数理解析研究所講究録 1273 巻 2002 年 16-28

Asymptotics on the Maximum Probability Estimators in Statistical Experiments

筑波大・数学　大谷内奈穂 (Nao Ohyauchi)
筑波大・数学　赤平　昌文 (Masafumi Akahira)

1　はじめに

ある母数 θ をもつ母集団分布からの無作為標本に基づいて θ の推定を行うが, 正則条件が必ずしも満たされていないような非正則な場合には, 最尤推定量は良い推定量とは限らない. そのような場合には (一般) ベイズ推定量や Weiss and Wolfowitz [WW74] による最大確率推定量が漸近的に有効であることが多い (Akahira and Takeuchi [AT81], Ibragimov and Has'minskii [IH81], Akahira [A91], Shono [S96]). また, Blyth [B82] は, 指数分布や一様分布の場合に, 小標本に基づいて最大確率推定量, 最尤推定量等の有効性について論じた (Weiss [W86]). 一方, Bayes リスクの観点から, 一様分布, 切断正規分布の位置母数の推定についても論じられ (Ohyauchi and Akahira [OA01], Ohyauchi [O02]), さらに一般の非正則分布族の場合にも推定量の Bayes リスクに関する情報不等式が得られている (Akahira and Ohyauchi [AO02]).

本論では, まず, 母数 θ を推論するための統計的実験において, ある条件を満たす密度をもつ分布族における位置母数の推定について考える. 特に, 最大確率推定量が最尤推定量よりも漸近的に良くなることを漸近平均 2 乗誤差, 漸近平均絶対誤差, 誤差確率による相対効率を用いて示す. また, 尺度母数をもつ分布族において, その最大確率推定量を漸近的に求める方法を提案し, 具体的には, 指数分布, 正規分布, ガンマ分布, ワイブル分布等の場合に MPE を求める. 特に, 指数分布の場合には, 最尤推定量, Bayes 推定量等の (漸近的) 有効性についても考察する.

2　設定

まず, $X_1, X_2, \cdots, X_n, \cdots$ を互いに独立にいずれも (ある σ-有限測度 μ に関して絶対連続な) 確率密度関数 (p.d.f.) $p(x,\theta)$ ($\theta \in \Theta$) をもつ分布に従う確率変数列とし, Θ は $\mathbf{R}^1$ の開区間とする. このとき, $\mathbf{X} := (X_1, \cdots, X_n) = \mathbf{x} := (x_1, \cdots, x_n)$ を与えたときの, θ の尤度関数は

$$L_{\mathbf{x}}(\theta) = \prod_{i=1}^{n} p(x_i, \theta)$$

になる. ここで, $\mathbf{X}$ の標本空間を $\mathcal{X}$ とするとき, θ に関する推論のための実験 (experiment) を $(\mathbf{X},\ \mathcal{X},\ L_{\cdot}(\theta),\ \theta \in \Theta)$ で表し, 本論ではこの実験上で考察する. 次に, θ の推定量

$\hat{\theta}_n := \hat{\theta}_n(\mathbf{X})$ の一致性のオーダーを $\{c_n\}$ とする．そして，任意に固定した正数 r に対して

$$\int_{d-(r/c_n)}^{d+(r/c_n)} L_{\mathbf{x}}(\theta)d\theta$$

を最大にする d を θ の最大確率推定量 (maximum probability estimator, 略して MPE) といい，$\hat{\theta}^r_{MP}$ で表す ([WW74])．ここで，MPE は r に依存することに注意．適当な正則条件の下では，c_n は $\sqrt{n}$ であるが，非正則な場合には $n^{1/\alpha}$ $(0 < \alpha < 2)$, $\sqrt{n \log n}$ 等になる (詳しくは Akahira and Takeuchi [AK81] 参照)．なお，MPE は損失関数として

$$l_r(\theta, d) = \begin{cases} 0 & (|\theta - d| \leq r/c_n), \\ 1 & (|\theta - d| > r/c_n) \end{cases}$$

をとり，事前測度としてルベーグ測度をとったときに一般ベイズ推定量 (generalized Bayes estimator, 略して GBE) になっている．ただし，$\int_\Theta L_{\mathbf{x}}(\theta)d\theta < \infty$ とする．

3　位置母数の推定

第2節の設定の下で，(ルベーグ測度に関する) p.d.f. p が次の条件 (A1) ～ (A3) を満たすとする．
(A1)　$p(x) > 0$ $(a < x < b)$, $p(x) = 0$ (その他)．
(A2)　p は区間 (a, b) において連続微分可能で，$p'(x) < 0$ $(a < x < b)$ である．
(A3)　$k := \lim_{x \to a+0} p(x) (> 0)$ が存在する．
ただし，$-\infty < a < b \leq \infty$ とする．いま，$\Theta = \mathbf{R}^1$ とし，θ が位置母数の場合，すなわち X_1 の p.d.f. が $p(x - \theta)$ の場合を考える．このときには，$a = 0$ として一般性を失わない．この場合の位置母数 θ の推定は非正則な場合になり，$c_n = n$ となる ([AT81])．さて，θ の尤度関数は

$$L_{\mathbf{x}}(\theta) = \prod_{i=1}^{n} p(x_i - \theta) > 0 \qquad (\underline{\theta} < \theta < \bar{\theta})$$

