Metrika (2000) 52: 163–171

Metrika

Sequential point estimation of normal mean under LINEX loss function

Yoshikazu Takada

Department of Computer Science, Kumamoto University, Kumamoto 860-8555, Japan
(e-mail: takada@kumamoto-u.ac.jp)

Received: December 1999

Abstract. A sequential point estimation of the mean of a normal distribution is considered under LINEX loss function. The regret of sequential procedures are obtained. Furthermore, it is shown that a sequential procedure with the sample mean as an estimate is asymptotically inadmissible. An accerelated stopping time is also considered.

Key words: LINEX loss function, sequential estimation, accelerated stopping time, inadmissibility.

1 Introduction

There are some cases for which it is appropriate to use asymmetric loss functions. In this paper we consider LINEX (Linear Exponential) loss function proposed by Varian (1975). Zellner (1986) showed that the sample mean is inadmissible for estimating the mean of a normal distribution with known variance under the loss function. Furthermore, he showed that the inadmissibility holds even though the variance is unknown.

This paper considers a sequential point estimation of the mean of a normal distribution under LINEX loss function. In Section 2 we give an asymptotic expansion of the regret of the proposed sequential procedure. It is shown that the inadmissibility of the sample mean also holds asymptotically for the case of the sequential estimation. In Section 3 we consider an accelerated stopping time, which was first proposed by Hall (1983) to save sampling operations. Mukhopadhyay and Solanky (1991) studied a general accelerated stopping time. We get results similar to those in the previous section.

Recently, Chattopadhyay (1998) considered the sequential estimation under LINEX loss function from the point of a bounded risk problem and obtained the analogous result.

2 Sequential estimation

Let $X_1, X_2, \ldots$ be a sequence of independent observations from a normal population with unknown mean μ and unknown variance σ^2. We consider estimating the mean under LINEX loss function

$$L(d,\mu) = \exp[a(d-\mu)] - a(d-\mu) - 1 \tag{2.1}$$

and cost $c(>0)$ for each observation, where $a \neq 0$.

If σ^2 were known, we would estimate μ by

$$\delta_n = \overline{X}_n - \frac{a\sigma^2}{2n} \tag{2.2}$$

instead of the sample mean $\overline{X}_n$ for a sample of fixed size n. This is due to the fact that the sample mean is dominated by δ_n under LINEX loss function and δ_n is admissible and minimax. See Zellner (1986) and Rojo (1987) for the details. Then the risk would be

$$\begin{aligned} R_n &= E_\theta\{L(\delta_n,\mu) + cn\} \\ &= \frac{a^2\sigma^2}{2n} + cn, \end{aligned}$$

where $\theta = (\mu, \sigma^2)$. The risk would be minimized at $n_c = \sqrt{\frac{a^2\sigma^2}{2c}}$ with $R_{n_c} = 2cn_c$. Unfortunately, σ^2 is unknown and hence the best fixed sample size procedure can not be used.

Motivated by the formula for n_c, we propose the following stopping time

$$T_c = \inf\{n \geq m; n > b\ell_n\hat{\sigma}_n\}, \tag{2.3}$$

where $m \geq 2$ is the initial sample size, $b = \sqrt{\frac{a^2}{2c}}$, $\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$, and $\{\ell_n\}$ is a sequence of constants such that

$$\ell_n = 1 + \frac{1}{n}\ell_0 + o\left(\frac{1}{n}\right) \quad \text{as } n \to \infty. \tag{2.4}$$

First we consider a sequential procedure with the stopping time T_c that estimates μ by

$$\hat{\delta}_{T_c} = \overline{X}_{T_c} - \frac{a\hat{\sigma}_{T_c}^2}{2T_c}. \tag{2.5}$$

Then the risk of the procedure is

$$R_{T_c} = E_\theta\{L(\hat{\delta}_{T_c},\mu) + cT_c\} \tag{2.6}$$

and the regret is $R_{T_c} - R_{n_c}$.

