Asymptotic concentration probabilities of the Pitman estimator and weighted estimators in the non-regular case

Nao Ohyauchi
Institute of Mathematics, University of Tsukuba, Ibaraki, Japan, 305-8571
e-mail: naopu@math.tsukuba.ac.jp

Abstract

For a family of truncated distributions with a location parameter, the asymptotic expansion of the Pitman estimator and its asymptotic variance are given by Akahira, Ohyauchi and Takeuchi (2007). In this article, the asymptotic distributions of the Pitman estimator and weighted estimators are derived, and their asymptotic concentration probabilities are also computed. Further the asymptotic comparison of their estimators for some truncated distributions is discussed.

Key Words: Truncated distributions, Asymptotic expansion, Asymptotic distribution

1. Introduction

The asymptotic theory of statistical inference is usually discussed under suitable regularity conditions. But, it is interesting to consider the non-regular case when the regularity conditions do not hold, since such a case is sometimes appeared in a practical situation (Akahira and Takeuchi (1995)).

In the paper, we consider the estimation of a location parameter for a two-sided truncated family of distributions. In Akahira et al. (2007), the asymptotic expansion of the Pitman estimation and its asymptotic variance are obtained and its comparison with the weighted estimator formed from the extreme statistics by the values of the density at the truncated points is done. Here, we obtain the asymptotic distributions of the Pitman estimator and the weighted estimator, from which their asymptotic concentration probabilities are derived. From the numerical viewpoint, we compare them.

2. The asymptotic distributions of the weighted estimator and the Pitman estimator

Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of independently and identically distributed (i.i.d.) random variables with a density $f_0(x-\theta)$ w.r.t. the Lebesgue measure, where $x \in \mathbf{R}^1$ and $\theta \in \mathbf{R}^1$. We also assume the following conditions.
(A1) $f_0(x) > 0$ for $(a < x < b)$; $f_0(x) = 0$, otherwise, where a and b are finite.
(A2) $f_0(\cdot)$ is continuously differentiable in the interval (a, b),

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

Let $\underline{\theta} := \max_{1\le i\le n} X_i - b$, $\bar{\theta} := \min_{1\le i\le n} X_i - a$, $U := n(\bar{\theta}-\theta)$ and $V := n(\underline{\theta}-\theta)$. Then we consider as the weighted estimator

$$\hat{\theta}^* = \frac{c_1}{c_1+c_2}\bar{\theta} + \frac{c_2}{c_1+c_2}\underline{\theta}, \tag{1}$$

which yields $T := n(\hat{\theta}^* - \theta) = \{c_1/(c_1+c_2)\}U + \{c_2/(c_1+c_2)\}V$. On the other hand, the asymptotic expansion of the Pitman estimator is given by

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

(see Akahira et al. (2007)). As is seen from (2), it is noted that the Pitman is not a weighted estimator formed from the extremes statistics.

Now, letting $W = U - V$ and $V = V$, from (2) we have the asymptotic cumulative distribution function (as.c.d.f.) of the Pitman estimator $\hat{\theta}_{PT}$

$$F_{\hat{\theta}_{PT}}(t) = E\left[P\left\{V \leq t + \frac{1}{c_1-c_2} - \frac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1} \,\middle|\, W\right\}\right]. \tag{3}$$

Since the as. conditional density of V given W is

$$f_{V|W}(v|w) = \begin{cases} \dfrac{(c_1-c_2)e^{-(c_1-c_2)v}}{e^{(c_1-c_2)w}-1} & \text{for } -w < v < 0, \\ 0 & \text{otherwise,} \end{cases}$$

it follows that

$$\begin{aligned} &P_\theta\left\{V \leq t + \frac{1}{c_1-c_2} - \frac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1} \,\middle|\, W\right\} \\ &= \begin{cases} 0 & \text{for } t + \dfrac{1}{c_1-c_2} - \dfrac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1} \leq -W, \\ \dfrac{1}{e^{(c_1-c_2)W}-1}\left\{e^{(c_1-c_2)W} - e^{-(c_1-c_2)\left(t+\frac{1}{c_1-c_2}-\frac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1}\right)}\right\} & \\ & \text{for } -W < t + \dfrac{1}{c_1-c_2} - \dfrac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1} < 0, \\ 1 & \text{for } 0 \leq t + \dfrac{1}{c_1-c_2} - \dfrac{We^{(c_1-c_2)W}}{e^{(c_1-c_2)W}-1}. \end{cases} \end{aligned} \tag{4}$$

Let $c_1 > c_2$ and $S := e^{(c_1-c_2)W}$. Suppose that $s_0^{(t)} (> 1)$ and $s_1^{(t)}$ satisfy the following

$$\begin{aligned} &\log s_0^{(t)} - ((c_1-c_2)t+1)(s_0^{(t)}-1) = 0, \\ &((c_1-c_2)t+1)(s_1^{(t)}-1) - s_1^{(t)}\log s_1^{(t)} = 0, \end{aligned}$$

respectively. From (3) and (4) we have

$$F_{\hat{\theta}_{PT}}(t) = \begin{cases} 1 + \dfrac{c_2}{c_1-c_2}(s_1^{(t)})^{-\frac{c_1}{c_1-c_2}} - \dfrac{c_1c_2}{(c_1-c_2)^2}e^{-(c_1-c_2)t-1}I(s_1^{(t)}) & \text{for } t > 0, \\ \dfrac{c_1}{c_1-c_2}(s_0^{(t)})^{-\frac{c_2}{c_1-c_2}} - \dfrac{c_1c_2}{(c_1-c_2)^2}e^{-(c_1-c_2)t-1}I(s_0^{(t)}) & \text{for } -\frac{1}{c_1-c_2} < t \leq 0, \\ 0 & \text{for } t \leq -\frac{1}{c_1-c_2}, \end{cases} \tag{5}$$

where $I(a) := \int_a^\infty s^{-\frac{c_1}{c_1-c_2}} \cdot s^{\frac{1}{s-1}} ds$.