になる．ただし，$\underline{\theta} := \max_{1 \leq i \leq n} x_i - b$, $\bar{\theta} := \min_{1 \leq i \leq n} x_i$ とする．このとき，条件 (A2) より，$\underline{\theta} < \theta < \bar{\theta}$ について

$$\frac{\partial}{\partial \theta} \log L_{\mathbf{x}}(\theta) = -\sum_{i=1}^{n} \frac{p'(x_i - \theta)}{p(x_i - \theta)} > 0$$

となるから，$\log L_{\mathbf{x}}(\theta)$ は区間 $(\underline{\theta},\ \bar{\theta})$ において単調増加関数になる．よって，θ の最尤推定量 (maximum likelihood estimator, 略して MLE) $\hat{\theta}_{ML}$ および MPE $\hat{\theta}^r_{MP}$ はそれぞれ

$$\hat{\theta}_{ML} = \min_{1 \leq i \leq n} X_i = \bar{\theta},$$
$$\hat{\theta}^r_{MP} = \min_{1 \leq i \leq n} X_i - \frac{r}{n} = \bar{\theta} - \frac{r}{n}$$

になる. ここで $r \to 0$ とすれば $\hat{\theta}^r_{MP}$ が $\hat{\theta}_{ML}$ に収束することに注意. いま, $\bar{\theta}$ について, $n(\bar{\theta}-\theta)$ の漸近 (asymptotic (as.)) p.d.f. は

$$f_{\bar{\theta}}(t) = \begin{cases} ke^{-kt} + o(1) & (t > 0), \\ 0 & (t \leq 0) \end{cases} \tag{3.1}$$

になる. よって, (3.1) より $\hat{\theta}_{ML}$, $\hat{\theta}^r_{MP}$ の漸近平均 2 乗誤差 (AMSE) は

$$\mathrm{AMSE}_\theta(\hat{\theta}_{ML}) = E_\theta[\{n(\bar{\theta}-\theta)\}^2] = \frac{2}{k^2} + o(1), \tag{3.2}$$

$$\mathrm{AMSE}_\theta(\hat{\theta}^r_{MP}) = E_\theta\left[\left\{n\left(\bar{\theta} - \frac{r}{n} - \theta\right)\right\}^2\right] = \left(r - \frac{1}{k}\right)^2 + \frac{1}{k^2} + o(1) \tag{3.3}$$

となる. 従って, (3.3) より $\mathrm{AMSE}_\theta(\hat{\theta}^r_{MP})$ を最小にする r は $r = 1/k$ となり, これを最良 (best) MPE という. 最良 MPE $\hat{\theta}^{1/k}_{MP}$ の AMSE は

$$\mathrm{AMSE}_\theta(\hat{\theta}^{1/k}_{MP}) = E_\theta\left[\left\{n\left(\bar{\theta} - \frac{1}{kn} - \theta\right)\right\}^2\right] = \frac{1}{k^2} + o(1) \tag{3.4}$$

となり, (3.2), (3.4) より

$$\frac{\mathrm{AMSE}_\theta(\hat{\theta}^{1/k}_{MP})}{\mathrm{AMSE}_\theta(\hat{\theta}_{ML})} = \frac{1}{2} + o(1)$$

となり, 最良 MPE に対する MLE の漸近的な相対効率が 50% であることを意味している.

また, 上記の AMSE の代わりに漸近平均絶対誤差 (asymptotic mean absolute error, 略して AMAE) を用いれば, θ の MLE $\hat{\theta}_{ML}$, MPE $\hat{\theta}^r_{MP}$ について, (3.1) より

$$\mathrm{AMAE}_\theta(\hat{\theta}_{ML}) = E_\theta[n|\bar{\theta}-\theta|] = \frac{1}{k} + o(1), \tag{3.5}$$

$$\mathrm{AMAE}_\theta(\hat{\theta}^r_{MP}) = E_\theta[n|\hat{\theta}^r_{MP}-\theta|] = r + \frac{2}{k}e^{-kr} - \frac{1}{k} + o(1) \tag{3.6}$$

となり, (3.6) より $\mathrm{AMAE}_\theta(\hat{\theta}^r_{MP})$ を最小にする r は $r = (1/k)\log 2 =: r_0$ になる. このとき, $\hat{\theta}^{r_0}_{MP}$ について

$$\mathrm{AMAE}_\theta(\hat{\theta}^{r_0}_{MP}) = \frac{1}{k}\log 2 + o(1)$$

になり, (3.5) より

$$\frac{\mathrm{AMAE}_\theta(\hat{\theta}^{r_0}_{MP})}{\mathrm{AMAE}_\theta(\hat{\theta}_{ML})} \approx \log 2 \ (\approx 0.693) \qquad (n \to \infty)$$

となり, $\hat{\theta}^{r_0}_{MP}$ に対する $\hat{\theta}_{ML}$ の漸近的な相対効率が約 70% であることを意味する.

