

Sequential Point Estimation of Location Parameter in Location-Scale Family of Non-Regular Distributions

Ken-ichi Koike

Institute of Mathematics, University of Tsukuba, Tsukuba, Ibaraki, Japan

Abstract: In this paper, we consider sequential estimation of the location parameter based on the midrange in the presence of unknown scale parameter when the underlying distribution has a bounded support. The estimation is done under squared loss plus cost of sampling. Stopping rules based on the range are proposed and they are shown to be asymptotically efficient. The risks of the sequential procedures are compared with the Robbins' sequential estimation procedure based on the sample mean. The formers are shown to be asymptotically more efficient than the latter in the sense of the sample size when the density function changes sharply at the end points of the support. Koike (2007) observed a similar asymptotic superiority of the sequential estimation procedure based on the midrange in the sequential interval estimation procedure under the same condition.

Keywords: Extreme value; Non-regular case; Robbins' procedure; Sequential point estimation.

Subject Classifications: 62L12; 62F10.

1. INTRODUCTION

Suppose that X_1, X_2, \dots are independent and identically distributed (i.i.d.) with $E(X_1) = \mu$ and $V(X_1) = \sigma^2 > 0$. We consider the estimation problem of μ under the squared loss plus cost. If μ is estimated by the sample mean $\bar{X}_n = \sum_{i=1}^n X_i/n$, then the risk is given by

Address correspondence to Ken-ichi Koike, Institute of Mathematics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan; Fax: 81-29-853-6501; E-mail: koike@math.tsukuba.ac.jp

$$r'_n := E(\bar{X}_n - \mu)^2 + dn = \sigma^2/n + dn,$$

where $d(> 0)$ is the cost per observation. If σ is known, then the risk is minimized at the integer closest to $n'_d := \sigma/\sqrt{d}$. For simplicity, we will assume that n'_d is an integer. Then the minimized risk is $r'_{n'_d} = 2dn'_d = 2\sqrt{d}\sigma$. However, unless σ is known, one can not attain this risk with a non-sequential procedure. For normal random variables, Robbins (1959) proposed the following stopping rule:

$$T'_d := \{n \geq m'_d \mid n^2 \geq v_n/d\} \quad \left(v_n := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right),$$

where m'_d is the initial sample size. Ghosh and Mukhopadhyay (1979) and Chow and Yu (1981) showed that, under some conditions, the sequential estimation procedure $(T'_d, \bar{X}_{T'_d})$ is asymptotically risk efficient, that is, $r'_{T'_d}/r'_{n'_d} \rightarrow 1$ as $d \rightarrow 0$, without normality assumption. Chow and Martinsek (1982) showed that $(T'_d, \bar{X}_{T'_d})$ has bounded regret in the sense $(r'_{T'_d} - r'_{n'_d})/d = O(1)$. For another major reference, see Ghosh et al. (1997).

As a typical non-regular case, some sequential estimation procedures are obtained for the uniform distribution by Akahira and Koike (2005), Akahira and Takeuchi (2003), Chaturvedi et al. (2001), Govindarajulu (1997), Mukhopadhyay et al. (1983), Mukhopadhyay (1987) and Wald (1950) among others. Basawa et al. (1990) also discusses non-regular cases for the bounded risk point estimation under a general setting.

In Section 2, we consider sequential estimation of the location parameter based on the midrange in the presence of unknown scale parameter when the underlying distribution has a bounded support. The estimation is done under squared loss plus cost of sampling. In Subsection 2.1, we consider the case when the underlying density function has positive limit values at the end points of the support. A stopping rule based on the range is proposed and it is shown to be asymptotically efficient. The risks of the sequential procedures are compared with the Robbins' sequential estimation procedure based on the sample mean. The former is shown to be asymptotically more efficient than the latter in the sense of the sample size. In Subsection 2.2, we consider a case when the underlying density function converges to 0 at the end points of the support. We investigate a sequential estimation procedure based on the range, compare with the Robbins' sequential estimation procedure and show an asymptotic superiority of the estimation procedure

based on the midrange to the Robbins' procedure when the density function changes sharply at the end points of the support.

Koike (2007) observed a similar asymptotic superiority of the sequential estimation procedure based on the midrange in the sequential interval estimation procedure under the same condition.

2. SEQUENTIAL ESTIMATION PROCEDURES

In this section we consider sequential estimation procedures for two cases below.

Let Z_1, Z_2, \dots be a sequence of i.i.d. random variables according to the density function $f_0(x - \theta)$ ($\theta \in \mathbb{R}^1$) with respect to the Lebesgue measure. We assume throughout the paper that $f_0(x)$ has a bounded support $(-a, a)$ ($a > 0$), i.e., $f_0(x) > 0$ for $-a < x < a$, and $f_0(x) = 0$ otherwise, and is twice continuously differentiable in $(-a, a)$. Note that if the support of f_0 is $(-a, b)$ ($a \neq b$), then the normalized midrange does not converge to θ in probability as $n \rightarrow \infty$.

