

Asymptotic comparison of estimators for a family of truncated distributions

筑波大・数理物質 大谷内奈穂 (Nao Ohyauchi)

(Graduate School of Pure and Applied Sciences, University of Tsukuba)

筑波大・数理物質 赤平昌文 (Masafumi Akahira)

(Graduate School of Pure and Applied Sciences, University of Tsukuba)

1. はじめに

統計的推測の漸近理論において、通常は正則条件を仮定して理論を展開することが多いが、必ずしも正則条件が成り立たない場合も少なくないので、そのような場合に、非正則推定の構造を解明することは重要と考えられる ([AT95]).

本稿において、非正則分布族の典型として位置母数をもつ切断分布族において、その母数の推定問題を考える。従来、このような切断分布族の場合に Pitman 推定量の漸近展開を導出し、その漸近分散を求め、さらに切断点での密度の値による荷重推定量の漸近分布による比較も行われた ([AOT07]). 一方、同じ分布族で極値統計量は 1 次の漸近情報量損失は 0 になるが、2 次の漸近情報量損失は正值になることが分かり、さらに、極値統計量と漸近補助統計量の組から成る統計量の 2 次の漸近情報量損失が 0 となることも示されている ([AKO12]). 他方、位置共変推定の観点からは、極値統計量からつくられる荷重推定量は、2 次の漸近情報量損失が 0 になるのではないかという問題*が生じ得るが、第 2 節において、その問題に答える。次に、Pitman 推定量の漸近情報量損失を計算するためには、その漸近密度を必要とするが、[AOT07] における Pitman 推定量の漸近展開からそれを求めることは難しい。しかし、その漸近展開から Pitman 推定量の漸近分布関数を求めることはでき (第 3 節)、また、極値統計量から成る荷重推定量の漸近分布も求めることができるので、それらの推定量の漸近集中確率を求めて数値計算的観点から比較を試みる (第 4 節).

2. 荷重推定量の漸近分布と漸近情報量損失

まず、 $X_1, X_2, \dots, X_n, \dots$ をたがいに独立に、いずれも (Lebesgue 測度に関する) 密度 $f_0(x - \theta)$ をもつ実確率変数列とする。ただし、 $x \in \mathbf{R}^1, \theta \in \mathbf{R}^1$ とする。また、 $f_0(\cdot)$ に次の条件を仮定する。

(A1) $f_0(x) > 0$ ($a < x < b$); $f_0(x) = 0$ ($x \leq a, x \geq b$) で、 a, b は有限とする。

(A2) $f_0(\cdot)$ は开区間 (a, b) 上で連続微分可能とし

$$c_1 := \lim_{x \rightarrow a+0} f_0(x) = f_0(a+0) > 0, \quad c_2 := \lim_{x \rightarrow b-0} f_0(x) = f_0(b-0) > 0.$$

ここで、極値統計量を

$$\underline{\theta} := \max_{1 \leq i \leq n} X_i - b, \quad \bar{\theta} := \min_{1 \leq i \leq n} X_i - a$$

*この問題は 2012 年 7 月に開催された ims-APRM2012 における著者らの講演に対するコメントとして竹内啓先生によって提起された。

とおき,

$$\hat{\theta}^* := \lambda \bar{\theta} + (1 - \lambda) \underline{\theta} \quad (0 < \lambda < 1)$$

とすれば, $\hat{\theta}^*$ は θ の位置共変推定量になる. このとき $\hat{\theta}^*$ の漸近情報量損失を計算する. ここで用いる情報量として, 任意の $\theta_1, \theta_2 \in \Theta$ について $f_0(\cdot - \theta_1)$ と $f_0(\cdot - \theta_2)$ の間の X_1 の一般情報量を

$$I_{X_1}^{(\alpha)}(\theta_1, \theta_2) := -\frac{8}{1 - \alpha^2} \log \int_{-\infty}^{\infty} f_0(x - \theta_1)^{\frac{1-\alpha}{2}} f_0(x - \theta_2)^{\frac{1+\alpha}{2}} dx \quad (|\alpha| < 1)$$

によって定義し, 同様にして統計量 T の密度 $f_T(t, \theta)$ を用いて T の一般情報量 $I_T^{(\alpha)}(\theta_1, \theta_2)$ も定義する.