さらに, 比較基準として, θ の推定量 $\hat{\theta} = \hat{\theta}(\mathbf{X})$ について, 任意の $\varepsilon > 0$ に対して, $\hat{\theta}$ の誤

差確率 (error probability, 略して EP) $\mathrm{EP}_\theta(\hat\theta) = P_\theta\{n|\hat\theta - \theta| > \varepsilon\}$ を考えると, θ の MLE $\hat\theta_{ML}$, MPE $\hat\theta^r_{MP}$ は

$$\mathrm{EP}_\theta(\hat\theta_{ML}) = e^{-k\varepsilon} + o(1), \tag{3.7}$$

$$\mathrm{EP}_\theta(\hat\theta^r_{MP}) = \begin{cases} 1 - e^{-kr}(e^{k\varepsilon} - e^{-k\varepsilon}) + o(1) & (0 < \varepsilon < r), \\ e^{-k(\varepsilon + r)} + o(1) & (r \le \varepsilon) \end{cases} \tag{3.8}$$

となり, (3.8) より $\mathrm{EP}_\theta(\hat\theta^r_{MP})$ を最小にする r は $r = \varepsilon$ となるから最良 MPE $\hat\theta^\varepsilon_{MP}$ は

$$\mathrm{EP}_\theta(\hat\theta^\varepsilon_{MP}) = e^{-2k\varepsilon} + o(1)$$

となり, (3.7) より

$$\frac{\mathrm{EP}_\theta(\hat\theta^\varepsilon_{MP})}{\mathrm{EP}_\theta(\hat\theta_{ML})} \approx e^{-k\varepsilon} < 1 \tag{3.9}$$

となり, 最良 MPE $\hat\theta^\varepsilon_{MP}$ が MLE より漸近的に良いことが分かる.

以上のことから, AMSE, AMAE についての最良 MPE に対する MLE の漸近的な相対効率は, それぞれ約 50%, 約 70% と求められたが, EP についてのその値は, 任意に与えられる ε と $k = \lim_{x \to a+0} p(x)$ に依存することがわかる. そこで, EP について具体的に例を挙げて考察する.

例 3.1　$p(x)$ が指数分布の p.d.f.

$$p(x) = \begin{cases} e^{-x} & (x > 0), \\ 0 & (x \le 0) \end{cases}$$

であれば, 条件 (A1)〜(A3) は満たされ, $k = 1$ となる. よって, (3.9) より

$$\frac{\mathrm{EP}_\theta(\hat\theta^\varepsilon_{MP})}{\mathrm{EP}_\theta(\hat\theta_{ML})} \approx e^{-\varepsilon}$$

となるから, $\varepsilon = 1/2$ なら $e^{-\varepsilon} \approx 0.607$ となり, $\hat\theta^{1/2}_{MP}$ に対する $\hat\theta_{ML}$ の漸近的な相対効率が約 61% であることを意味する. なお, Blyth [B82] は, 小標本論の立場からこの相対効率の値を求めている.

例 3.2　$p(x)$ が片側正規分布の p.d.f.

$$p(x) = \begin{cases} \sqrt{2/\pi}\, e^{-x^2/2} & (x > 0), \\ 0 & (x \le 0) \end{cases}$$

であれば, 条件 (A1)〜(A3) は満たされ, $k = \sqrt{2/\pi}$ となる. よって, (3.9) より

$$\frac{\mathrm{EP}_\theta(\hat\theta^\varepsilon_{MP})}{\mathrm{EP}_\theta(\hat\theta_{ML})} \approx e^{-\varepsilon\sqrt{2/\pi}}$$

となるから, $\varepsilon = 1/2$ なら $e^{-\varepsilon\sqrt{2/\pi}} = e^{-\sqrt{1/(2\pi)}} \approx 0.671$ となり, $\hat{\theta}_{MP}^{1/2}$ に対する $\hat{\theta}_{ML}$ の漸近的な相対効率が約 67% であることが分かる.

また, 例 3.1, 3.2 における漸近相対効率 $e^{-\varepsilon}$, $e^{-\varepsilon\sqrt{2/\pi}}$ について, $\varepsilon = 0.05(0.05)0.50$ の場合にそれらの値を求めた (表 3.1 参照). その結果, ε が 0 に近い場合は最良 MPE と MLE の良さにそれほど差はないが, ε が大きくなるにつれて, その良さに差が出てくるということも分かる.

表 3.1 EP についての最良 MPE に対する MLE の漸近相対効率

ε	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
例 3.1	0.95	0.90	0.86	0.82	0.78	0.74	0.70	0.67	0.64	0.61
例 3.2	0.96	0.92	0.89	0.85	0.82	0.79	0.76	0.73	0.70	0.67

4 尺度母数の推定

第 2 節の設定の下で, (ルベーグ測度に関する) p.d.f. p が

$$p(x, \theta) = \frac{1}{\theta} q\left(\frac{x}{\theta}\right) \tag{4.1}$$

の形であるとし, q は $\mathbf{R}^1$ 上で 3 回連続微分可能な関数とする. ただし, $\theta \in \Theta = (0, \infty)$ とする. このとき, θ は尺度母数になり, この場合の θ の推定は正則な場合となり, $c_n = \sqrt{n}$ になる.

