Journal of Korean
Data & Information Science Society
2005, Vol. 16, No. 3, pp. 657~663

Some Properties of Sequential Point Estimation of the Mean[1)]

Kiheon Choi[2)]

Abstract

Under the minimum risk point estimation formulation of Robbins(1959), we consider the sequential point estimation problem for normal population $N(\theta, \theta)$ with unknown parameter θ. In the case of completely unknown θ, Stein's(1945) two-stage procedure is known to enjoy the consistency property, but it is not even first-order efficient. In the case when $\theta > \theta_L$ where $\theta_L (>0)$ is known, the revised two-stage procedure is shown to enjoy all the usual second-order properties.

Keywords : asymptotic second-order efficient, minimum risk point estimation stopping times, two-stage procedure

1. Introduction

One of the most celebrated contributions to the sequential estimation due to Robbins(1959) used the framework provided by decision theory in a rather different way from that just described. In that paper, Robbins formulated a minimum risk point estimation approach for the mean of a normal population whose variance was also unknown. He used as stopping rules natural ones arising from this formulation based on a continuous upgrading of the parameter estimate. Such a rule might take the form

$$N = \inf\{ n; n \geq \max[m, g(\hat{\theta}_n)]\} \tag{1.1}$$

where m is the size of some initial sample. The problem has a long history,

---

1) This work was supported by the Institute of Natural Science Fund, Duksung Women's University, 2005.

2) Professor, Department of Statistics, Duksung Women's University, Seoul 132-714, Korea.
E-mail : khchoi@duksung.ac.kr

much of which is included in Khan(1969), Ghosh, Mukhopadhyay, and Sen(1997), and Choi, Weng, and Woodroofe(2001) have given a method for determining stopping rules in order to obtain fixed-width confidence intervals of prescribed coverage probability for an unknown parameter of a distribution possibly involving some unknown nuisance parameters. The results are only asymptotic, and rely on the asymptotic of Chow and Robbins(1965).

We suppose that we have at our disposal a sequence of independent observations $X_1, X_2, \cdots$ from a $N(\theta, \theta)$ distribution where $\theta(>0)$ is an unknown parameter. Mukhopadhyay and Cicconetti(2004) investigated purely sequential and two-stage bounded risk methodologies for estimating θ. They showed the first-order asymptotic properties of both purely sequential as well as sample sizes with regard to the loss function given by $L_n(\hat{\theta}_n, \theta) \equiv A(\hat{\theta}_n - \theta)^2$. This shows that the two-stage procedure is asymptotic second-order efficient.

Choi(2005) investigated some properties in sequential point estimation of the mean under the loss function given by

$$L_n(\hat{\theta}_n, \theta) \equiv A(\hat{\theta}_n - \theta)^2 + cn, \tag{1.2}$$

where $A(>0)$ is a known weight and $c(>0)$ is known cost per unit sample. The goal is to determine the sample size minimizing the risk function

$$E[L_n(\hat{\theta}_n, \theta)] \equiv AE(\hat{\theta}_n - \theta)^2 + cn. \tag{1.3}$$

If θ were known, this could be achieved by a sample of size $n^* = (A/c)^{1/2}$, the corresponding minimal risk being

$$R_{n^*}(c) \equiv E_q[L_n(\hat{\theta}_{n^*}, \theta)] = 2cn^*. \tag{1.4}$$

For convenience, we assume that n^* is an integer. It is known that no fixed sample size procedure will minimize the risk $E[L_n(\hat{\theta}_n), \theta)]$ uniformly in θ. For unknown θ there does not exist any fixed-sample size for which the risk can be minimized simultaneously for all θ. Motivated by the expression of n^* for θ known, it is natural the following two-stage procedure we proposed in Stein(1945). While estimating θ by the MLE $\hat{\theta}_n$, we have

$$R_n(\hat{\theta}_n, \theta) \approx An^{-1}I_X^{-1}(\theta) + cn, \tag{1.5}$$

which suggests that the required sample size n must be chosen to exceed

$$n^* \equiv n^*(c) = (Ac^{-1} I_X^{-1}(\theta))^{1/2} = (Ac^{-1} 2\theta^2 (2\theta + 1)^{-1})^{1/2}. \quad (1.6)$$

However, we tacitly disregard the fact that n^* may not be an integer. It is clear that the magnitude of n^* remains unknown. Incidentally, we describe asymptotic analysis as $c \to 0$. One should note that making c approach zero is merely a mathematical device that allows the optimal fixed-sample-size n^* as well as both initial and final sample size m and N from Eq. (1.7) to grow up to infinity. For a general overview of the area of sequential estimation, on may refer to Ghosh, Mukhopadhyay, and Sen(1997), a comprehensive resource.

Now, we discuss sampling in at most two batches. We again consider the initial sample size m, the pilot observations $X_1, \cdots, X_m$, and defined the stopping variable or final sample size as follows: Let

$$N \equiv N(c) = Max\{ m, [(Ac^{-1} 2\hat{\theta}_m^2 (2\hat{\theta}_m + 1)^{-1})^{1/2}]\}. \quad (1.7)$$

where [u] stands for the largest integer smaller than u.

