Open access • Journal Article • DOI:10.1080/03610927908827789

Sequential point estimation of the mean when the distribution is unspecified

— Source link

Malay Ghosh, Nitis Mukhopadhyay

Institutions: Iowa State University, University of Missouri

Published on: 01 Jan 1979 - Communications in Statistics-theory and Methods (Marcel Dekker, Inc.)

Topics: Mean squared error, Independent and identically distributed random variables, Asymptotic distribution, Moment (mathematics) and Point estimation

SEQUENTIAL POINT ESTIMATION OF THE MEAN WHEN THE DISTRIBUTION IS UNSPECIFIED

By

Nitis Mukhopadhyay
University of Minnesota, Minneapolis

Technical Report No. 312

AMS Classification: Primary 62L12, Secondary 60F05.

Key Words and Phrases: Distribution-free, Point estimation, Sequential methods, Squared error loss, Minimum risk, Ergodic, Uniform integrability, Risk efficiency,

Summary

Consider a sequence $X_1, X_2, \ldots$ of iid random variables with a distribution function F, not necessarily normal. Let μ and σ^2 be respectively the mean and variance of F, both being unknown. We assume $0 < \sigma_0 \leq \sigma < \infty$ for a known σ_0, and $E(|X|^{2+\delta}) < \infty$ for some $\delta > 0$. The loss structure is the cost of observations cn plus the squared error loss $A(\bar{X}_n - \mu)^2$ due to error in estimating μ by $\bar{X}_n$. A sequential procedure has been proposed to achieve the minimum risk (approximately). It is shown that this procedure is asymptotically risk efficient, as c approaches zero.

1. Introduction

Suppose $X_1, X_2, \ldots$ are iid with a distribution function F, having $\sigma^2 = E(X - \mu)^2$ positive and finite, μ being the mean of X. We assume that μ and σ^2 are both unknown. Having recorded n observations $X_1, X_2, \ldots, X_n$, suppose the loss incurred in estimating μ by $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$ is given by

$$(1.1) \qquad L_n = A(\bar{X}_n - \mu)^2 + cn$$

where A and c are known positive quantities, c being the cost per observation.

Our object is to minimize the risk,

$$(1.2) \qquad R_n(c) = E(L_n) = A\sigma^2 \cdot n^{-1} + cn$$

associated to (1.1). $R_n(c)$ is minimum for $n = n^* = b\sigma$, where $b = (A/c)^{\frac{1}{2}}$, and the minimum risk $R_{n^*}(c) = 2cn^*$.

Now, since σ is unknown, no fixed sample size procedure will minimize $R_n(c)$, uniformly for all σ. So, we consider the possibility of utilizing a sample of random size N. The associated risk of the procedure will be,

$$(1.3) \qquad E(L_N) = AE(\bar{X}_N - \mu)^2 + cE(N) .$$

We would like to examine if the risk efficiency, $\eta(c) = E(L_N)/R_{n^*}(c)$ converges to 1, as $c \to 0$. This problem is quite old, and sequential point estimation procedures for some specific non-normal populations are studied in [4], [6], [7]. Also, each problem requires a separate analysis. There is no unified non-parametric approach (in the sense of F being unknown) to the present problem, unlike the one of Chow and

Robbins (1965) for the fixed-width confidence intervals. The present work may be regarded as an attempt towards that goal. It has been shown that for our stopping rule, $\eta(c) \to 1$ as $c \to 0$, under the following assumptions:

Assumptions:

(A1) $\sigma_0 \leq \sigma < \infty$ for a known positive σ_0.

(A2) $E(|X|^{2+\delta})$ is finite, for some $\delta > 0$.

The assumption (A1) seems to be reasonable from a practical point of view. From earlier experiences, some lower bound on σ could be guessed in most situations. We stress that we do not however need σ_0 to be close to σ. The assumption (A2) is for mathematical convenience. This issue has been discussed again on page no. 4.

Let us now discuss the sequential procedure and study its properties.

2. Stopping rule and its properties.

Let $n_0 = \max\{[b\sigma_0],2\}$, where $[y]$ is the largest integer $< y$. Define $u_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$ for $n \geq 2$. We take the estimate of σ^2 as s_n^2, which is u_n^2 if F is continuous, and $u_n^2 + n^{-1}$ if F is discrete. We define the stopping rule $\mathcal{R}$ as follows:

$\mathcal{R}$: The stopping number $N \equiv N_c$ is the first integer n $(\geq n_0)$ for which

$$n \geq b \max\{s_n, \sigma_0\}, \tag{2.1}$$

n_0 being the starting sample size. When we stop, we propose $\bar{X}_N$ as the estimate of μ.

The stopping time N is well-defined. It is easy to see

<u>Lemma 1.</u> N is non-increasing in c, $P(N < \infty) = 1$ for any fixed $c > 0$.

<u>Remark 1.</u> Since $n_0 \to \infty$ as $c \to 0$, $\lim_{c \to 0} N = \infty$ a.s. We have an important result in the following lemma.