さて, θ の尤度関数は

$$L_{\mathbf{x}}(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} q\left(\frac{x_i}{\theta}\right)$$

となるから,

$$h(d) := \int_{d-(r/\sqrt{n})}^{d+(r/\sqrt{n})} \frac{1}{\theta^n} \prod_{i=1}^{n} q\left(\frac{x_i}{\theta}\right) d\theta \tag{4.2}$$

を最大にする d, すなわち θ の MPE $\hat{\theta}_{MP}^{r}$ を求めよう. そのために (4.1) から

$$h'(d) = \left(d + \frac{r}{\sqrt{n}}\right)^{-n} \prod_{i=1}^{n} q\left(\frac{x_i}{d + \frac{r}{\sqrt{n}}}\right) - \left(d - \frac{r}{\sqrt{n}}\right)^{-n} \prod_{i=1}^{n} q\left(\frac{x_i}{d - \frac{r}{\sqrt{n}}}\right) = 0 \tag{4.3}$$

となる d を求める. そこで, (4.3) から

$$\log\left(d + \frac{r}{\sqrt{n}}\right) - \frac{1}{n}\sum_{i=1}^{n} \log q\left(\frac{x_i}{d + \frac{r}{\sqrt{n}}}\right) = \log\left(d - \frac{r}{\sqrt{n}}\right) - \frac{1}{n}\sum_{i=1}^{n} \log q\left(\frac{x_i}{d - \frac{r}{\sqrt{n}}}\right)$$

となる. ここで

$$\log\left(d+\frac{r}{\sqrt{n}}\right)=\log d+\log\left(1+\frac{r}{d\sqrt{n}}\right)$$
$$=\log d+\frac{r}{d\sqrt{n}}-\frac{r^2}{2d^2n}+\frac{r^3}{3d^3n^{3/2}}+o\left(n^{3/2}\right)$$

になる. また,

$$z_1(\mathbf{x}):=-\frac{1}{n}\sum_{i=1}^{n}x_i\left(\log q(x_i)\right)',$$
$$z_2(\mathbf{x}):=-\frac{1}{n}\sum_{i=1}^{n}x_i^2\left(\log q(x_i)\right)'',$$
$$z_3(\mathbf{x}):=-\frac{1}{n}\sum_{i=1}^{n}x_i^3\left(\log q(x_i)\right)'''$$

とすれば

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\log q\left(\frac{x_i}{d+\frac{r}{\sqrt{n}}}\right)\\&=\frac{1}{n}\sum_{i=1}^{n}\log q\left(\frac{x_i}{d}\right)+\frac{r}{d\sqrt{n}}z_1\left(\frac{1}{d}\mathbf{x}\right)-\frac{r^2}{d^2n}\left\{z_1\left(\frac{1}{d}\mathbf{x}\right)+\frac{1}{2}z_2\left(\frac{1}{d}\mathbf{x}\right)\right\}\\&+\frac{r^3}{d^3n\sqrt{n}}\left\{z_1\left(\frac{1}{d}\mathbf{x}\right)+z_2\left(\frac{1}{d}\mathbf{x}\right)+\frac{1}{6}z_3\left(\frac{1}{d}\mathbf{x}\right)\right\}+o\left(\frac{1}{n\sqrt{n}}\right)\end{aligned}\tag{4.4}$$

となる. よって, (4.4) より, 次のことが成り立つ.

定理 $X_1,X_2,\cdots,X_n,\cdots$ が互いに独立に, いずれも p.d.f. (4.1) をもつ分布に従うとする. このとき,

$$d=dz_1\left(\frac{1}{d}\mathbf{X}\right)-\frac{r^2}{3dn}+\frac{r^2}{dn}\left\{z_1\left(\frac{1}{d}\mathbf{X}\right)+z_2\left(\frac{1}{d}\mathbf{X}\right)+\frac{1}{6}z_3\left(\frac{1}{d}\mathbf{X}\right)\right\}+o_p\left(\frac{1}{n}\right)\tag{4.5}$$

を満たす d は θ の漸近的な MPE の候補である.