We consider the following two cases as non-regular distribution.

(A1) $f_0(x)$ satisfies

$$\begin{cases} \lim_{x \rightarrow -a+0} f_0(x) = c_1(> 0), & \lim_{x \rightarrow a-0} f_0(x) = c_2(> 0), \\ \lim_{x \rightarrow -a+0} f'_0(x) = h_1, & \lim_{x \rightarrow a-0} f'_0(x) = h_2, \end{cases}$$

where c_1, c_2, h_1 and h_2 are some constants.

(A2) $f_0(x)$ satisfies

$$\lim_{x \rightarrow -a+0} (x + a)^{-\gamma} f_0(x) = g_1, \quad \lim_{x \rightarrow a-0} (a - x)^{-\gamma} f_0(x) = g_2,$$

where γ, g_1 and g_2 are some positive constants.

In (A2), if the converging order γ 's are different, then the normalized midrange does not converge to θ in probability as $n \rightarrow \infty$. Note that $f_0(x)$ satisfying (A2) converges to 0 with the order of $(x + a)^\gamma$ and $|x - a|^\gamma$ as $x \rightarrow -a + 0$ and $x \rightarrow a - 0$, respectively. So, the density changes sharply at the end points of the support if $0 < \gamma < 1$ and changes smoothly if $\gamma > 1$. These conditions are essentially the same as those in Akahira (1975a, b), Akahira and Takeuchi (1981, p. 31; 1995, pp. 81, 148) and Koike (2007).

We consider the cases of (A1) and (A2) in Subsections 2.1 and 2.2, respectively.

2.1. Estimation Procedure for (A1)

In this subsection we treat the case of (A1). At first, we consider the asymptotic distribution of the extreme values in a similar way to Akahira and Takeuchi (1995) and Koike (2007).

Under (A1), by putting $Z_{(1)} := \min_{1 \leq i \leq n} Z_i$, $Z_{(n)} := \max_{1 \leq i \leq n} Z_i$, $U := n(Z_{(1)} + a - \theta)$ and $V := n(Z_{(n)} - a - \theta)$, the joint density of (U, V) is expanded as

$$f_{U,V}^{(n)}(u, v) = \begin{cases} \exp\{-(uc_1 - vc_2)\} \left[c_1 c_2 + \frac{1}{n} \left\{ -c_1 c_2 + c_1 c_2 \left(2(uc_1 - vc_2) - \left(\frac{h_1 u^2}{2} - \frac{h_2 v^2}{2} \right) - \frac{1}{2}(uc_1 - vc_2)^2 \right) + h_1 uc_2 + h_2 vc_1 \right\} \right] \\ \quad + o\left(\frac{1}{n}\right) & (v < 0 < u), \\ 0 & (\text{otherwise}) \end{cases} \quad (2.1)$$

from Koike (2007).

Now, we consider the location-scale parameter family of distributions with a bounded support $(\theta - \xi a, \theta + \xi a)$. Suppose that X_1, X_2, \dots is a sequence of i.i.d. random variables with the density $(1/\xi)f_0((x - \theta)/\xi)$, where $\theta \in \mathbb{R}$ and $\xi > 0$. Put $Y_i := (X_i - \theta)/\xi$ for each $i = 1, 2, \dots$, and $Y_{(1)} := \min_{1 \leq i \leq n} Y_i$, $Y_{(n)} := \max_{1 \leq i \leq n} Y_i$. Letting $S := n(Y_{(1)} + Y_{(n)})/2$ and $T := n(Y_{(1)} - Y_{(n)} + 2a)/2$, we have the asymptotic joint density of (S, T)

$$f_{S,T}(s, t) = \begin{cases} 2c_1 c_2 \exp\{-(c_1 - c_2)s - (c_1 + c_2)t\} & (t > |s|), \\ 0 & (\text{otherwise}) \end{cases}$$

from (2.1). The asymptotic density of S is given by

$$f_S(s) = \begin{cases} K e^{-2c_1 s} & (s \geq 0), \\ K e^{2c_2 s} & (s < 0), \end{cases}$$

where $K = 2c_1 c_2 / (c_1 + c_2)$. So, the asymptotic expectations of S and S^2 are

$$\begin{aligned} E(S) &\approx K \left\{ \int_0^\infty s e^{-2c_1 s} ds + \int_{-\infty}^0 s e^{2c_2 s} ds \right\} = \frac{c_2 - c_1}{2c_1 c_2}, \\ E(S^2) &\approx K \left\{ \int_0^\infty s^2 e^{-2c_1 s} ds + \int_{-\infty}^0 s^2 e^{2c_2 s} ds \right\} = \frac{c_2^2 - c_1 c_2 + c_1^2}{2(c_1 c_2)^2}. \end{aligned} \quad (2.2)$$

So, we may assume the following condition.