<u>Lemma 2.</u> $\lim_{c \to 0} N/n^* = 1$ a.s.

<u>Proof:</u> We would distinguish the two cases.

<u>Case 1.</u> Suppose $[b\sigma_0] \leq 2$, in this case $n_0 = 2$. So we must have

$$N - 1 < b \max(s_{N-1}, \sigma_0) + (2 - 1)$$

(2.2) i.e. $N < b \max(s_{N-1}, \sigma_0) + 2$.

<u>Case 2.</u> Suppose $[b\sigma_0] > 2$, in this case $n_o = [b\sigma_0]$.

<u>Claim.</u> $P[N = n_0] = 0$

Suppose $P[N = n_0] > 0$, so it is possible to obtain

$$n_0 \geq b \max(s_{n_0}, \sigma_0) \geq b\sigma_0 .$$

Now $n_0 = [b\sigma_0] < b\sigma_0$, which is a contradiction, hence, the claim. We can immediately write

(2.3) $N - 1 < b \max(s_{N-1}, \sigma_0)$,

since $N \geq n_0 + 1$.

Combining (2.2) and (2.3), we arrive at the basic inequality

(2.4) $b \max(s_N, \sigma_0) \leq N < b \max(s_{N-1}, \sigma_0) + 2$.

Using the remark 1, the SLLN, one gets

$$\lim_{c \to 0} N/n^* = 1 \text{ a.s. as } c \to 0 ,$$

since $\sigma \geq \sigma_0$. Thus, the proof of Lemma 2 is complete.

We are now in a position to prove the following result.

Lemma 3. $\lim_{c\to 0} E(N/n^*) = 1$.

Proof: Looking at the right hand side of (2.4),

$$N < b\ (s_{N-1} + \sigma_0) + 2$$

$$< b\ (u_{N-1} + \sigma_0 + 1) + 2$$

$$(2.5) \qquad < b\ (u + \sigma_0 + 1) + 2$$

where $u = \sup_{n\geq 2}\{u_n\}$. Note that $u^2 \leq \sup_{n\geq 2}\{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\}$.

Under (A2), using dominated ergodic theorem [theorem 5, [8]]

$\sup_{n\geq 2}\{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\}$ is integrable. Hence $E(u^2) < \infty$. (2.5) leads to

$$(2.6) \qquad 0 < N/n^* < \frac{(u + \sigma_0 + 1)}{\sigma} + \frac{2}{\sigma} \quad \text{if } A > c .$$

Since N/n^* is dominated by an integrable function $(0 < c < A)$, using Lemma 2, dominated convergence theorem, we complete the proof of Lemma 3.

Remark 2. We needed (A2) to verify that $E(u^2) < \infty$. In fact, we really need to show that $E(u) < \infty$. We strongly feel that we can do away without (A2). But, it has not yet been proved.

Remark 3. Inequality (2.6) depicts the fact that $E(N) < \infty$ for all fixed c (> 0).

3. Main result and its proof.

THEOREM. Under (A1), (A2), $\lim_{c\to 0}\{E(L_N)/R_{n^*}(c)\} = 1$.

Proof: In view of (1.3) and Lemma 3, it suffices to show that

$$(3.1) \qquad \lim_{c\to 0}\{\frac{A\ E(\bar{X}_N - \mu)^2}{c\ n^*}\} = 1 .$$

Observe that

$$(\overline{X}_N - \mu)^2 = \frac{(\sum_{i=1}^{N} X_i - N\mu)^2}{n^{*2}} + \frac{(\sum_{i=1}^{N} X_i - N\mu)^2}{n^{*2}} \{\frac{n^{*2}}{N^2} - 1\}$$

which implies

$$(3.2) \qquad \frac{A}{cn^*} (\overline{X}_N - \mu)^2 = \frac{(\sum_{i=1}^{N} X_i - N\mu)^2}{\sigma^2 \cdot n^*} + \frac{(\sum_{i=1}^{N} X_i - N\mu)^2}{\sigma^2 \cdot n^*} \{\frac{n^{*2}}{N^2} - 1\}$$

$$= I + J \text{ , say.}$$

Using the result of Anscombe (1952), one gets

$$I \xrightarrow{\mathcal{L}} \{N(0,1)\}^2 \quad \text{as} \quad c \to 0 \text{ ,}$$

Also using a result of Chow et al. (1965),

$$E(I) = \frac{\sigma^2 \cdot E(N)}{\sigma^2 \cdot n^*} = \frac{E(N)}{n^*} \to 1 \quad \text{as} \quad c \to 0 \text{ ,}$$

by Lemma 3. Hence the family $\{I\}$ is uniformly integrable in the positive parameter c (loeve (1963), p. 183).

Now, look at the rule $\mathcal{R}$ in (2.1).

$N \geq b \max(s_N, \sigma_0) \geq b\sigma_0 = \frac{\sigma_0}{\sigma} n^*$ so that $n^*/N \leq \sigma/\sigma_0$.