系 上の定理において, $z_i(\mathbf{X})$ $(i=1,2,3)$ が尺度不変ならば,

$$\begin{aligned}d&=z_1(\mathbf{X})-\frac{r^2}{3dn}+\frac{r^2}{d^2n}\left\{z_1(\mathbf{X})+z_2(\mathbf{X})+\frac{1}{6}z_3(\mathbf{X})\right\}+o_p\left(\frac{1}{n}\right)\\&=z_1(\mathbf{X})-\frac{r^2}{3z_1(\mathbf{X})n}+\frac{r^2}{z_1(\mathbf{X})n}\left\{1+\frac{z_2(\mathbf{X})}{z_1(\mathbf{X})}+\frac{z_3(\mathbf{X})}{6z_1(\mathbf{X})}\right\}+o_p\left(\frac{1}{n}\right)\end{aligned}\tag{4.6}$$

注意 4.1　θ の尤度関数 $L_{\mathbf{x}}(\theta)$ が単峰ならば, 方程式 (4.3) の解は MPE になる. 実際,

$$h'(d) = L_{\mathbf{x}}\left(d + \frac{r}{\sqrt{n}}\right) - L_{\mathbf{x}}\left(d - \frac{r}{\sqrt{n}}\right)$$

より $L_{\mathbf{x}}(\theta)$ の単峰性から, $h'(d) = 0$ の解を $d = d^*$ とすれば, $d < d^*$ について $h'(d) > 0$, $d > d^*$ について $h'(d) < 0$ となり, h も単峰形となる. よって, $h(d)$ は d^* で最大値をとるので, d^* が MPE になることが分かる.

例 4.1　q として

$$q(x) = \begin{cases} e^{-x} & (x > 0), \\ 0 & (x \leq 0) \end{cases}$$

とすれば, (4.1) の $p(x, \theta) = (1/\theta)q(x/\theta)$ は, 尺度母数 θ をもつ指数分布の p.d.f. になる. このとき, $\log q(x) = -x$ となるから, $z_1(\mathbf{x}) = \bar{x} := (1/n)\sum_{i=1}^n x_i, z_2(\mathbf{x}) = z_3(\mathbf{x}) = 0$ となる. よって, $z_1(\mathbf{x}) = \bar{x}$ は尺度不変であるから, (4.6) から θ の MPE は, $n \to \infty$ のとき

$$\hat{\theta}^r_{MP} = \bar{X} + \frac{2r^2}{3n\bar{X}} + o_p\left(\frac{1}{n}\right) \tag{4.7}$$

になる. ただし $\bar{X} = (1/n)\sum_{i=1}^n X_i$ とする. ここで, $P(\bar{X} = 0) = 0$ であることに注意.

一方, θ の MLE は $\hat{\theta}_{ML} = \bar{X}$ になり, MPE の定義から $\hat{\theta}^r_{MP} \to \hat{\theta}_{ML}$ $(r \to 0)$ になるので, (4.7) からも確かめられる.

次に, c を正の定数として, $c\bar{X}$ の形の推定量を考えて, θ の周りでの集中確率

$$P_\theta\left\{|c\bar{X} - \theta| < \frac{r}{\sqrt{n}}\right\} \tag{4.8}$$

を最大にする c を求めてみよう. まず

$$g(c) := P_\theta\left\{\theta - \frac{r}{\sqrt{n}} < c\bar{X} < \theta + \frac{r}{\sqrt{n}}\right\}$$

とおくと

$$g(c) = P_\theta\left\{\frac{n}{c}\left(\theta - \frac{r}{\sqrt{n}}\right) < n\bar{X} < \frac{n}{c}\left(\theta + \frac{r}{\sqrt{n}}\right)\right\}$$

となり, $Z := \sum_{i=1}^n X_i$ の p.d.f. は

$$f_Z(z) = \frac{1}{\Gamma(n)}\frac{z^{n-1}}{\theta^n}e^{-z/\theta} \qquad (0 < z < \infty; \theta > 0)$$

となるから

$$g(c) = \int_{\frac{n}{c}(\theta - (r/\sqrt{n}))}^{\frac{n}{c}(\theta + (r/\sqrt{n}))} \frac{1}{\Gamma(n)}\frac{z^{n-1}}{\theta^n}e^{-\frac{z}{\theta}}dz$$

となり, $g'(c)=0$ となる c を, MPE を漸近的に求めたときと同様な方法で, 漸近的に

$$c = 1 - \frac{r^2}{3n\bar{X}^2} + o\left(\frac{1}{n}\right)$$

を得る. このことから, 集中確率 (4.8) を最大にする推定量は漸近的に

$$\hat{\theta}_n = \bar{X} - \frac{r^2}{3n\bar{X}} + o_p\left(\frac{1}{n}\right)$$

で与えられ, これは, MPE $\hat{\theta}^r_{MP}$ とは $1/n$ のオーダーで係数が異なる形になる.