(B1) There exists a positive constant A satisfying $E(S^2) \rightarrow A$ as $n \rightarrow \infty$.

Concerning this assumption, we have the following lemma.

Lemma 2.1. *(B1) and $E(S^4) = O(1)$ hold under (A1).*

Proof. At first, we show (B1). From Fatou's lemma and (2.2),

$$\liminf_{n \rightarrow \infty} E(S^2) \geq \frac{c_2^2 - c_1 c_2 + c_1^2}{2(c_1 c_2)^2}.$$

Since $S = n(Y_{(1)} + Y_{(n)})/2$ and $0 \leq E\{(Y_{(1)} + Y_{(n)})^2\} \leq 2[E\{(Y_{(1)} + a)^2\} + E\{(Y_{(n)} - a)^2\}]$, it suffices to show $E\{(Y_{(1)} + a)^2\} = O(n^{-2})$ and $E\{(Y_{(n)} - a)^2\} = O(n^{-2})$. The density of $Y_{(1)}$ is given by

$$f_{Y_{(1)}}(x) = n\{1 - F(x)\}^{n-1} f_0(x),$$

where F is the distribution function of Y_1 , hence,

$$\begin{aligned} E\{(Y_{(1)} + a)^2\} &= \left(\int_{-a}^{-a+\epsilon} + \int_{-a+\epsilon}^a \right) (x + a)^2 n\{1 - F(x)\}^{n-1} f_0(x) dx \\ &=: I_1 + I_2 \quad (\text{say}). \end{aligned}$$

Putting $y = n(x + a)$, we have, for a sufficiently small $\epsilon > 0$,

$$\begin{aligned} I_1 &= \int_0^{n\epsilon} (n^{-1}y)^2 n\{1 - F(-a + n^{-1}y)\}^{n-1} f_0(-a + n^{-1}y) n^{-1} dy \\ &= \int_0^{n\epsilon} n^{-2} y^2 \exp(-cy) \left\{ 1 - \frac{h_1}{2} y^2 n^{-1} + o(n^{-1}) \right\} \{c + h_1 n^{-1} + o(n^{-1})\} dy \\ &= n^{-2} \int_0^{n\epsilon} y^2 \exp(-cy) \left\{ c + n^{-1} \left(-\frac{c h_1}{2} y^2 + h_1 y \right) + o(n^{-1}) \right\} dy \\ &\leq C n^{-2}, \end{aligned}$$

where the second equality follows from the expansion of $f_0(x)$, and C is some positive constant. On the other hand, since

$$\begin{aligned} I_2 &= \int_{-a+\epsilon}^a (x + a)^2 n\{1 - F(x)\}^{n-1} f_0(x) dx \\ &\leq n\{1 - F(-a + \epsilon)\}^{n-1} \int_{-a+\epsilon}^a (x + a)^2 n f_0(x) dx \\ &= n\{1 - F(-a + \epsilon)\}^{n-1} E\{(Y_1 + a)^2\} = O(n^{-2}), \end{aligned}$$

we have $E\{(Y_{(1)} + a)^2\} = O(n^{-2})$.

In a similar way to the above, we have $E\{(Y_{(n)} - a)^2\} = O(n^{-2})$, thus $E(S^2) = O(1)$. And we can show $E(S^4) = O(1)$ similarly. Therefore we have the desired result. \square

If the population distribution is the uniform distribution $U(-1, 1)$, then the constant A is calculated exactly. In fact, an easy computation yields

$$E(S^2) = \frac{2n^2}{(n+1)(n+2)} \rightarrow 2,$$

$$E(S^4) = \frac{24n^4}{(n+1)(n+2)(n+3)(n+4)} \rightarrow 24.$$

If θ is estimated by the midrange $M_n = (X_{(1)} + X_{(n)})/2$, then the risk is given by

$$r_n := E(M_n - \theta)^2 + dn,$$

where $d(> 0)$ is the cost per observation. From $S = n(M_n - \theta)/\xi$ and (A2), r_n is approximated by $(A\xi^2/n^2) + dn$, which is minimized at the integer closest to $n = n_d^{(1)} := (2A\xi^2/d)^{1/3}$ and the minimized value is $r_{n_d^{(1)}} = 3(A\xi^2 d^2)^{1/3}/2^{2/3}$. However, unless ξ is known, one can not attain this risk with a non-sequential procedure. Since the range $R_n := X_{(n)} - X_{(1)}$ converges to $2a\xi$ almost surely as $n \rightarrow \infty$, therefore we consider the following stopping rule:

$$T_d^{(1)} := \inf \left\{ n \geq m_d^{(1)} \mid n^3 \geq AR_n^2/(2a^2d) \right\},$$

where $m_d^{(1)}$ is the initial sample size with $d^{-l} \leq m_d^{(1)} = o(d^{-1/3})$ ($0 < l < 1/3$). Then we have the following theorem.

