$\begin{aligned} (\theta_1, \theta_2) &\Rightarrow (\hat{\theta}_1, \hat{\theta}_2) \\ (x_1, x_2, \dots, x_m) &\Rightarrow (x_1^*, x_2^*, \dots, x_m^*) \\ (y_1, y_2, \dots, y_n) &\Rightarrow (y_1^*, y_2^*, \dots, y_n^*) \\ (\hat{\theta}_1, \hat{\theta}_2) &\Rightarrow (\hat{\theta}_1^*, \hat{\theta}_2^*) \\ (\bar{\theta}_1, \bar{\theta}_2) &\Rightarrow (\bar{\theta}_1^*, \bar{\theta}_2^*) \end{aligned}$$

のように置き換えを行う。ただし本稿では、標本サイズをより一般的に扱うため、上のブートストラップ法の標本サイズ m, n を m_B, n_B とする。また $\hat{\theta}_1^*, \hat{\theta}_2^*, \bar{\theta}_1^*, \bar{\theta}_2^*$ は $\hat{\theta}_1, \hat{\theta}_2, \bar{\theta}_1, \bar{\theta}_2$ と同じ手順により (5) から構成されるものである。このような $\bar{\theta}_1^*, \bar{\theta}_2^*$ の B 回の平均をとったものがブートストラップ推定値となる。さてここでは正規分布を想定してデータを生成するので、

$$\hat{\theta}_1^*(=\bar{X}^*) \sim N(\hat{\theta}_1, \lambda_1^2/m_B), \quad \hat{\theta}_2^*(=\bar{Y}^*) \sim N(\hat{\theta}_2, \lambda_2^2/n_B)$$

と見なせる. ここで $\sigma_{B1}^2 = \lambda_1^2/m_B$, $\sigma_{B2}^2 = \lambda_2^2/n_B$ とおく. ブートストラップ法で求めた推定値の期待値について, 先の(7), (8)と同じ関係が成り立つ. つまり

$$E[\hat{\theta}_1^* | \hat{\theta}_1, \hat{\theta}_2] = \hat{\theta}_1 - \frac{\sigma_{B1}^2}{\sqrt{\sigma_{B1}^2 + \sigma_{B2}^2}} \{\phi(\hat{\alpha}) - \hat{\alpha}\Phi(-\hat{\alpha})\},$$

$$E[\hat{\theta}_2^* | \hat{\theta}_1, \hat{\theta}_2] = \hat{\theta}_2 + \frac{\sigma_{B2}^2}{\sqrt{\sigma_{B1}^2 + \sigma_{B2}^2}} \{\phi(\hat{\alpha}) - \hat{\alpha}\Phi(-\hat{\alpha})\}$$

ここで,

$$\hat{\alpha} = (\hat{\theta}_2 - \hat{\theta}_1) / \sqrt{\sigma_{B1}^2 + \sigma_{B2}^2} \quad (10)$$

である.

ブートストラップ法は本来, 統計量の分布の導出が困難な場合などに適用されるものであり, 上のように求めたい指標 (ここでは $\hat{\theta}_1$, $\hat{\theta}_2$ の期待値) が式として陽に表される場合に適用されるものではない. ここでの目的はブートストラップ法の特性を評価することである点を注意しておく.

さてブートストラップ推定量の期待値を求めてみよう. (10)の $\hat{\alpha}$ を確率変数とみなすと, $\hat{\theta}_1 \sim N(\theta_1, \lambda_1^2)$, $\hat{\theta}_2 \sim N(\theta_2, \lambda_2^2)$ より

$$\hat{\alpha} \sim N(\mu, \gamma^2)$$

である. ここで

$$\mu = (\theta_2 - \theta_1) / \sqrt{\sigma_{B1}^2 + \sigma_{B2}^2}, \quad \gamma = \sqrt{(\sigma_1^2 + \sigma_2^2) / (\sigma_{B1}^2 + \sigma_{B2}^2)}$$

である. 今 $\hat{\alpha} = \mu + \gamma Z$ ($Z \sim N(0, 1)$) より, $E[\hat{\theta}_2^* | \hat{\theta}_1, \hat{\theta}_2]$ を $\hat{\theta}_1$, $\hat{\theta}_2$ について (よって Z について) 期待値をとると,

$$EE[\hat{\theta}_1^* | \hat{\theta}_1, \hat{\theta}_2] = \theta_1 - \frac{\sigma_{B1}^2}{\sqrt{\sigma_{B1}^2 + \sigma_{B2}^2}} E[\phi(\mu + \gamma Z) - (\mu + \gamma Z)\{1 - \Phi(\mu + \gamma Z)\}]$$

である. 以下では $EE[\hat{\theta}_1^* | \hat{\theta}_1, \hat{\theta}_2]$ などを単に $E[\hat{\theta}_1^*]$ などと表すことにする.