さらに, Bayes 的観点から θ の推定を考えてみよう. θ の一般事前密度 (improper prior density) として

$$\pi(\theta) = \begin{cases} 1/\theta & (\theta > 0), \\ 0 & (\theta \le 0) \end{cases}$$

をとり, $z := \sum_{i=1}^n x_i$ とすれば, θ の事後密度 (posterior density) は

$$\pi(\theta|z) = \frac{z^n}{\Gamma(n)}\frac{1}{\theta^{n+1}}e^{-z/\theta} \qquad (\theta > 0)$$

となる. このとき, θ の事後モード (posterior mode) $\hat{\theta}_{MO}$, 事後平均 (すなわち Bayes 推定量) $\hat{\theta}_{BA}$ は, それぞれ

$$\hat{\theta}_{MO} = \frac{1}{n+1}\sum_{i=1}^n X_i =: \frac{1}{n+1}Z, \quad \hat{\theta}_{BA} = \frac{1}{n-1}\sum_{i=1}^n X_i =: \frac{1}{n-1}Z$$

になる. いま, 損失関数として

$$l(\theta, d) = (d-\theta)^2/\theta^2$$

をとると, これは不変になる. このとき, θ の推定量 $\hat{\theta}_n := \hat{\theta}_n(Z)$ の事後リスクを考えると

$$r(\hat{\theta}_n|z) = E\left[\frac{(\hat{\theta}_n-\theta)^2}{\theta^2} \;\middle|\; z\right] = \hat{\theta}_n^2 E\left(\frac{1}{\theta^2} \;\middle|\; z\right) - 2\hat{\theta}_n E\left(\frac{1}{\theta} \;\middle|\; z\right) + 1$$

となり, これを最小にする $\hat{\theta}_n$ は

$$\hat{\theta}^*_n = E\left(\frac{1}{\theta} \;\middle|\; Z\right) \Big/ E\left(\frac{1}{\theta^2} \;\middle|\; Z\right) = \frac{1}{n+1}Z$$

となり, 事後モード $\hat{\theta}_{MO}$ に一致する. また, 損失関数として

$$l(\theta, d) = \frac{d}{\theta} - 1 - \log\frac{d}{\theta}$$

をとれば, これは不変になる. このとき, θ の推定量 $\hat{\theta}_n := \hat{\theta}_n(Z)$ の事後リスク

$$\begin{aligned} r(\hat{\theta}_n|z) &= E\left[\frac{\hat{\theta}_n}{\theta} - 1 - \log\frac{\hat{\theta}_n}{\theta} \;\middle|\; z\right] \\ &= \hat{\theta}_n E\left(\frac{1}{\theta} \;\middle|\; z\right) - 1 - \log\hat{\theta}_n + E(\log\theta|z) \end{aligned}$$

を最小にする $\hat{\theta}_n$ は

$$\hat{\theta}_n^{**} = 1 \Big/ E\left(\frac{1}{\theta} \;\middle|\; Z\right) = \frac{1}{n}Z$$

となり, MLE $\hat{\theta}_{ML}$ に一致する.

例 4.2　q として

$$q(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

とすれば, (4.3) の $p(x,\theta) = (1/\theta)q(x/\theta)$ は尺度母数 θ をもつ正規分布 $N(0,\theta^2)$ の p.d.f. になる. このとき,

$$\log q(x) = -\frac{1}{2}\log(2\pi) - \frac{x^2}{2}$$

より

$$(\log q(x))' = -x,\ (\log q(x))'' = -1,\ (\log q(x))''' = 0$$

となるから

$$z_1(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n x_i^2,\ z_2(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n x_i^2,\ z_3(\mathbf{x}) = 0$$

となる. z_1, z_2 は尺度不変でないので, (4.5) より θ の MPE は近似的に

$$\hat{\theta}_{MP}^r = \sqrt{M_2}\left(1 + \frac{r^2}{6M_2 n}\right) + o_p\left(\frac{1}{n}\right)$$

になる. ただし, $M_2 = (1/n)\sum_{i=1}^n X_i^2$ とする. 一方, θ の MLE は, $\hat{\theta}_{ML} = \sqrt{M_2}$ になる.

例 4.3　q として

$$q(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & (x > 0;\ \alpha > 0,\ \beta > 0), \\ 0 & (x \le 0) \end{cases}$$

とすれば, (4.3) の $p(x,\theta)=(1/\theta)q(x/\theta)$ は尺度母数 θ をもつガンマ分布の p.d.f. になる. このとき,

$$\log q(x)=\alpha\log\beta-\log\Gamma(\alpha)+(\alpha-1)\log x-\beta x \quad (x>0)$$

より

$$(\log q(x))'=\frac{\alpha-1}{x}-\beta,\ (\log q(x))''=\frac{-(\alpha-1)}{x^2},\ (\log q(x))'''=\frac{2(\alpha-1)}{x^3}$$

となるから

$$z_1(\mathbf{x})=-(\alpha-1)+\beta\bar{x},\ z_2(\mathbf{x})=\alpha-1,\ z_3(\mathbf{x})=-2(\alpha-1)$$

となる. z_1, z_2, z_3 は尺度不変でないので, (4.5) より θ の MPE は近似的に

$$\hat{\theta}^r_{MP}=\frac{\beta}{\alpha}\bar{X}+\frac{2\alpha r^2}{3n\beta\bar{X}}+o_p\left(\frac{1}{n}\right)$$

になる. 一方, θ の MLE は, $\hat{\theta}_{ML}=(\beta/\alpha)\bar{X}$ になる.