Theorem 2.1. *Under the conditions (A1) and (B1), as $d \rightarrow 0$, we have*

$$(i) \quad T_d^{(1)}/n_d^{(1)} \xrightarrow{\text{a.s.}} 1, \quad (ii) \quad E\left(T_d^{(1)}\right)/n_d^{(1)} \rightarrow 1, \quad (iii) \quad r_{T_d^{(1)}}/r_{n_d^{(1)}} \rightarrow 1.$$

Proof. At first, we note that

$$m_d^{(1)} \leq T_d^{(1)} \leq n_d^{(1)} + 1 \quad \text{with probability 1.} \quad (2.3)$$

In fact, since $0 \leq R_n \leq 2a\xi$ with probability 1, it holds $0 \leq (AR_n^2/(2a^2d))^{1/3} \leq (2A\xi^2/d)^{1/3}$ with probability 1. Hence, $n > (AR_n^2/(2a^2d))^{1/3}$ for n satisfying $n > (2A\xi^2/d)^{1/3}$. Therefore (2.3) holds. Since $T_d^{(1)} \xrightarrow{\text{a.s.}} \infty$ and $R_n \xrightarrow{\text{a.s.}} 2a\xi$, $R_{T_d^{(1)}} \xrightarrow{\text{a.s.}} 2a\xi$. By the definition of $T_d^{(1)}$,

$$\left(\frac{AR_{T_d^{(1)}}^2}{2a^2d} \right)^{1/3} \leq T_d^{(1)} < m_d^{(1)} + \left(\frac{AR_{T_d^{(1)}-1}^2}{2a^2d} \right)^{1/3}.$$

Dividing this by $n_d^{(1)}$, we have (i) as $d \rightarrow 0$ since $d^{-l} \leq m_d^{(1)} = o(d^{-1/3})$. To prove (ii), we have from (i) that

$$\liminf_{d \rightarrow 0} E \left(T_d^{(1)} / n_d^{(1)} \right) \geq 1.$$

by Fatou's lemma. On the other hand, by (2.3),

$$\frac{E \left(T_d^{(1)} \right)}{n_d^{(1)}} \leq \frac{(2A\xi^2/d)^{1/3} + 1}{(2A\xi^2/d)^{1/3}} \rightarrow 1 \quad (d \rightarrow 0),$$

hence $E \left(T_d^{(1)} \right) / n_d^{(1)} \rightarrow 1$ as $d \rightarrow 0$. So, we have (ii).

To prove (iii), we may assume $\theta = 0$ without loss of generality, since M_n is location equivariant. Putting $S_{k,n} := (k+n)M_{k+n} - kM_k$ ($k \geq 0, n \geq 1$), we have by Minkowski's inequality, that

$$\begin{aligned} 0 &\leq (E|S_{k,n}|^4)^{1/4} = (E|(k+n)M_{k+n} - kM_k|^4)^{1/4} \\ &\leq (E|(k+n)M_{k+n}|^4)^{1/4} + (E|kM_k|^4)^{1/4} = O(1) \end{aligned} \quad (2.4)$$

from Lemma 2.1. Taking η and λ satisfying $0 < \lambda < (A\xi^2)^{1/3} < \eta$, we have $P((d/2)^{1/3}T_d^{(1)} \geq \eta) \rightarrow 0$ as $d \rightarrow 0$ from (i). By (2.4) and Theorem B of Serfling (1980),

$$E \max_{1 \leq i \leq n} |S_{k,i}|^4 = O(1) \quad \text{for } k \geq k_0, n \geq 1. \quad (2.5)$$