関係式(A.1), (A.2), (A.3)を用いると, 次を得る:

$$\begin{aligned} & E[\phi(\mu + \gamma Z) - (\mu + \gamma Z)\{1 - \Phi(\mu + \gamma Z)\}] \\ &= \sqrt{1 + \gamma^2} \phi\left(\frac{\mu}{\sqrt{1 + \gamma^2}}\right) - \mu \left\{1 - \Phi\left(\frac{\mu}{\sqrt{1 + \gamma^2}}\right)\right\} \end{aligned}$$

ここで

$$1 + \gamma^2 = \frac{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}{\sigma_{B1}^2 + \sigma_{B2}^2}, \quad \frac{\mu}{\sqrt{1 + \gamma^2}} = \frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}$$

である. したがって次式を得る:

$$\begin{aligned} E[\tilde{\theta}_1^*] &= \theta_1 - \frac{\sigma_{B1}^2 \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}{\sigma_{B1}^2 + \sigma_{B2}^2} \left\{ \phi\left(\frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}\right) \right. \\ &\quad \left. - \frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}} \Phi\left(\frac{\theta_1 - \theta_2}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}\right) \right\}. \end{aligned} \quad (11)$$

同じく

$$\begin{aligned} E[\tilde{\theta}_2^*] &= \theta_2 + \frac{\sigma_{B2}^2 \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}{\sigma_{B1}^2 + \sigma_{B2}^2} \left\{ \phi\left(\frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}\right) \right. \\ &\quad \left. - \frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}} \Phi\left(\frac{\theta_1 - \theta_2}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}\right) \right\}. \end{aligned} \quad (12)$$

通常のブートストラップ法では, $m_B = m$, $n_B = n$ とするものとされる. これを仮定すると, $\sigma_{B1}^2 = \sigma_1^2$, $\sigma_{B2}^2 = \sigma_2^2$ (したがって, $\gamma^2 = 1$) となるので,

$$E[\tilde{\theta}_1^*] = \theta_1 - \frac{\sqrt{2}\sigma_1^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \left\{ \phi\left(\frac{\theta_2 - \theta_1}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}}\right) - \frac{\theta_2 - \theta_1}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}} \Phi\left(\frac{\theta_1 - \theta_2}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}}\right) \right\}$$

である. 期待値 $E[\tilde{\theta}_2^*]$ についても同様である (式は省略する). 特に, $\theta_1 = \theta_2 = \theta$ のとき次が成り立つ:

$$E[\tilde{\theta}_1^*] = \theta - \frac{\sigma_1^2}{\sqrt{\pi(\sigma_1^2 + \sigma_2^2)}}, \quad E[\tilde{\theta}_2^*] = \theta + \frac{\sigma_2^2}{\sqrt{\pi(\sigma_1^2 + \sigma_2^2)}},$$

これを(9)と比較すると, ブートストラップ法での推定量の期待値は, $\tilde{\theta}_1$, $\tilde{\theta}_2$ よりも $\sqrt{2}$ 倍大きく偏ることがわかる. このように $\theta_1 = \theta_2 = \theta$ のとき次の関係

が成り立つ:

$$E[\tilde{\theta}_1^*] < E[\tilde{\theta}_1] < \theta < E[\tilde{\theta}_2] < E[\tilde{\theta}_2^*]$$

上の(11), (12)式で, $m_B/n_B \rightarrow m/n$ となるように m_B, n_B を大きな値にとると, $\sigma_{B1}^2 \rightarrow 0, \sigma_{B2}^2 \rightarrow 0$ となり,

$$E[\tilde{\theta}_1^*] \rightarrow E[\tilde{\theta}_1], \quad E[\tilde{\theta}_2^*] \rightarrow E[\tilde{\theta}_2]$$

が成り立つ.