さて, $U := n(\bar{\theta} - \theta)$, $V := n(\underline{\theta} - \theta)$ とおくと $\hat{\theta}^*$ の定義より

$$T := n(\hat{\theta}^* - \theta) = \lambda n(\bar{\theta} - \theta) + (1 - \lambda)n(\underline{\theta} - \theta) = \lambda U + (1 - \lambda)V$$

となる. いま (U, V) の漸近密度は

$$f_{U,V}(u, v) = \begin{cases} c_1 c_2 e^{c_2 v - c_1 u} & (v < 0 < u) \\ 0 & (\text{その他}) \end{cases} \quad (2.1)$$

となるから, $t = \lambda u + (1 - \lambda)v$, $v = v$ とおいて変数変換すると

$$\frac{\partial(t, v)}{\partial(u, v)} = \begin{vmatrix} \lambda & 1 - \lambda \\ 0 & 1 \end{vmatrix} = \lambda$$

となることから, (V, T) の漸近密度は

$$\begin{aligned} f_{V,T}(v, t) &= \begin{cases} c_1 c_2 e^{c_2 v - \frac{c_1}{\lambda}(t - (1 - \lambda)v)} / \left| \frac{\partial(t, v)}{\partial(u, v)} \right| & (v < \min\{\frac{t}{1 - \lambda}, 0\}), \\ 0 & (\text{その他}) \end{cases} \\ &= \begin{cases} \frac{c_1 c_2}{\lambda} e^{(c_2 - \frac{c_1}{\lambda}(1 - \lambda))v - \frac{c_1}{\lambda}t} & (v < \min\{\frac{t}{1 - \lambda}, 0\}), \\ 0 & (\text{その他}) \end{cases} \end{aligned}$$

になる. よって, $n \rightarrow \infty$ のとき T の漸近周辺密度は

$$f_T(t) = \begin{cases} K e^{\frac{c_2}{1 - \lambda}t} + o(1) & (t < 0), \\ K e^{-\frac{c_1}{\lambda}t} + o(1) & (t \geq 0) \end{cases}$$

となる. ただし $K := c_1 c_2 / \{\lambda c_2 + (1 - \lambda)c_1\}$ とする. ここで, $\Delta > 0$ とすると

$$f_T(t - \Delta) = \begin{cases} K e^{\frac{c_2}{1 - \lambda}(t - \Delta)} + o(1) & (t < \Delta), \\ K e^{-\frac{c_1}{\lambda}(t - \Delta)} + o(1) & (t \geq \Delta) \end{cases}$$

となるから, $n \rightarrow \infty$ のとき

$$\begin{aligned} J_\alpha(n\Delta, \lambda) &:= \int_{-\infty}^{\infty} f_T(t)^{\frac{1-\alpha}{2}} f_T(t-n\Delta)^{\frac{1+\alpha}{2}} dt \\ &= \int_{-\infty}^0 K e^{\frac{c_2}{1-\lambda}t} e^{-\frac{c_2(1+\alpha)}{2(1-\lambda)}n\Delta} dt + \int_0^{n\Delta} K e^{-\frac{1}{2}\left\{\frac{c_1(1-\alpha)}{\lambda} - \frac{c_2(1+\alpha)}{1-\lambda}\right\}t} e^{-\frac{c_2(1+\alpha)}{2(1-\lambda)}n\Delta} dt \\ &\quad + \int_{n\Delta}^{\infty} K e^{-\frac{c_1}{\lambda}t} e^{\frac{(1+\alpha)c_1}{2\lambda}n\Delta} dt + o(1) \end{aligned}$$

になる. ここで

$$\lambda = \frac{(1-\alpha)c_1}{(1-\alpha)c_1 + (1+\alpha)c_2} =: \lambda_\alpha$$

とおくと

$$K = \frac{1}{2}\{(1-\alpha)c_1 + (1+\alpha)c_2\} =: K_\alpha$$

になる. このとき, $\Delta = O(1/n)$ とすれば, $n \rightarrow \infty$ のとき

$$\begin{aligned} J_\alpha(n\Delta, \lambda_\alpha) &= K_\alpha \left\{ e^{-K_\alpha n\Delta} \int_{-\infty}^0 e^{2K_\alpha t/(1+\alpha)} dt + e^{-K_\alpha n\Delta} \int_0^{n\Delta} dt \right. \\ &\quad \left. + e^{(1+\alpha)K_\alpha n\Delta/(1-\alpha)} \int_{n\Delta}^{\infty} e^{-2K_\alpha t/(1-\alpha)} dt \right\} + o(1) \\ &= K_\alpha \left(\frac{1}{K_\alpha} + n\Delta \right) e^{-K_\alpha n\Delta} + o(1) \\ &= (1 + K_\alpha n\Delta) e^{-K_\alpha n\Delta} + o(1) \end{aligned}$$

となる. よって, $n \rightarrow \infty$ のとき

$$\begin{aligned} I_{\hat{\theta}^*}^{(\alpha)}(\theta, \theta + \Delta) &= I_T^{(\alpha)}(\theta, \theta + n\Delta) = -\frac{8}{1-\alpha^2} \log J_\alpha(n\Delta, \lambda_\alpha) \\ &= \frac{1}{1-\alpha^2} \{2K_\alpha n\Delta - 8 \log(1 + K_\alpha n\Delta)\} + o(1) \end{aligned} \quad (2.2)$$

