A Class of Purely Sequential Minimum Risk Point Estimation Methodologies with Second-Order Properties

JUN HU
Oakland University
Department of Mathematics and Statistics
Rochester, Michigan
USA
junhu@oakland.edu

NITIS MUKHOPADHYAY
University of Connecticut
Department of Statistics
Storrs, Connecticut
USA
nitis.mukhopadhyay@uconn.edu

Abstract: Under the squared error loss plus linear cost of sampling, we revisit the *minimum risk point estimation* (MRPE) problem for an unknown normal mean μ when the variance σ^2 also remains unknown. We begin by defining a new class of purely sequential MRPE methodologies based on a general estimator W_n for σ satisfying a set of conditions in proposing the requisite stopping boundary. A number of desirable asymptotic first-order and second-order properties associated with this new class of estimation methodologies have been investigated. After such general considerations, we include a number of substantial illustrations where we respectively substitute appropriate multiples of Gini's mean difference and the mean absolute deviation in the place of the general estimator W_n.

Key–Words: Minimum Risk Point Estimation, Regret Expansion, Risk Efficiency, Sequential Sampling, Simulations

1 Introduction

In literature, purely sequential estimation methodologies date back to path-breaking papers of Anscombe (1953), Ray (1957), and Chow and Robbins (1965). They gave a solid foundation to establish purely sequential fixed-width confidence interval estimation methodologies for an unknown normal mean μ when the population variance σ^2 remained unknown. Indeed, Chow and Robbins (1965) brought forward the fundamental nature of the theory of purely sequential nonparametric fixed-width confidence interval estimation methodologies.

The far reaching purely sequential minimum risk point estimation methodology was originally put forward by Robbins (1959). It was subsequently broadened by Starr (1966) and Starr and Woodroofe (1969), where asymptotic properties such as efficiency and risk efficiency were summarized. Second-order properties were further studied in Woodroofe (1977).

Instead of using the customary sample variance (or sample standard deviation) as an estimator of the unknown parameter σ^2 (or σ) in the stopping rules, Sen and Ghosh (1981) considered nonparametric sequential point estimation based on U-statistics. They concluded the asymptotic *first-order efficiency* and *risk efficiency* as well as other elegant asymptotics. As Mukhopadhyay (1982) used a broader class of nonparametric estimators of σ^2, Chattopadhyay and Mukhopadhyay (2013), and Mukhopadhyay and Hu (2017, 2018) recently looked into appropriate functions of *Gini's Mean Difference* (GMD) or *Mean Absolute Deviance* (MAD) as possible substitutes of the traditional sample variance (or sample standard deviation).

The object of this paper is to revisit in depth the purely sequential minimum risk point estimation methodologies involving GMD or MAD established in Mukhopadhyay and Hu (2017), Hu and Mukhopadhyay (2019) and Hu and Zhuang (2019). Having proposed a new purely sequential methodology based on nonparametric estimators with some certain conditions satisfied, we develop asymptotic second-order results which are considerably stronger than those reported. The formulations of the newly proposed methodology are presented in Section 2. The main theorems are laid down in Section 3, along with substantial proofs. In Section 4, some illustrations are provided followed by summaries from simulations presented in Section 5. We end with more discussions in Section 6.

2 Formulations

Assuming that we have a sequence of independent observations $X_1, X_2, ...$ from a $N\left(\mu, \sigma^2\right)$ population with $-\infty < \mu < \infty$ and $0 < \sigma < \infty$, both unknown.

Having recorded $X_1, X_2, ..., X_n, n \geq 2$, we denote the customarily used unbiased estimator for μ by

$$\text{Sample mean: } \overline{X}_n \equiv n^{-1}\Sigma_{i=1}^{n} X_i,$$

and denote an appropriate consistent nonparametric estimator for σ by $W_n \equiv W_n(X_1, ..., X_n)$, where the sequence of $\{W_n, n \geq 2\}$ further satisfies the following conditions:

(C1) *Independence*: $\overline{X}_n$ and $\{W_k : 2 \leq k \leq n\}$ are distributed independently for all $n \geq 2$.