ところが, \hat{a} の分布を見ると, m_B, n_B を大きくとったとき, 期待値も分散も大きく変動することがわかる. K を適当な大きさの正数とすると,

$$\begin{aligned} P(\hat{a} \geq K) &= \int_K^\infty \frac{1}{\gamma} \phi((x-\mu)/\gamma) / \gamma dx = \Phi((\mu-K)/\gamma) \\ &= \Phi\left(\frac{(\theta_2 - \theta_1) - K\sqrt{\sigma_{B1}^2 + \sigma_{B2}^2}}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \end{aligned}$$

となるが, これは $\sigma_{B1}^2 \rightarrow 0, \sigma_{B2}^2 \rightarrow 0$ のとき $\Phi((\theta_2 - \theta_1)/\sqrt{\sigma_1^2 + \sigma_2^2})$ に近づく. つまり, K の値に依存しない値に近づき, これは $\theta_2 - \theta_1$ に依存した二項確率である. 適当な K の値に対して, 上の確率は 0 に近づくべきところが, そうなっていない. これは, \hat{a} について, 推定量としての一致性が成り立たないことを意味する.

4 情報量規準への適用について

Akaike (1973, 1974) による赤池情報量規準 (AIC) は, 推定された確率分布 $\hat{g}(x)$ の, 真の確率分布 $g(x)$ からの乖離度を Kullback-Leibler の情報量

$$I(g, \hat{g}) = \int g(x) \log g(x) dx - \int g(x) \log \hat{g}(x) dx \quad (13)$$

で測ろうとする手法である. マイナス符号を含めた第 2 項の値が小さいほど $\hat{g}(x)$ が真の確率分布 $g(x)$ により近いことが示唆される. 第 2 項の $\log \hat{g}(x)$ は想定した分布密度の対数尤度であり, 本稿では対数尤度関数 $l(\theta_1, \theta_2)$ の未知母数 θ_1, θ_2 に最尤推定量を代入したものである. (13)式第 2 項では対数尤

度関数の θ_1, θ_2 に代入した MLE (あるいは ORMLE) を定数とみなし, 他の変数について期待値を取るが, 一般にモデルの未知母数に依存する. そこでこれを再び最大対数尤度で推定を行うが, この推定は不偏でなく偏り (バイアス) を生じる. 通常の AIC ではこの偏りを補正するために未知母数の個数を用いるのであるが, これは用いられる MLE の漸近分布によるものである (坂元他 (1983) 等を参照されたい).

ここでの最大対数尤度は, 順序制約を前提とした ORMLE を代入した $l(\tilde{\theta}_1, \tilde{\theta}_2)$ である. 本稿の設定ではマイナスを除いた第 2 項は次のようになる:

$$E[l([\tilde{\theta}_1], [\tilde{\theta}_2])] = c_1 - E\left[\frac{1}{2}\sum_{i=1}^m \frac{(X_i - [\tilde{\theta}_1])^2}{\lambda_1^2} - \frac{1}{2}\sum_{j=1}^n \frac{(Y_j - [\tilde{\theta}_2])^2}{\lambda_2^2}\right]$$

ここで c_1 は (4) で定義したものである. また ORMLE に [] をつけたのは, ここでの期待値計算においては定数として扱われることを示すためである. 具体的には

$$\begin{aligned} E\left[\frac{(X_i - [\tilde{\theta}_1])^2}{\lambda_1^2}\right] &= E\left[\frac{(X_i - \theta_1 + \theta_1 - [\tilde{\theta}_1])^2}{\lambda_1^2}\right] \\ &= E\left[\frac{(X_i - \theta_1)^2}{\lambda_1^2}\right] + \frac{(\theta_1 - \tilde{\theta}_1)^2}{\lambda_1^2} = 1 + \frac{(\theta_1 - \tilde{\theta}_1)^2}{\lambda_1^2} \end{aligned}$$

となる. Y_j についても同じ計算をすると次を得る:

$$E[l([\tilde{\theta}_1], [\tilde{\theta}_2])] = c_1 - \frac{m+n}{2} - \frac{1}{2} \frac{(\tilde{\theta}_1 - \theta_1)^2}{\sigma_1^2} - \frac{1}{2} \frac{(\tilde{\theta}_2 - \theta_2)^2}{\sigma_2^2}.$$