3. The comparison of the estimators by the asymptotic concentration probability

From the conditions (A1), (A2) and (1) we have the as.c.d.f. of $T = n(\hat{\theta}^* - \theta)$

$$F_T(t) = \lim_{n\to\infty} P_\theta \left\{ n(\hat{\theta}^* - \theta) \le t \right\} = \begin{cases} \frac{1}{2} e^{(c_1+c_2)t} & \text{for } t \le 0, \\ 1 - \frac{1}{2} e^{-(c_1+c_2)t} & \text{for } t > 0. \end{cases} \tag{6}$$

From (5) and (6) we can obtain the asymptotic concentration probabilities (ACPs) of $\hat{\theta}_{PT}$ and $\hat{\theta}^*$ around θ

$$\lim_{n\to\infty} P_\theta\{n|\hat{\theta}_{PT} - \theta| \le a\} \text{ and } \lim_{n\to\infty} P_\theta\{n|\hat{\theta}^* - \theta| \le a\}$$

for $a > 0$.

Example 1 (Truncated exponential distribution). Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables with a density $f(x) = e^{1-x}/(e-1)$ for $0 < x < 1$; $= 0$ otherwise. Then it follows from (5) and (6) that the ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$ are given by

$$\begin{aligned} \mathrm{ACP}_{\hat{\theta}^*}(a) &:= \lim_{n\to\infty} P_\theta \left\{ n|\hat{\theta}^* - \theta| \le a \right\} = 1 - e^{-\frac{e+1}{e-1}a}, \\ \mathrm{ACP}_{\hat{\theta}_{PT}}(a) &:= \lim_{n\to\infty} P_\theta \left\{ n|\hat{\theta}_{PT} - \theta| \le a \right\} \\ &= \frac{1}{e-1}(s_1^{(a)})^{-\frac{e}{e-1}} - \frac{1}{(e-1)^2} e^{-a} I(s_1^{(a)}) \\ &\quad - \frac{e}{e-1}(s_0^{(-a)})^{-\frac{1}{e-1}} + \frac{1}{(e-1)^2} e^{a} I(s_0^{(-a)}) + 1 \end{aligned}$$

for $0 < a < 1$, respectively.

Table 1 Numerical values of ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$

a	0	0.1	0.2	0.3	0.4	0.5
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0	0.194583	0.351304	0.477529	0.579193	0.661075
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	0	0.185141	0.341385	0.474264	0.587527	0.683721

a	0.6	0.7	0.8	0.9	1
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0.727024	0.780141	0.822922	0.857378	0.88513
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	0.764536	0.830986	0.883492	0.921764	0.943809

Example 2 Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables with a density $f(x) = (3/2) - x$ for $0 < x < 1$; $= 0$ otherwise. Then it follows from (5) and (6) that the ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$ are given by

$$\begin{aligned} \mathrm{ACP}_{\hat{\theta}^*}(a) &:= 1 - e^{-2a}, \\ \mathrm{ACP}_{\hat{\theta}_{PT}}(a) &:= \frac{1}{2}(s_1^{(a)})^{-3/2} - \frac{3}{4e} e^{-a} I(s_1^{(a)}) \\ &\quad - \frac{3}{2}(s_0^{(-a)})^{-1/2} + \frac{3}{4e} e^{a} I(s_0^{(-a)}) + 1 \end{aligned}$$

for $0 < a < 1$, respectively.

Table 2 Numerical values of ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$

a	0	0.1	0.2	0.3	0.4	0.5
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0	0.181269	0.32968	0.451188	0.550671	0.632121
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	0	0.17082	0.317893	0.445593	0.556861	0.653641

a	0.6	0.7	0.8	0.9	1
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0.698806	0.753403	0.798103	0.834701	0.864065
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	0.737120	0.807848	0.865726	0.909745	0.936236

Example 3 (Truncated normal distribution). Suppose that $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of i.i.d. random variables with a density $f(x) = ke^{-x^2/2}$ for $0 < x < 1$; $= 0$ otherwise, where $k = 1/(\sqrt{2\pi}\{\Phi(1) - (1/2)\})$ and Φ is c.d.f. of the standard normal distribution. Then it follows from (5) and (6) that the ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$ are given by

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

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

for $0 < a < 1/\{k(1 - e^{-1/2})\}$, respectively.

Table 3 Numerical values of ACPs of $\hat{\theta}^*$ and $\hat{\theta}_{PT}$

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0	0.171187	0.31307	0.430663	0.528127	0.608905	0.675856	0.731345
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	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
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0.777336	0.815453	0.847045	0.873229	0.894931	0.912917	0.927825	0.94018
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	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	$1/\{k(1-e^{-1/2})\}$
$\mathrm{ACP}_{\hat{\theta}^*}(a)$	0.950421	0.958908	0.965942	0.971773	0.976605	0.98061	0.983143
$\mathrm{ACP}_{\hat{\theta}_{PT}}(a)$	0.964534	0.971416	0.976676	0.980667	0.9837	0.986045	0.987489

As is seen from Tables 1, 2 and 3, it is seen that the Pitman estimator $\hat{\theta}_{PT}$ is not uniformly better than the weighted estimator $\hat{\theta}^*$ in the sense of ACP.

References

Akahira, M., Ohyauchi, N. and Takeuchi, K. (2007). On the Pitman estimator for a family of non-regular distributions. *Metron*, **65**(1), 113–127.

Akahira, M. and Takeuchi, K. (1995). *Non-Regular Statistical Estimation.* Lecture Notes in Statistics 107, Springer, New York.