(C2) *Convergence in probability*:

$$W_n \overset{P_{\mu,\sigma}}{\to} \sigma \text{ as } n \to \infty.$$

(C3) *Aysmptotic normality*:

$$\sqrt{n}\left(\sigma^{-1}W_n - 1\right) \overset{L}{\to} N\left(0, \delta^2\right) \text{ as } n \to \infty,$$

for some $\delta(> 0)$.

(C4) *Uniform continuity in probability* (u.c.i.p): For every $\varepsilon > 0$, there exists a large $\nu \equiv \nu(\varepsilon)$ and small $\gamma > 0$ for which

$$P_{\mu,\sigma}\left(\max_{0 \leq k \leq n\gamma} |W_{n+k} - W_n| \geq \varepsilon\right) < \varepsilon$$

holds for any $n \geq \nu$.

(C5) *Kolmogorov's inequality*: For every $\varepsilon > 0$, and some $2 \leq n_1 \leq n_2$,

$$P_{\mu,\sigma}\left(\max_{n_1 \leq n \leq n_2} |W_n \geq \sigma| \geq \varepsilon\right)$$
$$\leq \varepsilon^{-r} E_{\mu,\sigma}\left[|W_{n_1} - \sigma|^r\right], \text{ with } r \geq 2.$$

(C6) *Order of central absolute moments*: For $n \geq 2$ and $r \geq 2$,

$$E_{\mu,\sigma}\left[|W_n - \sigma|^r\right] = O(n^{-r/2}).$$

(C7) Wiener's condition:

$$E_{\mu,\sigma}\left[\sup_{n \geq 2} W_n\right] < \infty.$$

Now we are in a position to propose a new purely sequential minimum risk point estimation methodology for an unknown normal mean based on W_n under a widely-used loss function given by

$$L_n \equiv L_n\left(\mu, \overline{X}_n\right) = A\left(\overline{X}_n - \mu\right)^2 + cn, \quad (1)$$

where $A(> 0)$ and c are both known. Here, A is an appropriate weight function, c is the unit cost of each observation, and n is the sample size. Associated with the loss function in (1), we can write the risk function as follows:

$$R_n(c) \equiv E_{\mu,\sigma}\left\{L_n\left(\mu, \overline{X}_n\right)\right\} = A\sigma^2 n^{-1} + cn, \quad (2)$$

by minimizing which we obtain the optimal fixed sample size n^* given by

$$n^* \equiv n^*(c) = \sigma\sqrt{A/c}, \quad (3)$$

had σ been known. And we tacitly disregard the fact that n^* may not be an integer.

Beginning with the pilot data $X_1, X_2, ..., X_m$ of size $m\,(\geq 2)$, we sample one additional observation at a time sequentially as needed until we terminate according to the following stopping rule:

$$\begin{aligned} N &\equiv N(c) \\ &= \inf\left\{n \geq m : n \geq \sqrt{A/c}\left(W_n + n^{-\lambda}\right)\right\}, \end{aligned} \quad (4)$$

where $\lambda(> 1/2)$ is held fixed. That is, if we have a pilot sample of size m such that $m \geq \sqrt{A/c}\left(W_m + m^{-\lambda}\right)$ already holds, no additional observations will be recorded and the final sample size is $N = m$. Otherwise, we record one more observation at a time and update n in the stopping rule (4). We terminate the sampling procedure at the first time that $N = n\,(\geq m)$ is observed such that $n \geq \sqrt{A/c}\left(W_n + n^{-\lambda}\right)$. Finally with $\{N, X_1, ..., X_m, ..., X_N\}$, we establish the minimum risk point estimator for μ as follows:

$$\overline{X}_N \equiv N^{-1}\Sigma_{i=1}^{N} X_i. \quad (5)$$

For the new class of sequential methodologies (4)-(5), it is obvious that $P_{\mu,\sigma}(N < \infty) = 1$ and $N \uparrow \infty$ w.p.1 as $c \downarrow 0$.