次にこれを最大対数尤度 $l(\tilde{\theta}_1, \tilde{\theta}_2)$ で推定する. そこで

$$T = l(\tilde{\theta}_1, \tilde{\theta}_2) - E[l([\tilde{\theta}_1], [\tilde{\theta}_2])]$$

を定義する. この統計量 T の期待値が補正すべきバイアスになる. $l(\tilde{\theta}_1, \tilde{\theta}_2)$ の期待値は次のようになる:

$$E[l(\tilde{\theta}_1, \tilde{\theta}_2)] = c_1 - \frac{1}{2} E\left[\sum_{i=1}^m \frac{(X_i - \bar{X})^2}{\lambda_1^2} + \sum_{j=1}^n \frac{(Y_j - \bar{Y})^2}{\lambda_2^2} + \sum_{i=1}^2 \frac{(\hat{\theta}_i - \tilde{\theta}_i)^2}{\sigma_i^2}\right].$$

ここで

$$E \left[\sum_{i=1}^m \frac{(X_i - \bar{X})^2}{\lambda_1^2} \right] = m - 1, \quad E \left[\sum_{j=1}^n \frac{(Y_j - \bar{Y})^2}{\lambda_2^2} \right] = n - 1$$

である。また $\hat{\theta}_1 < \hat{\theta}_2$ のとき、 $\sum_{j=1}^n (\hat{\theta}_j - \bar{\theta}_j)^2 / \sigma_j^2 = 0$ であり、 $\hat{\theta}_1 > \hat{\theta}_2$ のとき、

$$\sum_{i=1}^2 \frac{(\hat{\theta}_i - \bar{\theta}_i)^2}{\sigma_i^2} = \frac{(\hat{\theta}_2 - \hat{\theta}_1)^2}{\sigma_1^2 + \sigma_2^2}$$

である。また(5)より次が得られる：

$$\begin{aligned} E \left[\sum \frac{(\hat{\theta}_i - \bar{\theta}_i)^2}{\sigma_i^2} \right] &= \int_{x>y} \frac{(y-x)^2}{\sigma_1^2 + \sigma_2^2} \frac{1}{\sigma_1} \phi \left(\frac{x-\theta_1}{\sigma_1} \right) \frac{1}{\sigma_2} \phi \left(\frac{y-\theta_2}{\sigma_2} \right) dx dy \\ &= \int_{u>a} (u-a)^2 \phi(u) \phi(v) dudv = (1+a^2) \Phi(-a) - a\phi(a). \end{aligned}$$

ただし、 a は(6)式のものである。同様に、(A.4)より次が得られる：

$$\begin{aligned} E \left[\sum \frac{(\tilde{\theta}_i - \theta_i)^2}{\sigma_i^2} \right] &= \int_{x \leq y} \left\{ \frac{(x-\theta_1)^2}{\sigma_1^2} + \frac{(y-\theta_2)^2}{\sigma_2^2} \right\} \\ &\quad \times \frac{1}{\sigma_1} \phi \left(\frac{x-\theta_1}{\sigma_1} \right) \frac{1}{\sigma_2} \phi \left(\frac{y-\theta_2}{\sigma_2} \right) dx dy \\ &\quad + \int_{x>y} \left\{ \frac{1}{\sigma_1^2} \left(\frac{\sigma_2^2 x + \sigma_1^2 y}{\sigma_1^2 + \sigma_2^2} - \theta_1 \right)^2 + \frac{1}{\sigma_2^2} \left(\frac{\sigma_2^2 x + \sigma_1^2 y}{\sigma_1^2 + \sigma_2^2} - \theta_2 \right)^2 \right\} \\ &\quad \times \frac{1}{\sigma_1} \phi \left(\frac{x-\theta_1}{\sigma_1} \right) \frac{1}{\sigma_2} \phi \left(\frac{y-\theta_2}{\sigma_2} \right) dx dy = \int_{u \leq a} (u^2 + v^2) \phi(u) \phi(v) dudv \\ &\quad + \int_{u > a} \left\{ \frac{1}{\sigma_1^2} \left(\frac{\sigma_1 \sigma_2 v}{\sqrt{\sigma_1^2 + \sigma_2^2}} + \frac{\sigma_1^2 (\theta_2 - \theta_1)}{\sigma_1^2 + \sigma_2^2} \right)^2 + \frac{1}{\sigma_2^2} \left(\frac{\sigma_1 \sigma_2 v}{\sqrt{\sigma_1^2 + \sigma_2^2}} - \frac{\sigma_2^2 (\theta_2 - \theta_1)}{\sigma_1^2 + \sigma_2^2} \right)^2 \right\} \\ &\quad \times \phi(u) \phi(v) dudv \\ &= \int_{u \leq a} (u^2 + v^2) \phi(u) \phi(v) dudv + \int_{u > a} (v^2 + a^2) \phi(u) \phi(v) dudv \\ &= \int_{u \leq a} u^2 \phi(u) du + 1 + \int_{u > a} a^2 \phi(u) du = 1 - \{\Phi(-a) + a\phi(a)\} + 1 + a^2 \{1 - \Phi(a)\} \\ &= 1 + \Phi(a) - a\phi(a) + a^2 - a^2 \Phi(a), \end{aligned}$$