Since $T_d^{(1)} \geq m_d^{(1)}$ with probability 1,

$$\begin{aligned} &\eta^{-2}(d/2)^{2/3} E \left\{ \left(T_d^{(1)} M_{T_d^{(1)}} \right)^2 I \left(\lambda \leq (d/2)^{1/3} T_d^{(1)} \leq \eta \right) \right\} \\ &\leq E \left(M_{T_d^{(1)}}^2 \right) \\ &\leq E \left\{ M_{T_d^{(1)}}^2 I \left(T_d^{(1)} \leq \lambda(2/d)^{1/3} \right) \right\} \\ &\quad + \lambda^{-2}(d/2)^{2/3} E \left\{ \left(T_d^{(1)} M_{T_d^{(1)}} \right)^2 I \left(\lambda \leq (d/2)^{1/3} T_d^{(1)} \leq \eta \right) \right\} \\ &\quad + E \left\{ M_{T_d^{(1)}}^2 I \left(T_d^{(1)} \geq \eta(2/d)^{1/3} \right) \right\}, \end{aligned} \quad (2.6)$$

where $I(A)$ is the indicator function of an event A . By Schwarz's inequality and (2.5),

$$\begin{aligned} &E \left\{ M_{T_d^{(1)}}^2 I \left(T_d^{(1)} \geq \eta(2/d)^{1/3} \right) \right\} \\ &\leq \eta^{-2}(d/2)^{2/3} \sum_{j=0}^{\infty} 2^{-2j} \left[E \left\{ \max_{2^j \eta(2/d)^{1/3} \leq n \leq 2^{j+1} \eta(2/d)^{1/3}} |nM_n|^4 \right\} \right]^{1/2} \end{aligned}$$

$$\begin{aligned}
& \cdot \left[P \left\{ 2^j \eta(2/d)^{1/3} \leq T_d^{(1)} \leq 2^{j+1} \eta(2/d)^{1/3} \right\} \right]^{1/2} \\
& = o \left(d^{2/3} \sum_{j=0}^{\infty} 2^{-2j} 2^j d^{-1/3} \right) = o(d^{1/3})
\end{aligned}$$

since $P \left(T_d^{(1)} \geq \eta(2/d)^{1/3} \right) \rightarrow 0$ as $d \rightarrow 0$. For an $\varepsilon > 0$ satisfying $\lambda^3 < (A\xi^2) - \varepsilon$,

$$\begin{aligned}
& P \left\{ T_d^{(1)} \leq \lambda(2/d)^{1/3} \right\} \\
& \leq P \left\{ \lambda(2/d)^{1/3} \geq \left(\frac{AR_n}{2a^2d} \right)^{1/3} \text{ for some } m_d^{(1)} \leq n \leq \lambda(2/d)^{1/3} \right\} \\
& = P \left\{ \lambda^3 \geq \frac{AR_n^2}{4a^2} \text{ for some } m_d^{(1)} \leq n \leq \lambda(2/d)^{1/3} \right\} \\
& \leq P \left\{ 1 - \left(\frac{R_n}{2a\xi} \right)^2 > \frac{\varepsilon}{A\xi^2} \text{ for some } m_d^{(1)} \leq n \leq \lambda(2/d)^{1/3} \right\} \\
& = P \left\{ 0 \leq \frac{R_n}{2a\xi} < \sqrt{1 - \frac{\varepsilon}{A\xi^2}} \text{ for some } m_d^{(1)} \leq n \leq \lambda(2/d)^{1/3} \right\} \\
& \leq P \left(0 \leq \frac{R_{m_d^{(1)}}}{2a\xi} < \sqrt{1 - \frac{\varepsilon}{A\xi^2}} \right) \quad (\text{by the monotonicity of } R_n \text{ w.r.t. } n) \\
& = O \left(\alpha^{m_d^{(1)}} \right), \tag{2.7}
\end{aligned}$$

where $\alpha \in (0, 1)$ is a constant. (2.7) follows from the estimation of the probability of the event $\{R_n \leq l\}$ ($l > 0$). In fact, putting $R'_n = Z_{(n)} - Z_{(1)}$, we have

$$\begin{aligned}
P(R'_n \leq l) &= P(Z_{(n)} - Z_{(1)} \leq l) \leq P \left\{ \{Z_{(1)} \geq -l/2\} \cup \{Z_{(n)} \leq l/2\} \right\} \\
&\leq P\{Z_{(1)} \geq -l/2\} + P\{Z_{(n)} \leq l/2\}.
\end{aligned}$$

Let G be the distribution function of Z_1 . Since $P\{Z_{(1)} \geq -l/2\} = \{1 - G(-l/2)\}^n$