In the subsequent discussion, we need some results provided in Woodroofe (1982). It is easy to see that the stopping time T_c of (2.3) assumes the form (10.1) with $\alpha = 1/2$ in Woodroofe (1982). Hence Theorem 10.1 and Table 10.1 in Woodroofe (1982) leads to the following result.

Theorem 1. *If* $m \geq 3$, *then*

$$E_\theta(T_c) = n_c + v - \tfrac{3}{4} + \ell_o + o(1), \quad \text{as } c \to 0$$

with $v = 0.633$.

Since the event $\{T_c = n\}$ is independent of $\overline{X}_n$,

$$E_\theta\{\exp[a(\hat{\delta}_{T_c} - \mu)]\} = E_\theta\left\{\exp\left[-\frac{a^2}{2T_c}(\hat{\sigma}^2_{T_c} - \sigma^2)\right]\right\}$$

and

$$E_\theta\{\hat{\delta}_{T_c} - \mu\} = -aE_\theta\left\{\frac{\hat{\sigma}^2_{T_c}}{2T_c}\right\}.$$

So from (2.1) and (2.6),

$$R_{T_c} = E_\theta\left\{\exp\left[-\frac{a^2}{2T_c}(\hat{\sigma}^2_{T_c} - \sigma^2)\right] + \frac{a^2\hat{\sigma}^2_{T_c}}{2T_c} - 1 + cT_c\right\}. \tag{2.7}$$

Letting

$$\phi(x) = \frac{1}{x} + x, \quad x > 0, \tag{2.8}$$

we can get the following result along the proof of Theorem 10.2 in Woodroofe (1982).

Lemma 1. *Let* $0 < \varepsilon < 1$ *and let* N_c *be any stopping time such that*

(i) $\lim_{c\to 0} n_c^2 P(N_c \leq \varepsilon n_c) = 0$,

(ii) $\lim_{c\to 0} \frac{E(N_c - n_c)^2}{n_c} = \tau^2$.

Then

$$\lim_{c\to 0} n_c E\left\{\phi\left(\frac{N_c}{n_c}\right) - \phi(1)\right\} = \tau^2$$

Now we proceed to derive an asymptotic expansion of the regret as $c \to 0$. Let $C = \{T_c > \varepsilon n_c\} \cap \{|\hat{\sigma}^2_{T_c} - \sigma^2| \leq \delta\}$ for some $0 < \varepsilon < 1$ and $\delta > 0$. Let C' be the compliment of C.

Lemma 2.

$$P_\theta(C') = O(c^{(m-1)/2}) \quad \text{as } c \to 0.$$

Proof. Note that

$$P_\theta(C') \le P_\theta(T_c \le \varepsilon n_c) + P_\theta\left(\sup_{n>\varepsilon n_c} |\hat{\sigma}_n^2 - \sigma^2| > \delta\right).$$

It follows from Lemma 10.3 in Woodroofe (1982) that

$$P_\theta(T_c \le \varepsilon n_c) = O(c^{(m-1)/2}). \tag{2.9}$$

Hence it is enough to show that

$$P_\theta\left(\sup_{n>\varepsilon n_c} |\hat{\sigma}_n^2 - \sigma^2| > \delta\right) = O(c^{(m-1)/2}),$$

which follows from the reverse submartingale property of $\{|\hat{\sigma}_n^2 - \sigma^2|^q\}$ with $q = 2(m-1)$ and

$$E|\hat{\sigma}_{\varepsilon n_c}^2 - \sigma^2|^q = O(n_c^{-(m-1)}).$$

So the proof is completed.

Rewrite (2.7) as

$$R_{T_c} = E_\theta\left\{\exp\left[-\frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2)\right] + \frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2) - 1\right\} + E_\theta\left\{\frac{a^2\sigma^2}{2T_c} + cT_c\right\}$$
$$= I + II \quad \text{(say)}. \tag{2.10}$$