よって

$$\begin{aligned}
 E[T] &= 1 - \frac{1}{2} E \left[\frac{(\tilde{\theta}_1 - \hat{\theta}_1)^2}{\sigma_1^2} + \frac{(\tilde{\theta}_2 - \hat{\theta}_1)^2}{\sigma_2^2} \right] + \frac{1}{2} E \left[\frac{(\tilde{\theta}_1 - \theta_1)^2}{\sigma_1^2} + \frac{(\tilde{\theta}_2 - \theta_1)^2}{\sigma_2^2} \right] \\
 &= 1 - \frac{1}{2} \{ (1+a^2) \Phi(-a) - a\phi(a) \} + \frac{1}{2} [1 + \Phi(a) - a\phi(a) + a^2 \{1 - \Phi(a)\}] \\
 &= 1 + \Phi(a) \tag{14}
 \end{aligned}$$

となり、きわめて簡潔な式が得られる。これがバイアス補正項になる。なお、 $a=0$ のときこの値は $3/2$ となるが、これは Anraku (1999) の補正項についての結果に対応している。

次に、 $E[\tilde{\theta}_1^*]$ 、 $E[\tilde{\theta}_2^*]$ を求めた場合と同じようにブートストラップ法を適用してみる。

$$X_1^*, X_2^*, \dots, X_{m_B}^* (\sim N(\hat{\theta}_1, \lambda_1^2)); Y_1^*, Y_2^*, \dots, Y_{n_B}^* (\sim N(\hat{\theta}_2, \lambda_2^2))$$

とし、 σ_{B1}^2 、 σ_{B2}^2 も先と同じとする。ここで

$$T^* = l(\tilde{\theta}_1^*, \tilde{\theta}_2^*) - E[l([\tilde{\theta}_1^*], [\tilde{\theta}_2^*])]$$

とおく。これを B 回生成する。ここで、(2)より

$$l(\tilde{\theta}_1^*, \tilde{\theta}_2^*) = c_1 - \frac{1}{2} \sum_{i=1}^{m_B} \frac{(x_i^* - \tilde{\theta}_1^*)^2}{\lambda_1^2} - \frac{1}{2} \sum_{j=1}^{n_B} \frac{(y_j^* - \tilde{\theta}_2^*)^2}{\lambda_2^2}$$

$$E[l([\tilde{\theta}_1^*], [\tilde{\theta}_2^*])] = E[l([\hat{\theta}_1^*], [\hat{\theta}_1^*])] = c_1 - \frac{m_B + n_B}{2} - \frac{1}{2} \frac{(\hat{\theta}_1^* - \hat{\theta}_1)^2}{\sigma_{B1}^2} - \frac{1}{2} \frac{(\hat{\theta}_2^* - \hat{\theta}_2)^2}{\sigma_{B2}^2}$$

である。なお先の結果(14)により、 T^* の期待値は

$$E^*[T^* | \hat{\theta}_1, \hat{\theta}_2] = 1 + \Phi(\hat{a})$$

となる。これを \hat{a} について期待値をとると(A.1)から次が得られる：

$$E[T^*] = 1 + \Phi\left(\frac{\theta_2 - \theta_1}{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_{B1}^2 + \sigma_{B2}^2}}\right).$$