$= \alpha_1^n$ (say) and $P\{Z_{(n)} \leq l/2\} = G^n(l/2) = \alpha_2^n$ (say), $P(R'_n \leq l) \leq \alpha_1^n + \alpha_2^n$, hence (2.7) holds. By Schwarz's inequality and (2.7),

$$\begin{aligned}
& E \left\{ M_{T_d^{(1)}}^2 I(T_d^{(1)} \leq \lambda(2/d)^{1/3}) \right\} \\
& \leq \left\{ E |M_{T_d^{(1)}}|^4 \right\}^{1/2} \left[P \left\{ T_d^{(1)} \leq \lambda(2/d)^{1/3} \right\} \right]^{1/2} \\
& \leq \sum_{j: 2^j \geq m_d^{(1)}} 2^{-2j} \left\{ E \left(\max_{2^j \leq n \leq 2^{j+1}} |nM_n|^4 \right) \right\}^{1/2} \left[P \left\{ T_d^{(1)} \leq \lambda(2/d)^{1/3} \right\} \right]^{1/2}
\end{aligned}$$

$$=D \sum_{j:2^j \geq m_d^{(1)}} 2^{-2j} \left(O \left(\alpha^{m_d^{(1)}} \right) \right)^{1/2} = O \left(m_d^{(1)-1} \alpha^{m_d^{(1)}/2} \right),$$

where D is some constant. On the other hand, since $|a^2 - b^2| \leq |a - b|^2 + 2|b||a - b|$ for $a, b \in \mathbb{R}$,

$$\begin{aligned} & \left| E \left\{ \left(T_d^{(1)} M_{T_d^{(1)}} \right)^2 I(\lambda \leq (d/2)^{1/3} T_d^{(1)} \leq \eta) \right\} \right. \\ & \quad \left. - E \left\{ ([\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]})^2 \right\} \right| \\ & \leq E \left\{ \max_{\lambda(2/d)^{1/3} \leq n \leq \eta(2/d)^{1/3}} \left| (n M_n)^2 - ([\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]})^2 \right| \right\} \\ & \quad + E \left[([\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]})^2 \right. \\ & \quad \cdot \left\{ I \left((d/2)^{1/3} T_d^{(1)} < \lambda \right) + I \left((d/2)^{1/3} T_d^{(1)} > \eta \right) \right\} \Big] \\ & \leq \left\{ E \left(\max_{\lambda(2/d)^{1/3} \leq n \leq \eta(2/d)^{1/3}} |n M_n - [\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]}|^4 \right) \right\}^{1/2} \\ & \quad + 2 \left[E \left\{ ([\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]})^2 \right\} \right]^{1/2} \\ & \quad \cdot \left\{ E \left(\max_{\lambda(2/d)^{1/3} \leq n \leq \eta(2/d)^{1/3}} |n M_n - [\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]}|^4 \right) \right\}^{1/4} \\ & \quad + \left\{ E |[\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]}|^4 \right\}^2 \\ & \quad \cdot \left\{ P^{1/2} \left((d/2)^{1/3} T_d^{(1)} < \lambda \right) + P^{1/2} \left((d/2)^{1/3} T_d^{(1)} > \eta \right) \right\} \end{aligned}$$

from Schwarz's inequality. Therefore, since $E \left([\lambda(2/d)^{1/3}] M_{[\lambda(2/d)^{1/3}]} \right)^2 \sim A\xi^2$ as $d \rightarrow 0$, and η and λ can be taken arbitrary close to $(A\xi^2)^{1/3}$,

$$E \left(\hat{\theta}_{T_d^{(1)}} - \theta \right)^2 \sim (A\xi^2)^{1/3} (d/2)^{2/3}. \quad (2.8)$$

By (ii) and (2.8), we have (iii). \square

Remark 2.1. The above proof of (iii) is basically based on that of Theorem 1 of Lai (1996), which shows that, in a general setting, a sequential estimation procedure based on a \sqrt{n} -consistent estimate is risk-efficient, whereas we treat a sequential estimation procedure based on an n -consistent estimate.

From Theorem 2.1 and Chow and Yu (1981), as $d \rightarrow 0$,

$$\frac{r_{T_d^{(1)}}}{r'_{T_d'}} \approx \frac{3(A\xi^2 d^2)^{1/3} / 2^{2/3}}{2\sqrt{d}\sigma} \rightarrow 0,$$

where $\sigma^2 = V(X_1)$. So, the estimation procedure $(T_d^{(1)}, M_{T_d^{(1)}})$ is asymptotically better than $(T'_d, \bar{X}_{T'_d})$. A similar phenomenon that the sequential interval estimation procedure based on the midrange is asymptotically better than the sample mean can be found in Koike (2007).

2.2. Estimation Procedure for (A2)

In this subsection, we consider the case of (A2).

By putting $U' := n^{1/(\gamma+1)}(Z_{(1)} + a - \theta)$ and $V' := n^{1/(\gamma+1)}(Z_{(n)} - a - \theta)$ with $Z_{(1)} = \min_{1 \leq i \leq n} Z_i$ and $Z_{(n)} = \max_{1 \leq i \leq n} Z_i$, under the condition (A2), the joint density $f_{U', V'}^{(n)}(u, v)$ of (U', V') satisfies

$$f_{U', V'}^{(n)}(u, v) \rightarrow \begin{cases} g_1 g_2 (-uv)^\gamma \exp\left\{-\frac{g_2}{\gamma+1}(-v)^{\gamma+1} - \frac{g_1}{\gamma+1}u^{\gamma+1}\right\} & (v < 0 < u), \\ 0 & (\text{otherwise}) \end{cases}$$

as $n \rightarrow \infty$ (Koike (2007)).