例 4.4　q として

$$q(x)=\frac{1}{2}\,e^{-|x|} \quad (-\infty<x<\infty)$$

とすれば, (4.3) の $p(x,\theta)=(1/\theta)q(x/\theta)$ は尺度母数 θ をもつ両側指数分布の p.d.f. になる. このとき,

$$\log q(x)=-\log 2-|x|$$

より

$$(\log q(x))'=-sgn x, \quad (\log q(x))''=(\log q(x))'''=0$$

となるから

$$z_1(\mathbf{x})=\overline{|x|}\ ,\ z_2(\mathbf{x})=z_3(\mathbf{x})=0$$

となる. ただし, $\overline{|x|}:=(1/n)\sum_{i=1}^n|x_i|$ とする. z_1 は尺度不変だから, (4.6) より θ の MPE は近似的に

$$\hat{\theta}^r_{MP}=\overline{|X|}+\frac{2r^2}{3n\overline{|X|}}+o_p\left(\frac{1}{n}\right)$$

になる. ただし, $\overline{|X|}:=(1/n)\sum_{i=1}^n|X_i|$ とする. 一方, θ の MLE は, $\hat{\theta}_{ML}=\overline{|X|}$ になる

注意 4.2　例4.4 における $q(x)$ は, $x=0$ において微分可能ではないので, q を $\mathbf{R}^1-\{0\}$ において考えればよい.

例4.5　q として

$$q(x)=e^{-x}(1+e^{-x})^{-2} \quad (-\infty<x<\infty)$$

とすれば, (4.3) の $p(x,\theta)=(1/\theta)q(x/\theta)$ は尺度母数 θ をもつロジスティック分布 (logistic distribution) の p.d.f. になる. ここで, $q(x)$ は $x=0$ に関して対称であることに注意. このとき,

$$\log q(x)=-x-2\log(1+e^{-x})$$

より

$$(\log q(x))'=-1+\frac{2}{e^x+1}, \quad (\log q(x))''=-\frac{2e^x}{(e^x+1)^2},$$
$$(\log q(x))'''=\frac{2e^x(e^x-1)}{(e^x+1)^3},$$

となるから

$$z_1(\mathbf{x})=\bar{x}-\frac{2}{n}\sum_{i=1}^n\frac{x_i}{e^{x_i}+1}, \quad z_2(\mathbf{x})=\frac{2}{n}\sum_{i=1}^n\frac{x_i^2e^{x_i}}{(e^{x_i}+1)^2},$$
$$z_3(\mathbf{x})=-\frac{2}{n}\sum_{i=1}^n\frac{x_i^3e^{x_i}(e^{x_i}-1)}{(e^{x_i}+1)^3}$$

となる. ただし, $\bar{x}=(1/n)\sum_{i=1}^n x_i$ とする. z_1, z_2, z_3 は尺度不変でないので, (4.5) より

$$\begin{aligned} d=\bar{x}-\frac{2}{n}\sum_{i=1}^n\frac{x_i}{e^{x_i/d}+1}-\frac{r^2}{3dn}+\frac{r^2}{dn}\Bigg\{&\frac{\bar{x}}{d}-\frac{2}{n}\sum_{i=1}^n\frac{x_i/d}{e^{x_i/d}+1} \\ &+\frac{2}{n}\sum_{i=1}^n\frac{(x_i/d)^2e^{x_i/d}}{(e^{x_i/d}+1)^2}-\frac{1}{3n}\sum_{i=1}^n\frac{(x_i/d)^3e^{x_i/d}(e^{x_i/d}-1)}{(e^{x_i/d}+1)^3}\Bigg\}+o_p\left(\frac{1}{n}\right) \\ =:\bar{x}-\frac{2}{n}\sum_{i=1}^n\frac{x_i}{e^{x_i/d}+1}-\frac{r^2}{3dn}&+\frac{r^2}{dn}\Delta\left(\frac{1}{d}\mathbf{x}\right)+o_p\left(\frac{1}{n}\right) \end{aligned} \tag{4.9}$$

になる. q は単峰形であるから (4.9) を満たす d は θ の MPE になる. また,

$$\theta=\bar{X}-\frac{2}{n}\sum_{i=1}^n\frac{X_i}{e^{X_i/\theta}+1}$$

の解 $\theta=\hat{\theta}_{ML}$ は θ の MLE になる. 一般に, (4.9) を満たす d を陽に求めることは難しいが, 次のような手順を用いれば数値的に近似を得ることができるであろう. まず,

$$x^*:=\bar{x}-\frac{2}{n}\sum_{i=1}^n\frac{x_i}{e^{x_i/\bar{x}}+1}$$

$$d_1 := \bar{x} - x^*$$

とおく. そして, n を固定して

$$d_{k+1} = \bar{x} - \frac{2}{n}\sum_{i=1}^{n} \frac{x_i}{e^{x_i/d_k}+1} - \frac{r^2}{3d_k n} + \frac{r^2}{d_k n}\Delta\left(\frac{1}{d_k}\mathbf{x}\right) + o_p\left(\frac{1}{n}\right) \quad (k = 1, 2, \cdots)$$

のように反復していけば, 適当な k について, d_k の値は θ の MPE の近似値と見なすことができるであろう.