この式から m_B 、 n_B を大きくするとき、 $\sigma_{B1}^2 \rightarrow 0$ 、 $\sigma_{B2}^2 \rightarrow 0$ となり、 $E[T^*] \Rightarrow E[T]$ となる。しかし、任意の $\varepsilon (0 < \varepsilon < 1)$ に対して

$$P(1 + \Phi(\hat{a}) \geq 1 + \epsilon) = P(\Phi(\hat{a}) \geq \epsilon) = P(\hat{a} \geq c(\epsilon)).$$

ここで、 $c(\epsilon)$ は $\int_d^\infty \phi(u) du = \epsilon$ により決まる d を表すものとする、

$$P(\hat{a} \geq c(\epsilon)) = \int_{c(\epsilon)}^\infty \frac{1}{\gamma} \phi\left(\frac{x-\mu}{\gamma}\right) dx = \Phi\left(\frac{C(\epsilon)\sqrt{\sigma_{B1}^2 + \sigma_{B2}^2} - (\theta_2 - \theta_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

である。この場合も m_B, n_B を大きくするとき、 $\sigma_{B1}^2 \rightarrow 0, \sigma_{B2}^2 \rightarrow 0$ となり、上の式は $\Phi(-a)$ に近づく。この確率も ϵ に依存しない二項確率と考えられる。

このようにブートストラップ推定では、 $m_B = m, n_B = n$ とするとき、 $E(T^*) = 1 + \Phi\left(\frac{\alpha}{\sqrt{2}}\right) \approx 1 + \Phi(\alpha) = E(T)$ である。また m_B, n_B を大きくすると、ブートストラップ推定量の期待値は期待値 $E[T]$ に収束するが、一致性は成り立たないことがわかる。

5 おわりに

本稿では、2つの正規分布の平均母数順序制約がある場合の最尤推定量 (ORMLE) と情報量基準に関連して導出される統計量の期待値を求めた。またこれらをブートストラップ法で推定した場合の期待値を求めたが、少なくともここで設定した方法ではうまく機能しないことが示された。多標本の場合でも同じ状況が予想される。

またブートストラップ標本を生成する際に、 θ_1, θ_2 を、制約のない MLE $\hat{\theta}_1, \hat{\theta}_2$ に置き換えたが、ORMLE $\bar{\theta}_1, \bar{\theta}_2$ を用いることも考えられる。しかし、ここでは用いなかった。理由としては、計算の困難さと、バイアス等の結果がいつそう著しくなることが予想されるからである。

なお、ブートストラップ法がうまく機能しない場合が様々に指摘されている (Bickel and Freedman (1981), Beran (1997), Ishiguro et al.(1997), Andrews (2000) 他)。一方このような不具合を改善すべくいくつかの方法が提案されている (例えば, Davison and Hinkley (1997), Politis et al.(1999), 下平 (2012) などを参照されたい)。しかし、このような方法で本稿で取りあげた問題が改善されるのか、今のところ (少なくとも著者にとっては) 不明である。今後の課題としたい。

6 補遺

1. μ, γ を定数とし, $Z \sim N(0, 1)$ とするとき,

$$E[\Phi(\mu + \gamma Z)] = \int_{-\infty}^{\infty} \Phi(\mu + \gamma z) \phi(z) dz = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\mu + \gamma z} \phi(w) dw \right\} \phi(z) dz$$

直交変換

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -\frac{\gamma}{\sqrt{1+\gamma^2}} & \frac{1}{\sqrt{1+\gamma^2}} \\ -\frac{1}{\sqrt{1+\gamma^2}} & -\frac{\gamma}{\sqrt{1+\gamma^2}} \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix}$$

を行うと, (x, y) の積分領域は

$$\{(z, w) \mid w \leq \gamma z + \mu\} \Rightarrow \{(u, v) \mid u \leq \frac{\mu}{\sqrt{1+\gamma^2}}\}$$

と変換されるので, 次の関係式を得る:

$$E[\Phi(\mu + \gamma Z)] = \Phi\left(\frac{\mu}{\sqrt{1+\gamma^2}}\right). \quad (\text{A.1})$$