になる. 一方, $\mathbf{X} = (X_1, \dots, X_n)$ の一般情報量は

$$I_{\mathbf{X}}^{(\alpha)}(\theta, \theta + \Delta) = n I_{X_1}^{(\alpha)}(\theta, \theta + \Delta) = \frac{2K_\alpha n\Delta}{1-\alpha^2} + o(1) \quad (2.3)$$

となる. ゆえに, (2.2) と (2.3) から推定量 $\hat{\theta}^*$ の 1 次の漸近情報量損失は, $n \rightarrow \infty$ のとき

$$I_{\mathbf{X}}^{(\alpha)}(\theta, \theta + \Delta) - I_{\hat{\theta}^*}^{(\alpha)}(\theta, \theta + \Delta) = \frac{8}{1-\alpha^2} \log(1 + K_\alpha n\Delta) + o(1)$$

となり, $\hat{\theta}^*$ は 1 次の漸近情報量損失

$$\frac{8}{1-\alpha^2} \log(1 + K_\alpha n\Delta) (> 0)$$

をもつことが分かる. このことは, 1次元の推定量 $\hat{\theta}^*$ は元のデータ \mathbf{X} がもつ情報量を漸近的に把えきれないことも意味し, 第1節に述べた問題に対する否定的解答になる.

例1 $X_1, X_2, \dots, X_n, \dots$ をたがいに独立に, いずれも密度

$$f_0(x - \theta) = \begin{cases} ce^{-(x-\theta)} & (0 < x - \theta < 1), \\ 0 & (\text{その他}) \end{cases}$$

をもつ切断指数分布に従う確率変数列とする. ただし, $c = e/(e-1)$ とする. このとき, $X_{(1)} := \min_{1 \leq i \leq n} X_i$, $X_{(n)} := \max_{1 \leq i \leq n} X_i$ とすれば

$$\underline{\theta} = X_{(n)} - 1, \quad \bar{\theta} = X_{(1)}$$

となり,

$$\hat{\theta}^* = \lambda \bar{\theta} + (1 - \lambda) \underline{\theta} = \lambda X_{(1)} + (1 - \lambda)(X_{(n)} - 1) \quad (0 < \lambda < 1)$$

になる. ここで, 上記のことから

$$\lambda_\alpha = \frac{1 - \alpha}{1 - \alpha + (1 + \alpha)e^{-1}}$$

とすれば

$$K_\alpha = \frac{c}{2} \{1 - \alpha + (1 + \alpha)e^{-1}\}$$

になる. このとき, $\hat{\theta}^*$ の1次の漸近情報量損失は

$$\frac{8}{1 - \alpha^2} \log \left(1 + \frac{c}{2} (1 - \alpha + (1 + \alpha)e^{-1}) n \Delta \right)$$

になる. ただし $\Delta = O(1/n)$ とする. なお, この場合には統計量 $(\underline{\theta}, \bar{\theta})$ は2次の漸近情報量損失は無いことが示されている ([A96]).

3. Pitman 推定量とその漸近分布

一般に, X_1, \dots, X_n をたがいに独立にいずれも位置母数 $\theta (\in \mathbf{R}^1)$ をもつ密度 $f_0(x - \theta)$ をもつ分布に従う実確率変数と仮定する. このとき, $\mathbf{X} := (X_1, \dots, X_n)$ に基づく θ の推定量 $\hat{\theta} = \hat{\theta}(\mathbf{X}) = \hat{\theta}(X_1, \dots, X_n)$ について, 任意の $c \in \mathbf{R}^1$ に対して

$$\hat{\theta}(X_1 + c, \dots, X_n + c) = \hat{\theta}(X_1, \dots, X_n) + c$$

が成り立つとき, $\hat{\theta}$ を位置共変 (location equivariant) 推定量という. また, 位置共変推定量全体のクラスの中で平均2乗誤差を最小にする位置共変推定量を最良位置共変推定量といい, Pitman([P39]) は最良位置共変推定量は

$$\hat{\theta}_{PT} := \frac{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \theta \prod_{i=1}^n f_0(X_i - \theta) d\theta}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \prod_{i=1}^n f_0(X_i - \theta) d\theta}$$

となることを示したので、この推定量は (θ の) Pitman 推定量とも呼ばれている ([LC98]). さて、第2節の条件 (A1), (A2) の下では、 θ の Pitman 推定量 $\hat{\theta}_{PT}$ の漸近展開は

$$\begin{aligned} n(\hat{\theta}_{PT} - \theta) &= \frac{1}{2}(U + V) + \frac{(U - V)(e^{(c_1 - c_2)(U - V)} + 1)}{2(e^{(c_1 - c_2)(U - V)} - 1)} - \frac{1}{c_1 - c_2} + o_p(1) \\ &= \frac{(U - V)e^{(c_1 - c_2)(U - V)}}{e^{(c_1 - c_2)(U - V)} - 1} + V - \frac{1}{c_1 - c_2} + o_p(1) \end{aligned}$$