上記の例 4.5 以外の結果および, 同様にして求めた他の分布の結果をまとめると, 表 4.1 になる.

表 4.1 $p(x,\theta) = (1/\theta)q(x/\theta)$ の尺度母数 θ に対する漸近的な MPE と MLE

分布	$q(x)$	$\hat{\theta}^r_{MP}$	$\hat{\theta}_{ML}$
指数分布	$e^{-x}\ (x>0)$	$\bar{X} + \frac{2r^2}{3n\bar{X}} + o_p\left(\frac{1}{n}\right)$	$\bar{X}$
正規分布	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\ (-\infty < x < \infty)$	$\sqrt{M_2}\left(1 + \frac{r^2}{6M_2 n}\right) + o_p\left(\frac{1}{n}\right)$ $(M_2 := (1/n)\sum_{i=1}^n X_i^2)$	$\sqrt{M_2}$
ガンマ分布	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\ (x>0;\alpha>0,\beta>0)$	$\frac{\beta}{\alpha}\bar{X} + \frac{2\alpha r^2}{3n\beta\bar{X}} + o_p\left(\frac{1}{n}\right)$	$\frac{\beta}{\alpha}\bar{X}$
両側指数分布	$\frac{1}{2}\,e^{-\lvert x\rvert}\ (-\infty < x < \infty)$	$\overline{\lvert X\rvert} + \frac{2r^2}{3n\overline{\lvert X\rvert}} + o_p\left(\frac{1}{n}\right)$ $(\overline{\lvert X\rvert} := (1/n)\sum_{i=1}^n \lvert X_i\rvert)$	$\overline{\lvert X\rvert}$
ワイブル分布	$\alpha x^{\alpha-1}e^{-x^\alpha}\ (x>0;\alpha>0)$	$\left(\overline{X^\alpha}\right)^{1/\alpha} + \frac{\alpha(\alpha+3)r^2}{6n\alpha\left(\overline{X^\alpha}\right)^{1/\alpha}} + o_p\left(\frac{1}{n}\right)$ $(\overline{X^\alpha} := (1/n)\sum_{i=1}^n X_i^\alpha)$	$\left(\bar{X^\alpha}\right)^{1/\alpha}$

5 おわりに

本論において, ある非正則分布族における位置母数の最大確率推定, また尺度母数をもつ分布族の最大確率推定, Bayes 推定等について論じて, 最良最大確率推定量 (best MPE) の近似的な良さを示した. しかし, 非正則推定問題を Bayes 的観点からもっとよく考える必要がある. 特に, 事前分布, 損失関数によって Bayes 推定量が異なってくるため, 他の推定量と比較する際に微妙な部分を残している.

参考文献

[A91] Akahira, M.(1991). The 3/2th and 2nd order asymptotic efficiency of maximum probability estimators in non-regular cases. *Ann. Inst. Statist. Math.*, **43**, 181–195.

[AO02] Akahira, M. and Ohyauchi, N.(2002). Information inequalities for the Bayes risk for a family of non-regular distributions. *Mathematical Research Note* 2002-001, Institute of Mathematics, University of Tsukuba. To appear in *Ann. Inst. Statist. Math.*

[AT81] Akahira, M. and Takeuchi, K.(1981). *Asymptotic Efficiency of Statistical Estimators: Concepts and Higher Order Asymptotic Efficiency.* Lecture Notes in Statistics 7, Springer, New York.

[B82] Blyth, C. R.(1982). Maximum probability estimation in small samples. *Festshrift for Erich Lehmann* (P. J. Bickel, K. A. Doksum, and J. L. Hodges, Jr., eds). Pacific Grove, CA: Wadsworth and Brooks/Cole.

[IH81] Ibragimov, I. A. and Has'minskii, R. Z.(1981). *Statistical Estimation: Asymptotic Theory.* Springer, New York.

[KKP01] Kotz, S, Kozubowski, T. J. and Podgórski, K. (2001). *The Laplace Distribution and Generalizations.* Birkhäuser, Boston.

[O01] Ohyauchi, N.(2001). Comparison of the Bayes risks of estimators for a family of truncated normal distributions. *Commun. Statist. -Theory and Meth.*, **31** (5), 699–718.

[OA01] Ohyauchi, N. and Akahira, M.(2001). On the lower bounds for the Bayes risk of estimators in the uniform and truncated normal cases. *Proc. Symps., Res. Inst. Math. Sci., Kyoto Univ.*, **1224**, 11–35.

[S96] Shono, H. (1996). Estimation of a location parameter of non-regular distributions. (In Japanese), Master's Thesis, University of Tsukuba.

[W86] Weiss, L.(1986). A note on small-sample maximum probability estimation. *Statistics & Probability Letters*, **4**, 109–111.

[WW74] Weiss, L. and Wolfowitz, J.(1974). *Maximum Probability Estimators and Related Topics.* Lecture Notes in Mathematics 424, Springer, Berlin.