2. 積分

$$\int_{-\infty}^{\infty} \phi(\gamma z + \mu) \phi(z) dz$$

の被積分関数は次のように表される:

$$\begin{aligned} \phi(\gamma z + \mu) \phi(z) &= \frac{1}{2\pi} \exp\left\{-\frac{(\gamma z + \mu)^2}{2} - \frac{z^2}{2}\right\} \\ &= \frac{1}{2\pi} \exp\left\{-\frac{1+\gamma^2}{2} \left(z + \frac{\gamma\mu}{1+\gamma^2}\right)^2 - \frac{\mu^2}{2(1+\gamma^2)}\right\} \end{aligned}$$

したがって

$$\begin{aligned} \int_{-\infty}^{\infty} \phi(\gamma z + \mu) \phi(z) dz &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1+\gamma^2}} \exp\left\{-\frac{\mu^2}{2(1+\gamma^2)}\right\} \\ &\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\{1/(1+\gamma^2)\}}} \exp\left\{-\frac{1+\gamma^2}{2} \left(z + \frac{\gamma\mu}{1+\gamma^2}\right)^2\right\} dz \end{aligned}$$

となり, 次が成り立つ:

$$\int_{-\infty}^{\infty} \phi(\gamma z + \mu) \phi(z) dz = \frac{1}{\sqrt{1+\gamma^2}} \phi\left(\frac{\mu}{\sqrt{1+\gamma^2}}\right) \quad (\text{A. 2})$$

3. 部分積分により

$$\int_{-\infty}^{\infty} z \Phi(\gamma z + \mu) \phi(z) dz = -\int_{-\infty}^{\infty} \Phi(\gamma z + \mu) \phi'(z) dz = \gamma \int \phi(\gamma z + \mu) \phi(z) dz.$$

よって、次を得る。

$$\int_{-\infty}^{\infty} z \Phi(\gamma z + \mu) \phi(z) dz = \frac{\gamma}{\sqrt{1+\gamma^2}} \phi\left(\frac{\mu}{\sqrt{1+\gamma^2}}\right) \quad (\text{A. 3})$$

4.

$$\int_{z>a} z^2 \phi(z) dz = -\int_{z>a} z \phi'(z) dz = -[z\phi(z)]_a^{\infty} + \int_{z>a} \phi(z) dz \quad (\text{A. 4})$$

より次が成り立つ:

$$\int_{z>a} z^2 \phi(z) dz = \Phi(-a) + a\phi(a) \quad (\text{A. 5})$$

参考文献

- Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle", In *2nd International Symposium on Information Theory*, Ed. B. N. Petrov and F. Csaki, pp. 267–81. Budapest: Akademiai Kiado. (*Breakthroughs in Statistics 1*, (1992). Ed. S. Kotz and N. L. Johnson, pp. 610–24. New York: Springer-Verlag. に再録)
- Akaike, H. (1974). A New Look at the Statistical Model Identification", *IEEE Transactions on Automatic Control*, AC-19, 716–723.
- Andrews, W.K. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* **68**, 399–405.
- Anraku, K. (1999). An Information Criterion for Parameters under a Simple Order Restriction", *Biometrika* **86**, 141–152.
- Bickel, P. J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap", *Annals of Statistics* **9**, 1196–1217.
- Beran, R. (1997). Diagnosing bootstrap success. *Annals of the Institute of Statistical Mathematics* **49**, 1–24.
- Chernoff, H. (1954). On the distribution of the likelihood ratio. *Annals of Mathematical*

- Statistics* **25**, 573–578.
- Davison, A. C. and Hinkley, D.V. (1997). *Bootstrap methods and their applications*. Cambridge University Press.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics* **7**, 1–26.
- Ishiguro, M., Sakamoto, Y. and Kitagawa, G. (1997). Bootstrapping log likelihood and EIC, an extension of AIC”, *Annals of the Institute of Statistical Mathematics* **49**, 411–434.
- Politis, D.N., Romano, J.P., and Wolf, M. (1999). *Subsampling*. Springer, New York.
- 坂元慶行, 石黒真木夫, 北川源四郎 (1983), 「情報量統計学」, 共立出版.
- 下平英寿 (2012). マルチスケール・ブートストラップの漸近理論, 日本統計学会誌, **41**, 445–473.