Choi(2005) proved the following properties for the two-stage estimation methodology.

Lemma 1.1. *Under the sole assumption* $I_X^{-1}(\theta) < \infty$, *the sequential procedure terminates with probability* 1, that is, closed sequential sampling procedure.

Since this stopping rule terminates with probability one for every fixed θ, A and c. Thus, we propose to estimate θ by $\hat{\theta}_N$, namely $\sqrt{T_N + 1/4} - 1/2$ with $T_N = N^{-1} \sum_{i=1}^{N} X_i$ based on $X_1, \cdots, X_N$. Some interesting asymptotic results are now summarized in the following theorems.

Theorem 1.1. *For the two-stage estimation procedure, for all fixed* $\theta > 0$ *and* $A > 0$, $\hat{\theta}_N$ *is unbiased estimator of* θ.

Theorem 1.2. *For the two-stage estimation procedure, for all fixed* $\theta > 0$ *and* $A > 0$, *we have as* $c \to 0$:

(i) $N / n^* \to 1$ w.p.1,

(ii) $E_\theta(N) / n^* \to 1$ (asymptotic efficiency).

2. Minimum Risk Point Estimation and Second-Order Analyses

Customarily, the random sample size N proposed in this problem area has been such that $I(N=n)$ and $\hat{\theta}_n$ turn out to be independent for all $n \geq 2$, while $\hat{\theta}_n$ is proposed as the final estimator of θ. The risk associated with $\hat{\theta}_N$ is then given by

$$R(c) = E(L_N) = c\{E(N) + E(n^{*2}/N)\}, \tag{2.1}$$

while the risk efficiency and regret of the methodology under consideration are respectively defined by

$$e(c) = \frac{R(c)}{R_{n*}(c)} = \frac{1}{2}\{E(N/n^*) + E(n^*/N)\} \tag{2.2}$$

$$w(c) = R(c) - R_{n*}(c) = cE\{(N-n^*)^2/N\} \tag{2.3}$$

For the purely sequential procedure of Robbins(1959), the second-order asymptotic(as $c \rightarrow 0$) expansion such as $E(N-n^*) = \eta' + o(1)$ and $w(c) = (1/2)c + o(c)$ were obtained in Woodroofe(1977) under appropriate conditions on the initial sample size, where η' is a computable real number. These are the sharpest known results in the purely sequential framework. The roles of two-stage and modified two-state procedures in the president problem are summarized by Mukhopadhay and Solanky(1994).

Let us assume that θ is unknown. But, in a practical application, on may know a viable lower bound $\theta_L(>0)$ for the unknown parameter θ. Observe that $I_X^{-1}(\theta) = 2\theta^2(2\theta+1)^{-1}$ is an increasing function of θ, so that we can conclude that $I_X^{-1}(\theta) \geq I_X^{-1}(\theta_L) = 2\theta_L^2(2\theta_L+1)^{-1}$, and pursue the route of two-stage sampling. Define

$$m = m(c) = \max\{m_0, [(\frac{A}{c} I_X^{-1}(\theta_L))^{\frac{1}{2}}] + 1\} \tag{2.4}$$

where $m_0(\geq 2)$ is a fixed integer. Let $X_1, \cdots, X_m$ be the pilot observations which provide $\hat{\theta}_m$. Then, let

$$N = N(c) = \max\{m, [(\frac{A}{c} I_X^{-1}(\hat{\theta}_m))^{\frac{1}{2}}] + 1\}, \tag{2.5}$$

and we implement this two-stage procedure as we did in the case of (1.7). Finally, we estimate θ by $\hat{\theta}_N$. Along the lines of Section 2, one can obtain the first-order results such as: $E(N/n^*) \to 1$ and $e(c) \to 1$ as $c \to 0$; that is, the two-stage procedure (2.4)-(2.5) is both asymptotically first-order efficient and risk efficient. But, let us focus more on the second-order analysis. First, we consider the following results. Details are omitted, and their proof can be found similar in Mukhopadhyay and Duggan(1997).

Lemma 2.1. *For the two-stage procedure (2.4)-(2.5), we have:*
(i) $P(N = m) = O(m^{-k})$ *where the expression for* m *was defined in (2.4);*
(ii) $n^{*-\frac{1}{2}}(N - n^*) \to N(0, \frac{1}{2}(I_X^{-1}(\theta) I_X(\theta_L))^{\frac{1}{2}})$ *in distribution as* $c \to 0$;
(iii) $(N - n^*)^2/n^*$ *is uniformly integrable for* $0 < c < c_0$ *with sufficiently small* c_0.

Now, the main results are summarized as follows.