Lemma 3. *If $m \ge 7$, then*

$$I/c = O(c^{1/2}) \quad \text{as } c \to 0.$$

Proof. Devide I into two parts such that

$$I = \int_C \left\{\exp\left[-\frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2)\right] + \frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2) - 1\right\} dP_\theta$$
$$+ \int_{C'} \left\{\exp\left[-\frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2)\right] + \frac{a^2}{2T_c}(\hat{\sigma}_{T_c}^2 - \sigma^2) - 1\right\} dP_\theta$$
$$= I_1 + I_2 \quad \text{(say)}.$$

The Taylor expansion of e^x shows

$$I_1 = \int_C \frac{1}{2}\left(\frac{a^2}{2T_c}\right)^2 (\hat{\sigma}_{T_c}^2 - \sigma^2)^2 \exp(\Delta)\, dP_\theta$$
$$= \frac{c^2 n_c}{2\sigma^4} \int_C \left(\frac{n_c}{T_c}\right)^3 T_c(\hat{\sigma}_{T_c}^2 - \sigma^2)^2 \exp(\Delta)\, dP_\theta,$$

where

$$|\Delta| \leq \frac{a^2}{2T_c}|\hat{\sigma}^2_{T_c} - \sigma^2| \leq \frac{a^2\delta}{2m}$$

on C. It is easy to see that the distribution of $\left(\frac{n_c}{T_c}\right)^3 T_c\left(\hat{\sigma}^2_{T_c} - \sigma^2\right)^2 \exp(\Delta_t)$ converges to that of $2\sigma^4\chi^2_1$ as $c \to 0$ and is uniformly integrable on C, where χ^2_1 denotes the chi-squared random variable with one degree of freedom. Hence we have

$$I_1 = \frac{c^2 n_c}{2\sigma^4}(2\sigma^4 + o(1))$$
$$= O(c^{3/2})$$

So it is enough to show that if $m \geq 7$, then

$$I_2 = O(c^{3/2}) \quad \text{as } c \to o. \tag{2.11}$$

Since $T_c \geq m$, we have

$$0 \leq I_2 \leq \exp\left(\frac{a^2\sigma^2}{2m}\right)P_\theta(C') + \frac{a^2}{2m}\int_{C'}|\hat{\sigma}^2_{T_c} - \sigma^2|\, dP_\theta$$
$$\leq \exp\left(\frac{a^2\sigma^2}{2m}\right)P_\theta(C') + \frac{a^2}{2m}K^{1/2}P_\theta(C')^{1/2},$$

where

$$K = E(\hat{\sigma}^2_{T_c} - \sigma^2)^2,$$

which is finite from the reverse submartingale property of $\{(\hat{\sigma}^2_n - \sigma^2)^2\}$. Hence (2.11) follows from Lemma 2.

Lemma 4. *If $m \geq 4$, then*

$$\frac{1}{c}(II - 2cn_c) = \frac{1}{2} + o(1) \quad \textit{as } c \to 0.$$

Proof. Note that

$$II - 2cn_c = E_\theta\left\{\frac{n_c^2 c}{T_c} + cT_c - 2cn_c\right\}$$
$$= cn_c E_\theta\left\{\phi\left(\frac{T_c}{n_c}\right) - \phi(1)\right\},$$

where ϕ is (2.8). It follows from Theorem 10.1 in Woodroofe (1982) that if $m \geq 3$,

$$\lim_{c\to 0}\frac{E_\theta(T_c-n_c)^2}{n_c}=\frac{1}{2}.$$

So from (2.9) the conditions of Lemma 1 are satsified with $\tau^2 = 1/2$ if $m \geq 4$. Hence the proof is completed.

Now the asymptotic expansion of the regret follows from (2.10) and Lemmas 3 and 4.

Theorem 2. *If $m \geq 7$, then*

$$\frac{R_{T_c}-R_{n_c}}{c}=\frac{1}{2}+o(1)\quad \textit{as } c\to 0.$$

Next we consider another sequential procedure with the same stopping time T_c as (2.3) that estimates μ by the sample mean $\overline{X}_{T_c}$. We denote the risk function of the procedure by $\tilde{R}_{T_c}$. Then

$$\begin{aligned}\tilde{R}_{T_c} &= E_\theta\{\exp[a(\overline{X}_{T_c}-\mu)]-a(\overline{X}_{T_c}-\mu)-1+cT_c\}\\ &= E_\theta\left\{\exp\left(\frac{a^2\sigma^2}{2T_c}\right)-1+cT_c\right\}. \qquad (2.12)\end{aligned}$$

Now we give an asymptotic expansion of the regret of the procedure.