Suppose that X_1, X_2, \dots is a sequence of i.i.d. random variables with the density $(1/\xi)f_0((x - \theta)/\xi)$, where $\theta \in \mathbb{R}$ and $\xi > 0$. Put $Y_i := (X_i - \theta)/\xi$ for each $i = 1, 2, \dots$, and $Y_{(1)} := \min_{1 \leq i \leq n} Y_i$, $Y_{(n)} := \max_{1 \leq i \leq n} Y_i$.

Letting $S' := n^{1/(\gamma+1)}(Y_{(1)} + Y_{(n)})/2$ and $T' := n^{1/(\gamma+1)}(Y_{(1)} - Y_{(n)} + 2a)/2$, we have the asymptotic joint density of (S', T') in a same manner to Subsection 2.1. So, the asymptotic expectation of S'^2 is $E(S'^2)$ can be calculated. So, we may assume the following condition.

(B2) There exists a positive constant B satisfying $E(S'^2) \rightarrow B$ as $n \rightarrow \infty$.

Concerning this assumption, we have the following lemma.

Lemma 2.2. *(B2) and $E(S'^4) = O(1)$ hold under (A2).*

The proof is omitted since it is similar to the one of Lemma 2.1.

Under the condition (B2), as $n \rightarrow \infty$,

$$E(n^{2/(\gamma+1)}M_n^2) \rightarrow B\xi^2. \quad (2.9)$$

If θ is estimated by the midrange $M_n = (X_{(1)} + X_{(n)})/2$, then the risk is given by

$$r_n = E(M_n - \theta)^2 + dn,$$

where $d(> 0)$ is the cost per observation. From $S' = n^{1/(\gamma+1)}(M_n - \theta)/\xi$ and (2.9), r_n is approximated by $B\xi^2 n^{-2/(\gamma+1)} + dn$, which is minimized at the

integer closest to $n = n_d^{(2)} := \{2B\xi^2/(d(\gamma+1))\}^{(\gamma+1)/(\gamma+3)}$ and the minimized value is

$$r_{n_d^{(2)}} = B\xi^2 \left(\frac{d(\gamma+1)}{2B\xi^2} \right)^{2/(\gamma+3)} \left(\frac{\gamma+3}{\gamma+1} \right).$$

However, unless ξ is known, one can not attain this risk with a non-sequential procedure. Since the range $R_n = X_{(n)} - X_{(1)}$ converges to $2a\xi$ almost surely as $n \rightarrow \infty$, therefore we consider the following stopping rule:

$$T_d^{(2)} := \inf \left\{ n \geq m_d^{(2)} \mid n^{(\gamma+3)/(\gamma+1)} \geq BR_n^2/(2a^2d(\gamma+1)) \right\},$$

where $m_d^{(2)}$ is the initial sample size with $d^{-l} \leq m_d^{(2)} = o(d^{-(\gamma+1)/(\gamma+3)})$ ($0 < l < (\gamma+1)/(\gamma+3)$). Then we have the following theorem.

Theorem 2.2. *Under the conditions (A2) and (B2), as $d \rightarrow 0$, we have*

$$(i) \quad T_d^{(2)}/n_d^{(2)} \xrightarrow{\text{a.s.}} 1, \quad (ii) \quad E(T_d^{(2)})/n_d^{(2)} \rightarrow 1, \quad (iii) \quad r_{T_d^{(2)}}/r_{n_d^{(2)}} \rightarrow 1.$$

The proof is omitted since it is similar to the one of Theorem 2.1.

From Theorem 2.2 and Chow and Yu (1981), as $d \rightarrow 0$,

$$\frac{r_{T_d^{(2)}}}{r_{T_d'}} \approx \frac{\left(\frac{d(\gamma+1)}{2B\xi^2} \right)^{2/(\gamma+3)} \left(1 + \frac{2B\xi^2}{\gamma+1} \right)}{2\sqrt{d}\sigma} \rightarrow \begin{cases} 0 & (0 < \gamma < 1), \\ \text{constant} & (\gamma = 1), \\ \infty & (\gamma > 1), \end{cases}$$

where $\sigma^2 = V(X_1)$. So, the estimation procedure $(T_d^{(2)}, M_{T_d^{(2)}})$ is asymptotically better than $(T_d', \bar{X}_{T_d'})$ for $0 < \gamma < 1$, and worse for $\gamma > 1$. In other words, $(T_d^{(2)}, M_{T_d^{(2)}})$ is asymptotically superior to $(T_d', \bar{X}_{T_d'})$ if the density changes sharply at the end points of the support. Koike (2007) observed a similar asymptotic superiority of the sequential estimation procedure based on the midrange in the sequential interval estimation procedure for θ under the same assumptions when the density changes steeply at the end points of the support. Note that similar results for the location family in the non-sequential case can be found in Akahira (1975a) and Akahira and Takeuchi (1981, 1995).