Theorem 2.1. *For all values of* $\theta(>\theta_L)$, *we have for the two-stage procedure (2.4) -(2.5) as* $c \to 0$:
(i) $-\frac{1}{4}(I_X^{-1}(\theta) I_X(\theta_L))^{\frac{1}{2}} + o(1) \le E(N - n^*) \le 1 - \frac{1}{4}(I_X^{-1}(\theta) I_X(\theta_L))^{\frac{1}{2}} + o(1)$;
(ii) $w(c) = \frac{1}{2}(I_X^{-1}(\theta) I_X(\theta_L))^{\frac{1}{2}} c + o(c)$, *with the regret function defined in (2.3).*

Proof. Part (i) follows from the basic inequality $(A c^{-1} I_X^{-1}(\hat{\theta}_m))^{1/2} \le N \le m I(N = m) + (A c^{-1} I_X^{-1}(\hat{\theta}_m))^{1/2}$ and the fact that $E((I_X^{-1}(\hat{\theta}_m))^{1/2}) = I_X^{-1}(\hat{\theta}_m)\{1 - \frac{1}{4} m^{-1} + O(m^{-2})\}$.

In order to prove part (ii), we split the expectation in (2.3) on the sets $\{N = m\}$ and $\{N > m\}$ as it was in the case of proving Theorem 3.2 of Mukhopadhyay and Duggan(1997). Now combing part (i) and (ii) of Lemma 2.1, we can write

$$E\{N^{-1}(N - n^*)^2 I(N > m)\} = \frac{1}{2}(I_X^{-1}(\theta) I_X(\theta_L))^{\frac{1}{2}} + o(1). \tag{2.6}$$

Next, $E\{N^{-1}(N-n^*)^2 I(N>m)\} \le O(m)P(N=m) = o(1)$ in view of part (i), Lemma 2.1. Combing this with (2.6), we now have part (ii).

Simulation. In order to justify the results of Theorem 2.1 we shall give brief simulation results. We are interested in the performance of our sequential procedure for various values of θ_L, and so we consider the cases when $\theta_L = 2$ and 4. Since the cost c is sufficiently small in our theorem, the values of c are chosen such that $n^* = (Ac^{-1}2\theta^2(2\theta+1)^{-1})^{1/2} = 30,\ 50,\ 100,\ 300,\ 800.$ The pilot sample size is set at $m=10$ which are sufficient. The simulation results in Tables 1 and 2 are based on 10,000 repetitions by means of the stopping rule defined by (2.3). From Theorem 2.1 we have that $E(N) = n^* + 1 + o(1)$ and $w(c)/c = \frac{1}{2}(I_X^{-1}(\theta)I_X(\theta_L))^{\frac{1}{2}} + o(1)$ as $c \to 0$. Table 1 and 2 show that these results are justified.

Table 1: $\theta_L = 2$

n^*	m	c	$E(N)$	$w(c)/c$
40	20	0.002222222	40.0454	0.03225806
200		0.00008888889	200	0
400		0.0002222222	401	0.002493766
500		0.00001422222	500	0

Table 2: $\theta_L = 4$

n^*	m	c	$E(N)$	$w(c)/c$
30	10	0.001777778	31	0.03225806
50		0.00064	50	0
100		0.00016	100	0
300		1.7777778×10^{-5}	301	0.003322259
800		2.5×10^{-6}	801	0.001248439

Reference

1. Choi, K., Weng, R. C., and Woodroofe, M. (2001). Sequential Confidence

Intervals for a Population Size with Fixed Proportional Accuracy. *Sequential Analysis*, 20 25-43.

2. Choi, K. (2005). On the Point Estimation of Stein's Two-Stage Procedure, *Journal of the Korean Data Analysis Society*, Vol. 7, No. 3, 729-735.
3. Chow, Y. S.; Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* 36, 457-462.
4. Ghosh, M., Mukhopadhay, N. and Sen, P. K. (1997). *Sequential Estimation*; John Wiley: New York.
5. Khan, R. A. (1969). A general method of determining fixed-width confidence intervals. *Ann. Math. Statist.* 40, 704-709.
6. Mukhopadhay, N. and Solanky, T.K.S. (1994). *Multistage Selection And Ranking Procedures*, Marcel Dekker, Inc.
7. Mukhopadhay, N. and Duggan, W. T. (1997). Can a two-stage procedure enjoy second-order properties? *Sankhya Ser. A* 59, 435-448.
8. Mukhopadhyay, N. and Cicconetti, G. (2004). Application of Sequential Estimating the Mean in a Normal Distribution Having Equal Mean and Variance, *Sequential Analysis*, 23, 625-665.
9. Robbins, H. (1959). Sequential estimation of the mean of a normal population. In *Probability and Statistics*; Grenander, U., Ed.; Almquist and Wiksell: Uppsala; H. Cramer Vol., 235-245.
10. Stein, C. (1945). A two sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* 16, 243-258.
11. Woodroofe, M. (1977). Second-order approximations for sequential point and interval estimation. *Ann. Statist.* 5, 984-995.

[received date : Apr. 2005, accepted date : Jul. 2005]