Hence

$$(3.3) \qquad -1 \leq \frac{n^{*2}}{N^2} - 1 \leq \frac{n^{*2}}{N^2} \leq \sigma^2/\sigma_0^2 \text{ .}$$

(3.3), together with the fact that $\{I\}$ is uniformly integrable, imply that, so is $\{J\}$. Using Lemma 2, $E(J) \to 0$ as $c \to 0$. Hence (3.1) is verified, and so the proof of the Theorem is complete.

4. What can be done without (A2)?

Suppose we taken $n_0 = \max\{[b\sigma_0], 3\}$, i.e. $n_0 \geq 3$. By looking at the rule $\mathcal{R}$,

$$(N - 1)^2 < b^2 \max(s_{N-1}^2), \sigma_0^2) + (3-1)^2$$

$$< b^2 (1 + u_{N-1}^2 + \sigma_0^2) + 4$$

$$\Rightarrow \quad (N-1)^2(N-2) < b^2_2 \left(N + \sum_1^{N-1} (X_i - \bar{X}_{N-1})^2 + \sigma_0^2(N-2)\right) + 4(N-2)$$

$$\leq b^2\{N + \sum_1^{N} (X_i - \mu)^2 + \sigma_0^2 \cdot N\} + 4N .$$

Now $(N-1)^2(N-2) \geq N(N-3)^2$, since $n_0 \geq 3$. So,

$$N(N-3)^2 \leq b^2\{\sum_1^N(X_i - \mu)^2 + \sigma_0^2 N + N\} + 4N .$$

Assume $E(N) < \infty$. Using Jensen's inequality, Wald's first equation, one gets

$$(EN)(EN-3)^2 \leq b^2(EN)(\sigma^2 + \sigma_0^2 + 1) + 4(EN)$$

$$\Rightarrow \quad (EN-3)^2 \leq b^2(\sigma^2 + \sigma_0^2 + 1) + 4$$

$$\Rightarrow \quad E(N) - 3 \leq b(\sigma + \sigma_0 + 1) + 2 , \quad \text{since } E(N) \geq 3 .$$

$$(4.1) \quad \Rightarrow \quad E(N) \leq b(\sigma + \sigma_0 + 1) + 5 .$$

If $E(N)$ is not finite, define $N_K = \min(N,K)$ for positive integers K . Now $N_K \uparrow N$ a.s. as $K \to \infty$. Also $E(N_K) \leq b(\sigma + \sigma_0 + 1) + 5$. Monotone convergence theorem will yield

$$(4.2) \qquad E(N) \leq b(\sigma + \sigma_0 + 1) + 5 ,$$

which shows that $E(N)$ is finite, even without (A2). But (4.2) does not quite lead to Lemma 3.

In the same way, one can get (when F is continuous)

$$N(N-3)^2 \leq b^2 \max\{\sum_1^N (X_i - \mu)^2, \sigma_0^2 N\} + 4N$$

$$= \frac{b^2}{2}\{\sum_1^N(X_i - \mu)^2 + \sigma_0^2 N + |\sum_1^N(X_i - \mu)^2 - \sigma_0^2 N|\} + 4N .$$

Somehow, one must show $E|\sum_1^N(X_i - \mu)^2 - \sigma_0^2 N| \to 0$, if at all it is true, as $c \to 0$, without using (A2).

Remark 4. It will be of much importance to examine the behavior of $\mathcal{R}$ for normal F , under (A1) for moderate c .

Remark 5. It will be of considerable practical importance to take a few non-normal F , and examine the behavior of $\mathfrak{R}$ for moderate c . In this case, one must take resort to simulation methods, it seems.

Remark 6. Under (A1) and (A2), we do not yet know the order of

$\lim_{c\to 0}\{E(L_N) - R_{n*}(c)\}$.

References

[1] Anscombe, F. J. (1952). Large sample theory of sequential estimation. Proc. Cam. Phil. Soc., 48, 600-607.

[2] Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. Ann. Math. Statist., 36, 457-462.

[3] Chow, Y. S., Robbins, H. and Teicher, H. (1965). Moments of randomly stopped sums. Ann. Math. Statist., 36, 789-799.

[4] Ghosh, M. and Mukhopadhyay, N. (1975). Sequential estimation of the parameter of a rectangular distribution. Cal. Statist. Assoc. Bull., 24, 117-122.

[5] Loeve, M. (1963). Probability Theory. Van Nostrand, Princeton.

[6] Robbins, H. and Siegmund, D. (1974). Sequential estimation of p in Bernoulli trials. Studies in Probability and Statistics (E. J. Williams ed.), Jerusalem Academic Press.

[7] Starr, N. and Woodroofe, M. (1972). Further remarks on sequential estimation: the exponential case. Ann. Math. Statist., 43, 1147-1154.

[8] Wiener, N. (1939). The ergodic theorem. Duke J. Math., 5, 1-18.