ACKNOWLEDGMENTS

The author would like to thank Professor Masafumi Akahira of University of Tsukuba for his valuable comments throughout the preparation of the paper.

The author also thanks the Editor and the referee for helpful comments which have led to improvements to the original version of the manuscript. This research was partly supported by the Grant-in-Aid for Scientific Research (C) 17540101, Japan Society for the Promotion of Science.

REFERENCES

- Akahira, M. (1975a). Asymptotic Theory for Estimation of Location in Non-Regular Cases, I: Order of Convergence of Consistent Estimators, *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers* 22: 8-26.
- Akahira, M. (1975b). Asymptotic Theory for Estimation of Location in Non-Regular Cases, II: Bounds of Asymptotic Distributions of Consistent Estimators, *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers* 22: 99-115.
- Akahira, M. and Koike, K. (2005). Sequential Interval Estimation of a Location Parameter with the Fixed Width in the Uniform Distribution with an Unknown Scale Parameter, *Sequential Analysis* 24: 63-75.
- Akahira, M. and Takeuchi, K. (1981). *Asymptotic Efficiency of Statistical Estimators: Concepts and Higher Order Asymptotic Efficiency*, Lecture Notes in Statistics 7, New York: Springer-Verlag.
- Akahira, M. and Takeuchi, K. (1995). *Non-Regular Statistical Estimation*, Lecture Notes in Statistics 107, New York: Springer-Verlag.
- Akahira, M. and Takeuchi, K. (2003). The Information Inequality in Sequential Estimation for the Uniform Case, *Sequential Analysis* 22: 223-232.
- Basawa, I. V., McCormick, W. P., and Sriram, T. N. (1990). Sequential Estimation for Dependent Observations with an Application to Non-Standard Autoregressive Processes, *Stochastic Processes and Their Applications* 35: 149-168.
- Chaturvedi, A., Surinder, K., and Sanjeev, K. (2001). Multi-Stage Estimation Procedures for the “Range” of Two-Parameter Uniform Distribution, *Metron* 59: 179-186.
- Chow, Y. S. and Martinsek, A. T. (1982). Bounded Regret of a Sequential Procedure for Estimation of the Mean, *Annals of Statistics* 10: 909-914.
- Chow, Y. S. and Yu, K. F. (1981). On the Performance of Sequential Procedure for the Estimation of the Mean, *Annals of Statistics* 9: 184-189.
- Ghosh, M. and Mukhopadhyay, N. (1979). Sequential Point Estimation of the Mean When the Distribution is Unspecified, *Communications in Statistics-Theory & Methods* 8: 637-652.

- Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*, New York: Wiley.
- Govindarajulu, Z. (1997). A Note on Two-Stage and Sequential Fixed-Width Intervals for the Parameter in the Uniform Density, *Statistics & Probability Letters* 36: 179-188. Erratum: *Statistics & Probability Letters* 42 (1999): 213-215.
- Koike, K. (2007). Sequential Interval Estimation of a Location Parameter with the Fixed Width in the Non-Regular Case, *Sequential Analysis* 26: 63-70.
- Lai, T. L. (1996). On Uniform Integrability and Asymptotically Risk-Efficient Sequential Estimation, *Sequential Analysis* 15: 237-251.
- Mukhopadhyay, N. (1987). A Note on Estimating the Range of a Uniform Distribution, *South African Statistical Journal* 21: 27-38.
- Mukhopadhyay, N., Ghosh, M., Hamdy, H. I., and Wackerly, D. D. (1983). Sequential and Two-Stage Point Estimation for the Range in a Power Family Distribution, *Sequential Analysis* 2: 259-288.
- Robbins, H. (1959). Sequential Estimation of the Mean of a Normal Population, in *Probability and Statistics (Harold Cramér Volume)*, U. Grenander, ed., pp. 235-245, Stockholm: Almqvist and Wiksell.
- Serfling, R. J. (1970). Moment Inequalities for the Maximum Cumulative Sum, *Annals of Mathematical Statistics* 41: 1227-1234.
- Wald, A. (1950). *Statistical Decision Function*, New York: Wiley.