で与えられる ([AOT07]). この表現から分かるように Pitman 推定量はやや複雑で極値統計量からつくられる荷重推定量の形に漸近的にはなっていないことに注意. このとき、 $W = U - V$, $V = V$ とすれば、 $\hat{\theta}_{PT}$ の漸近累積分布関数 (asymptotic(as.))c.d.f. $F_{\hat{\theta}_{PT}}(t)$ は

$$\begin{aligned} F_{\hat{\theta}_{PT}}(t) &= \lim_{n \rightarrow \infty} P_\theta \left\{ n(\hat{\theta}_{PT} - \theta) \leq t \right\} = P_\theta \left\{ \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} + V - \frac{1}{c_1 - c_2} \leq t \right\} \\ &= E \left[P_\theta \left\{ V \leq t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} \mid W \right\} \right] \end{aligned} \quad (3.1)$$

で求められる. ここで、 (U, V) の漸近密度 (2.1) において、 $W = U - V$, $V = V$ として変数変換すると (W, V) の漸近密度は

$$f_{W,V}(w, v) = \begin{cases} c_1 c_2 e^{-(c_1 - c_2)v - c_1 w} & (-w < v < 0, w > 0), \\ 0 & (\text{その他}) \end{cases}$$

になる. よって、 $n \rightarrow \infty$ のとき W の漸近周辺密度は

$$f_W(w) = \begin{cases} \frac{c_1 c_2}{c_1 - c_2} (e^{-c_2 w} - e^{-c_1 w}) & (w > 0), \\ 0 & (w \leq 0) \end{cases}$$

となるから、 W を与えたときの V の漸近条件付密度関数は

$$f_{V|W}(v|w) = \frac{f_{W,V}(w, v)}{f_W(w)} = \begin{cases} \frac{(c_1 - c_2)e^{-(c_1 - c_2)v}}{e^{(c_1 - c_2)w} - 1} & (-w < v < 0), \\ 0 & (\text{その他}) \end{cases}$$

になり、

$$\begin{aligned} &P_\theta \left\{ V \leq t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} \mid W \right\} \\ &= \begin{cases} 0 & \left(t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} \leq -W \right), \\ \frac{1}{e^{(c_1 - c_2)W} - 1} \left\{ e^{(c_1 - c_2)W} - e^{-(c_1 - c_2) \left(t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} \right)} \right\} & \left(-W < t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} < 0 \right), \\ 1 & \left(0 \leq t + \frac{1}{c_1 - c_2} - \frac{We^{(c_1 - c_2)W}}{e^{(c_1 - c_2)W} - 1} \right) \end{cases} \end{aligned} \quad (3.2)$$

が得られる。ここで、 $c_1 > c_2$ とし、 $S := e^{(c_1-c_2)W}$ とすれば、

$$\left\{ w \mid t + \frac{1}{c_1 - c_2} - \frac{we^{(c_1-c_2)w}}{e^{(c_1-c_2)w} - 1} \leq -w \right\} = \begin{cases} \{s \mid 1 < s \leq s_0^{(t)}\} & \left(-\frac{1}{c_1-c_2} < t < 0\right), \\ \{s \mid s > 1\} & \left(t \leq -\frac{1}{c_1-c_2}\right), \\ \phi & (t \geq 0), \end{cases}$$

$$\left\{ w \mid -w < t + \frac{1}{c_1 - c_2} - \frac{we^{(c_1-c_2)w}}{e^{(c_1-c_2)w} - 1} < 0 \right\} = \begin{cases} \{s \mid s > s_1^{(t)}\} & (t > 0), \\ \{s \mid s > s_0^{(t)}\} & \left(-\frac{1}{c_1-c_2} < t \leq 0\right), \\ \phi & \left(t \leq -\frac{1}{c_1-c_2}\right), \end{cases}$$

$$\left\{ w \mid 0 \leq t + \frac{1}{c_1 - c_2} - \frac{we^{(c_1-c_2)w}}{e^{(c_1-c_2)w} - 1} \right\} = \begin{cases} \{s \mid 1 < s < s_1^{(t)}\} & (t > 0), \\ \phi & (t \leq 0) \end{cases}$$

を得る。ただし、 ϕ は空集合を表わし、 $s_0^{(t)} (> 1)$ は $\log s_0^{(t)} - ((c_1 - c_2)t + 1)(s_0^{(t)} - 1) = 0$ を、 $s_1^{(t)} (> 1)$ は $((c_1 - c_2)t + 1)(s_1^{(t)} - 1) - s_1^{(t)} \log s_1^{(t)} = 0$ を満たすものとする。よって、(3.1)、(3.2) より $\hat{\theta}_{PT}$ の as.c.d.f. として

$$F_{\hat{\theta}_{PT}}(t) = \begin{cases} 1 + \frac{c_2}{c_1 - c_2} (s_1^{(t)})^{-\frac{c_1}{c_1 - c_2}} - \frac{c_1 c_2}{(c_1 - c_2)^2} e^{-(c_1 - c_2)t - 1} I(s_1^{(t)}) & (t > 0), \\ \frac{c_1}{c_1 - c_2} (s_0^{(t)})^{-\frac{c_2}{c_1 - c_2}} - \frac{c_1 c_2}{(c_1 - c_2)^2} e^{-(c_1 - c_2)t - 1} I(s_0^{(t)}) & \left(-\frac{1}{c_1 - c_2} < t \leq 0\right), \\ 0 & \left(t \leq -\frac{1}{c_1 - c_2}\right) \end{cases} \quad (3.3)$$