Theorem 3. *If $m \geq 4$, then*

$$\frac{\tilde{R}_{T_c}-R_{n_c}}{c}=\frac{a^2\sigma^2}{4}+\frac{1}{2}+o(1)\quad \textit{as } c\to 0.$$

Proof. From (2.12)

$$\tilde{R}_{T_c} = E_\theta\left\{\exp\left(\frac{a^2\sigma^2}{2T_c}\right)-1-\frac{a^2\sigma^2}{2T_c}\right\}+II,$$

where II is given in (2.10). Hence from Lemma 4 it is enough to show that

$$\frac{1}{c}E_\theta\left\{\exp\left(\frac{a^2\sigma^2}{2T_c}\right)-1-\frac{a^2\sigma^2}{2T_c}\right\}=\frac{a^2\sigma^2}{4}+o(1). \qquad (2.13)$$

The left hand side of (2.13) is equal to

$$\begin{aligned}&\frac{1}{c}\int\frac{1}{2}\left(\frac{a^2\sigma^2}{2T_c}\right)^2\exp(\Delta')\,dP_\theta\\ &\qquad=\frac{a^2\sigma^2}{4}\int\left(\frac{n_c}{T_c}\right)^2\exp(\Delta')\,dP_\theta\\ &\qquad=\frac{a^2\sigma^2}{4}III\quad \text{(say)},\end{aligned}$$

where

$$|\Delta'| \leq \frac{a^2\sigma^2}{2T_c} \leq \frac{a^2\sigma^2}{2m}. \tag{2.14}$$

Hence, in order to to prove (2.13), it suffices to show that if $m \geq 4$, then

$$III = 1 + o(1). \tag{2.15}$$

Write III as

$$\begin{aligned} III &= \int_{T<\varepsilon n_c} \left(\frac{n_c}{T_c}\right)^2 \exp(\Delta')\, dP_\theta + \int_{T\geq \varepsilon n_c} \left(\frac{n_c}{T_c}\right)^2 \exp(\Delta')\, dP_\theta \\ &= III_1 + III_2 \quad \text{(say)} \end{aligned} \tag{2.16}$$

for some $0 < \varepsilon < 1$. From (2.9) and (2.14)

$$\begin{aligned} III_1 &\leq \left(\frac{n_c}{m}\right)^2 \exp\left(\frac{a^2\sigma^2}{2m}\right) P_\theta(T_c < \varepsilon n_c) \\ &= O(c^{(m-3)/2}). \end{aligned} \tag{2.17}$$

It is easy to see that

$$III_2 = 1 + o(1). \tag{2.18}$$

Substituting (2.17) and (2.18) into (2.16), (2.15) follows if $m \geq 4$. So the proof is completed.

Zellner (1986) showed that the sample mean is inadmissible even though the variance is unknown. Theorems 2 and 3 show that the same result also asymptotically holds under the sequential setting.

3 Accelerated stopping time

In this section we consider an accelerated stopping time. Let $\kappa \geq 0$ and $0 < \rho < 1$ be fixed constants, and let

$$N_1 = \inf\{n \geq m; n > \rho b \hat{\sigma}_n\}.$$

See (2.3). Based on $X_1, \ldots, X_{N_1}$, we define

$$N_2 = [b\hat{\sigma}_{N_1} + \kappa] + 1,$$

where $[u]$ denotes the largest integer less than u. Let

$$M_c = \max(N_1, N_2).$$

We sample the difference $M_c - N_1$ in one single batch. The stopping time M_c is called an accelerated stopping time. The pooled sample is of size M_c and we estimate μ by $\hat{\delta}_{M_c}$ in (2.5) where T_c is replaced by M_c. Then the risk of the procedure is

$$R_{M_c} = E_\theta\{L(\hat{\delta}_{M_c}, \mu) + cM_c\}.$$

Similar to the proof of (4) (page 221) in Hall (1983), we can get the asymptotic expansion of the expectation of M_c. See also Corollary 2.1 of Mukhopadhyay and Solanky (1991).