が得られる。ただし、 $I(a) := \int_a^\infty s^{-\frac{c_1}{c_1 - c_2}} \cdot s^{\frac{1}{s-1}} ds$ とする。また、 $c_1 < c_2$ のときも上記と同様にすればよい。

4. 漸近集中確率による比較

第2節の荷重推定量 $\hat{\theta}^*$ において、特に $\lambda = c_1/(c_1 + c_2)$ として、 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ を集中確率の観点から比較を行う。まず、

$$T := n(\hat{\theta}^* - \theta) = \frac{c_1}{c_1 + c_2} U + \frac{c_2}{c_1 + c_2} V$$

となるから、 T の漸近周辺密度は

$$f_T(t) = \begin{cases} \frac{c_1 + c_2}{2} e^{(c_1 + c_2)t} + o(1) & (t < 0), \\ \frac{c_1 + c_2}{2} e^{-(c_1 + c_2)t} + o(1) & (t \geq 0) \end{cases}$$

となり、 T の as.c.d.f. $F_T(t)$ は

$$F_T(t) = \lim_{n \rightarrow \infty} P_\theta \left\{ n(\hat{\theta}^* - \theta) \leq t \right\} = \begin{cases} \frac{1}{2} e^{(c_1 + c_2)t} & (t \leq 0), \\ 1 - \frac{1}{2} e^{-(c_1 + c_2)t} & (t > 0) \end{cases} \quad (4.1)$$

になる. ここで, (4.1), (3.3) から $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の漸近集中確率 (asymptotic concentration probability (ACP)) $\lim_{n \rightarrow \infty} P_\theta \{n|\hat{\theta}^* - \theta| \leq a\}$, $\lim_{n \rightarrow \infty} P_\theta \{n|\hat{\theta}_{PT} - \theta| \leq a\}$ ($a > 0$) を求めたい. しかし前者は容易に計算できるが, 後者についてはそう簡単ではないので, 具体例において数値計算的観点から比較する.

例 2 (切断指数分布). X_1, \dots, X_n を互いに独立に, いずれも密度

$$f(x) = \begin{cases} \frac{e}{e-1} e^{-x} & (0 < x < 1), \\ 0 & (\text{その他}) \end{cases}$$

をもつ切断指数分布に従う確率変数とする. このとき, $c_1 = e/(e-1)$, $c_2 = 1/(e-1)$ より, (4.1), (3.3) から $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の as.c.d.f. を求めると, それぞれ

$$F_T(t) = \begin{cases} \frac{1}{2} e^{\frac{e+1}{e-1}t} & (t \leq 0), \\ 1 - \frac{1}{2} e^{-\frac{e+1}{e-1}t} & (t > 0) \end{cases}$$

$$F_{\hat{\theta}_{PT}}(t) = \begin{cases} 1 + \frac{1}{e-1} (s_1^{(t)})^{-\frac{e}{e-1}} - \frac{1}{(e-1)^2} e^{-t} I(s_1^{(t)}) & (t > 0), \\ \frac{e}{e-1} (s_0^{(t)})^{-\frac{1}{e-1}} - \frac{1}{(e-1)^2} e^{-t} I(s_0^{(t)}) & (-1 < t \leq 0), \\ 0 & (t \leq -1) \end{cases}$$

となる. ただし, $I(a) := \int_a^\infty s^{-\frac{e}{e-1}} \cdot s^{\frac{1}{e-1}} ds$ とする. このとき, $0 < a < 1$ に対して $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は, それぞれ

$$\text{ACP}_{\hat{\theta}^*}(a) := \lim_{n \rightarrow \infty} P_\theta \{n|\hat{\theta}^* - \theta| \leq a\} = F_T(a) - F_T(-a) = 1 - e^{-\frac{e+1}{e-1}a},$$

$$\text{ACP}_{\hat{\theta}_{PT}}(a) := \lim_{n \rightarrow \infty} P_\theta \{n|\hat{\theta}_{PT} - \theta| \leq a\} = F_{\hat{\theta}_{PT}}(a) - F_{\hat{\theta}_{PT}}(-a)$$

$$= \frac{1}{e-1} (s_1^{(a)})^{-\frac{e}{e-1}} - \frac{1}{(e-1)^2} e^{-a} I(s_1^{(a)})$$

$$- \frac{e}{e-1} (s_0^{(-a)})^{-\frac{1}{e-1}} + \frac{1}{(e-1)^2} e^a I(s_0^{(-a)}) + 1$$

で与えられ, $a = 0(0.1)1$ についての $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は表 1 のようになり, $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフは図 1 のようになる.

表 1 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP の値

a	0	0.1	0.2	0.3	0.4	0.5
$\text{ACP}_{\hat{\theta}^*}(a)$	0	0.194583	0.351304	0.477529	0.579193	0.661075
$\text{ACP}_{\hat{\theta}_{PT}}(a)$	0	0.185141	0.341385	0.474264	0.587527	0.683721
a	0.6	0.7	0.8	0.9	1	
$\text{ACP}_{\hat{\theta}^*}(a)$	0.727024	0.780141	0.822922	0.857378	0.88513	
$\text{ACP}_{\hat{\theta}_{PT}}(a)$	0.764536	0.830986	0.883492	0.921764	0.943809	

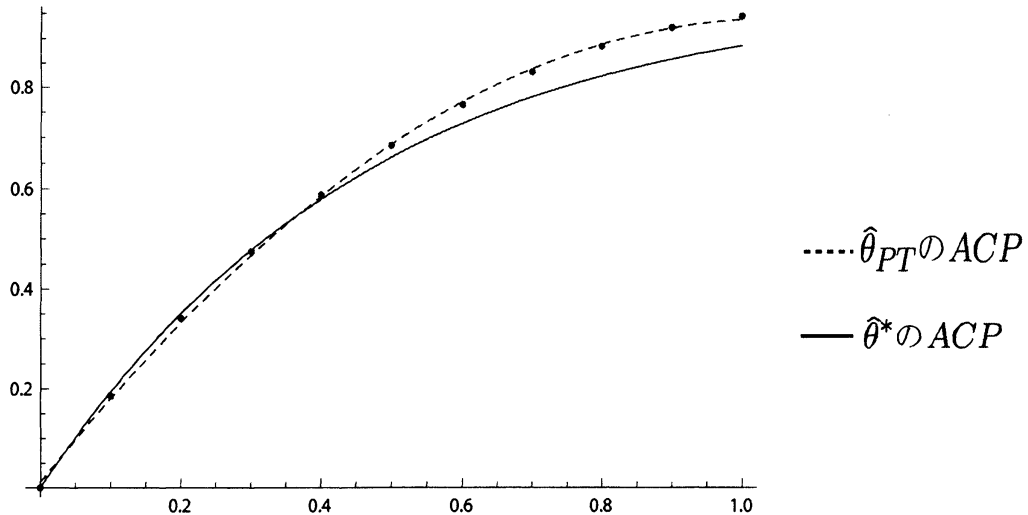


図1 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフ

グラフより, $0.37 < a < 1$ であれば, $\hat{\theta}_{PT}$ の方が $\hat{\theta}^*$ よりも ACP の意味で漸近的に良いがその他の a では悪いことが分かる.

例3 X_1, \dots, X_n を互いに独立に, いずれも密度

$$f(x) = \begin{cases} \frac{3}{2} - x & (0 < x < 1), \\ 0 & (\text{その他}) \end{cases}$$

をもつ分布に従う確率変数とする. このとき, $c_1 = 3/2, c_2 = 1/2$ より, (4.1), (3.3) から $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の as.c.d.f. はそれぞれ

$$F_T(t) = \begin{cases} \frac{1}{2}e^{2t} & (t \leq 0), \\ 1 - \frac{1}{2}e^{-2t} & (t > 0), \end{cases}$$

$$F_{\hat{\theta}_{PT}}(t) = \begin{cases} 1 + \frac{1}{2}(s_1^{(t)})^{-\frac{3}{2}} - \frac{3}{4}e^{-(t+1)}I(s_1^{(t)}) & (t > 0), \\ \frac{3}{2}(s_0^{(t)})^{-\frac{1}{2}} - \frac{3}{4}e^{-(t+1)}I(s_0^{(t)}) & (-1 < t \leq 0), \\ 0 & (t \leq -1) \end{cases}$$

となる. ただし, $I(a) := \int_a^\infty s^{-\frac{3}{2}} \cdot s^{\frac{1}{s-1}} ds$ とする. このとき, $0 < a < 1$ に対して $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は, それぞれ

$$\text{ACP}_{\hat{\theta}^*}(a) = F_T(a) - F_T(-a) = 1 - e^{-2a},$$

$$\begin{aligned} \text{ACP}_{\hat{\theta}_{PT}}(a) &= F_{\hat{\theta}_{PT}}(a) - F_{\hat{\theta}_{PT}}(-a) \\ &= \frac{1}{2}(s_1^{(a)})^{-\frac{3}{2}} - \frac{3}{4e^{-a}}I(s_1^{(a)}) \end{aligned}$$

$$-\frac{3}{2}(s_0^{(-a)})^{-\frac{1}{2}} + \frac{3}{4e}e^a I(s_0^{(-a)}) + 1$$

になる. $a = 0(0.1)1$ についての $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は表 2 のようになり, $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフは図 2 のようになる.