Theorem 4. *If $m \geq 3$, then*

$$\kappa - \frac{3}{4}\rho^{-1} \leq \liminf_{c\to 0} E_\theta(M_c - n_c) \leq \limsup_{c\to 0} E_\theta(M_c - n_c) \leq \kappa - \frac{3}{4}\rho^{-1} + 1.$$

Furthermore, if ρ^{-1} is an integer, then

$$E_\theta(M_c) = n_c + \kappa - \tfrac{3}{4}\rho^{-1} + q + o(1) \quad \text{as } c \to 0,$$

where $q = E_\theta\{[\rho^{-1}(2 - Z) + \kappa] + 1 - (\rho^{-1}(2 - Z) + \kappa)\}$ and Z is a random variable on $(0, 2)$.

Note that if $0 < \varepsilon < \rho$, then

$$P_\theta(M_c \leq \varepsilon n_c) \leq P_\theta\left(N_1 \leq \frac{\varepsilon}{\rho} n_c^*\right),$$

where $n_c^* = \rho n_c$ and $\varepsilon/\rho < 1$. So we can get

$$P_\theta(M_c \leq \varepsilon n_c) = O(c^{(m-1)/2}). \tag{3.1}$$

for $0 < \varepsilon < \rho$. See Theorem 2.3 of Mukohopadhyay and Solanky (1991). An analogous argument to that of (2) (page 220) in Hall (1983) shows the following result. See also Theorem 2.2 of Mukhopadhyay and Solanky (1991).

Lemma 5. *If $m \geq 3$, then*

$$\lim_{c\to 0} \frac{E_\theta(M_c - n_c)^2}{n_c} = \frac{1}{2\rho}.$$

It follows from (3.1) and Lemma 5 that the conditions of Lemma 1 are satisfied with $\tau^2 = \dfrac{1}{2\rho}$ if $m \geq 4$. So an argument paralleling that of Theorem 2 shows the following result.

Theorem 5. *If $m \geq 7$, then*

$$\frac{R_{M_c} - R_{n_c}}{c} = \frac{1}{2\rho} + o(1) \quad \text{as } c \to 0.$$

Next we consider another sequential procedure that estimates μ by $\overline{X}_{M_c}$. We denote the risk function by $\tilde{R}_{M_c}$. Then the argument used to prove Theorem 3 leads to the following result.

Theorem 6. *If* $m \geq 4$*, then*

$$\frac{\tilde{R}_{M_c} - R_{n_c}}{c} = \frac{a^2\sigma^2}{4} + \frac{1}{2\rho} + o(1) \quad as\ c \to 0.$$

From Theorems 5 and 6, the sample mean is asymptotically inadmissible under the accelerated stopping time as well.

Acknowledgement: The author thanks the referee for helpful suggestions.

References

Chattopadhyay S (1998) Sequential estimation of normal mean under asymmetric loss function with a shrinkage stopping rule. Metrika 48:53–59

Hall P (1983) Sequential estimation saving sampling operations. J. Roy. Statist. Soc. B 45:219–223

Mukhopadhyay N and Solanky TKS (1991) Second order properties of accelerated stopping times with applications in sequential estimation. Seq. Anal. 10:99–123

Rojo J (1987) On the admissibility of $c\bar{x} + d$ with respect to the LINEX loss function. Commun. Statist. A 16:3745–3748

Varian HR (1975) A Bayesian approach to real estate assessment. In: Fienberg SE and Zellner A (eds.) Studies in Bayesian Econometrics and Statistics in honor of Leonard J. Savage. North Holland, Amsterdam, pp. 195–208

Woodroofe M (1982) Nonlinear renewal theory in sequential analysis. SIAM, Philadelphia

Zellner A (1986) Bayesian estimation and prediction using asymmetric loss functions. J. Amer. Statist. Assoc. 81:446–451