表 2 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP の値

a	0	0.1	0.2	0.3	0.4	0.5
ACP $_{\hat{\theta}^*}(a)$	0	0.181269	0.32968	0.451188	0.550671	0.632121
ACP $_{\hat{\theta}_{PT}}(a)$	0	0.17082	0.317893	0.445593	0.556861	0.653641
a	0.6	0.7	0.8	0.9	1	
ACP $_{\hat{\theta}^*}(a)$	0.698806	0.753403	0.798103	0.834701	0.864065	
ACP $_{\hat{\theta}_{PT}}(a)$	0.737120	0.807848	0.865726	0.909745	0.936236	

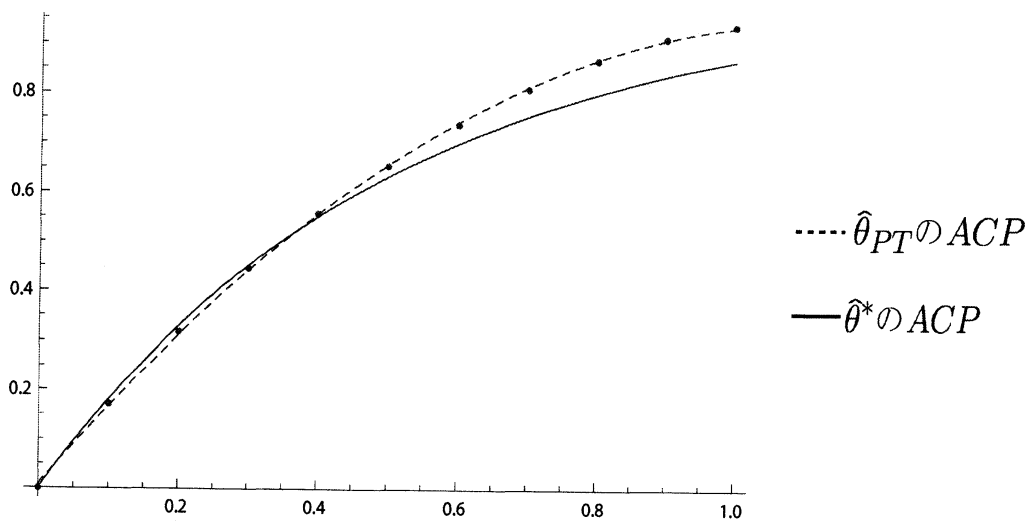


図 2 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフ

グラフより, $0.37 < a < 1$ であれば, $\hat{\theta}_{PT}$ の方が $\hat{\theta}^*$ よりも ACP の意味で漸近的に良いが, その他の a では悪いことが分かる.

例 4 (切断正規分布). X_1, \dots, X_n を互いに独立に, いずれも密度

$$f(x) = \begin{cases} ke^{-\frac{x^2}{2}} & (0 < x < 1), \\ 0 & (\text{その他}) \end{cases}$$

をもつ切断指数分布に従う確率変数とする. ただし, $k = 1/(\sqrt{2\pi}\{\Phi(1) - (1/2)\})$, Φ は標準正規分布の c.d.f. とする. このとき, $c_1 = k$, $c_2 = ke^{-\frac{1}{2}}$ より, (4.1), (3.3) から $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の as.c.d.f. を求めると, それぞれ

$$F_T(t) = \begin{cases} \frac{1}{2}e^{k(1+e^{-1/2})t} & (t \leq 0), \\ 1 - \frac{1}{2}e^{-k(1+e^{-1/2})t} & (t > 0) \end{cases}$$

$$F_{\hat{\theta}_{PT}}(t) = \begin{cases} 1 + \frac{1}{\sqrt{e}-1}(s_1^{(t)})^{-\frac{\sqrt{e}}{\sqrt{e}-1}} - \frac{1}{\sqrt{e}(\sqrt{e}-1)^2}e^{-k(1-e^{-1/2})t}I(s_1^{(t)}) & (t > 0), \\ \frac{\sqrt{e}}{\sqrt{e}-1}(s_0^{(t)})^{-\frac{1}{\sqrt{e}-1}} - \frac{1}{\sqrt{e}(\sqrt{e}-1)^2}e^{-k(1-e^{-1/2})t}I(s_0^{(t)}) & \left(-\frac{1}{k(1-e^{-1/2})} < t \leq 0\right), \\ 0 & (t \leq -1) \end{cases}$$

となる。ただし、 $I(a) := \int_a^\infty s^{-\frac{1}{1-e^{-1/2}}} \cdot s^{\frac{1}{s-1}} ds$ とする。このとき、 $0 < a < 1/\{k(1-e^{-1/2})\}$ に対して $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は、それぞれ

$$ACP_{\hat{\theta}^*}(a) := 1 - e^{-k(1+e^{-1/2})a},$$

$$ACP_{\hat{\theta}_{PT}}(a) := 1 + \frac{1}{\sqrt{e}-1}(s_1^{(a)})^{-\frac{\sqrt{e}}{\sqrt{e}-1}} - \frac{1}{\sqrt{e}(\sqrt{e}-1)^2}e^{-k(1-e^{-1/2})a}I(s_1^{(a)}) - \frac{\sqrt{e}}{\sqrt{e}-1}(s_0^{(a)})^{-\frac{1}{\sqrt{e}-1}} + \frac{1}{\sqrt{e}(\sqrt{e}-1)^2}e^{-k(1-e^{-1/2})a}I(s_0^{(a)})$$

で与えられ、 $a = 0(0.1)1/\{k(1-e^{-1/2})\}$ についての $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP は表3のようになり、 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフは図3のようになる。

表3 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP の値

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$ACP_{\hat{\theta}^*}(a)$	0	0.171187	0.31307	0.430663	0.528127	0.608905	0.675856	0.731345
$ACP_{\hat{\theta}_{PT}}(a)$	0	0.168558	0.309735	0.428186	0.527632	0.611091	0.681039	0.739526
a	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5
$ACP_{\hat{\theta}^*}(a)$	0.777336	0.815453	0.847045	0.873229	0.894931	0.912917	0.927825	0.94018
$ACP_{\hat{\theta}_{PT}}(a)$	0.78827	0.828722	0.862117	0.889513	0.911822	0.929833	0.944235	0.955627
a	1.6	1.7	1.8	1.9	2	2.1	$1/\{k(1-e^{-1/2})\}$	
$ACP_{\hat{\theta}^*}(a)$	0.950421	0.958908	0.965942	0.971773	0.976605	0.98061	0.983143	
$ACP_{\hat{\theta}_{PT}}(a)$	0.964534	0.971416	0.976676	0.980667	0.9837	0.986045	0.987489	

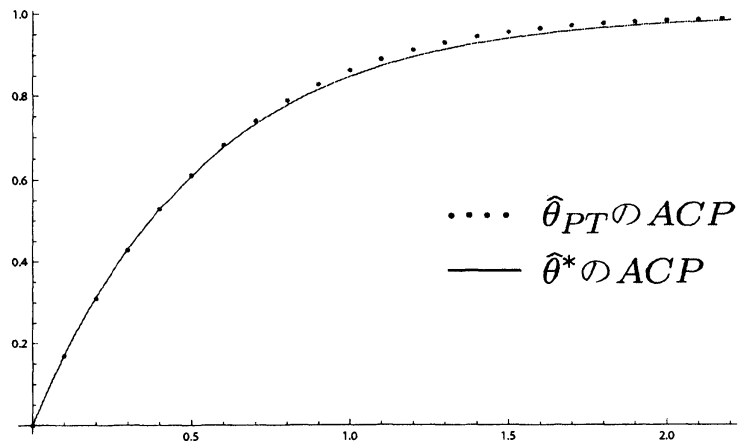


図3 $\hat{\theta}^*$ と $\hat{\theta}_{PT}$ の ACP のグラフ

グラフより, $0.5 < a < 1/\{k(1 - e^{-1/2})\}$ であれば, $\hat{\theta}_{PT}$ の方が $\hat{\theta}^*$ よりも ACP の意味で漸近的に良いがその他の a では悪いことが分かる.

5. おわりに

非正則分布族の典型として位置母数をもつ切断分布族を考え, 極値統計量からつくられる荷重推定量は位置共変推定量になるが, その漸近情報量損失は 0 にならないことを示した. また, Pitman 推定量と荷重推定量の漸近分布を求め, それらの漸近集中確率によって数値計算的見地から比較を試みたところ, Pitman 推定量が荷重推定量より ACP の意味で漸近的に一様に良いとは限らないことが分かった.

参考文献

- [A96] Akahira, M. (1996). Loss of information of a statistic for a family of non-regular distributions. *Ann. Inst. Statist. Math.*, **48**(2), 349–364.
- [AKO12] Akahira, M., Kim, H. G. and Ohyauchi, N. (2012). Loss of information of a statistic for a family of non-regular distributions, II: more general case. *Ann. Inst. Statist. Math.*, **64**(6), 1121–1138.
- [AOT07] Akahira, M., Ohyauchi, N. and Takeuchi, K. (2007). On the Pitman estimator for a family of non-regular distributions. *Metron*, **65**(1), 113–127.
- [AT95] Akahira, M. and Takeuchi, K. (1995). *Non-Regular Statistical Estimation*. Lecture Notes in Statistics 107, Springer, New York.
- [LC98] Lehmann, L. and Casella, G. (1998). *Theory of Point Estimation*. Springer, New York.