3 Main Results

In this section, we lay down a number of main lemmas and theorems associated with the new class of purely sequential minimum risk point estimation methodologies given by (4)-(5). For the stopping time N defined in (4), for all fixed μ, σ, and A, we have the following:

Theorem 1 (Asymptotic First-Order Efficiency) *With n^* defined by (3), we have*

$$\lim_{c \to 0} E_{\mu,\sigma}\left\{N/n^*\right\} = 1. \quad (6)$$

Proof: By the stopping rule defined in (4), we have the following two inequalities

$$N \geq \sqrt{A/c}\left(W_N + N^{-\lambda}\right), \tag{7}$$

as well as

$$N < m + \sqrt{A/c}\left(W_{N-1} + (N-1)^{-\lambda}\right), \tag{8}$$

from which we conclude

$$\frac{W_N + N^{-\lambda}}{\sigma} \leq \frac{N}{n^*} < \frac{W_{N-1} + (N-1)^{-\lambda}}{\sigma} + \frac{m}{n^*}. \tag{9}$$

Then, it is clear that as $c \to 0$, $N/n^* \overset{P_{\mu,\sigma}}{\to} 1$ under (C2). Also, we note that for some sufficiently small c, the right-hand side of (9) can be bounded as follows:

$$0 \leq N/n^* < \sigma^{-1}\left(\sup_n W_n + 1\right) + 1. \tag{10}$$

Under (C7) and by the dominated convergence theorem, therefore, $\lim_{c\to 0} E_{\mu,\sigma}[N/n^*] = 1$ holds immediately from (10). $\square$

Lemma 2 *For any arbitrary* $0 < \eta < 1$*, with* $r \geq 2$*, we have*

$$P_{\mu,\sigma}(N \leq \eta n^*) = O\left(n^{*^{-\frac{r}{2(1+\lambda)}}}\right). \tag{11}$$

Proof: Let $[u]$ denote the largest integer that is smaller than u and we define:

$$n_{1c} = [(A/c)^{\frac{1}{2(1+\lambda)}}] = O(c^{-\frac{1}{2(1+\lambda)}}),$$

and

$$n_{2c} = \eta n^* = \eta\sigma\sqrt{A/c}.$$

Clearly, $N \geq n_{1c}$ w.p.1 from the definition of N in (4). Next, we set out to obtain the rate at which $P_{\mu,\sigma}\{N \leq \eta n^*\}$ may converge to zero for small c:

$$\begin{aligned}
& P_{\mu,\sigma}\{N \leq \eta n^*\} \\
\leq\ & P_{\mu,\sigma}\left\{W_n \leq \eta\sigma \text{ for some } n \text{ s.t. } n_{1c} \leq n \leq n_{2c}\right\} \\
\leq\ & P_{\mu,\sigma}\left\{\max_{n_{1c}\leq n\leq n_{2c}} |W_n - \sigma| \geq (1-\eta)\,\sigma\right\} \\
\leq\ & ((1-\eta)\,\sigma)^{-r} E_{\mu,\sigma}\left|W_{n_{1c}} - \sigma\right|^r, \quad \text{by (C5)} \\
=\ & O\left(n_{1c}^{-r/2}\right) = O\left(n^{*-r/(2(1+\lambda))}\right), \quad \text{by (C6)}.
\end{aligned}$$

The proof is complete. $\square$

Theorem 3 (Asymptotic First-Order Risk Efficiency) *Define Risk Efficiency to be* $\xi(c) = R_N(c)/R_{n^*}(c)$*. Then, we have*

$$\lim_{c\to 0}\xi(c) = 1. \tag{12}$$

Proof: Under (C1), it is not hard to obtain that

$$E_{\mu,\sigma}\{n^*/N\} = E_{\mu,\sigma}\{J_1\} + E_{\mu,\sigma}\{J_2\}, \tag{13}$$

where $J_1 = \frac{n^*}{N} I\left(N > \frac{1}{2}n^*\right)$, $J_2 = \frac{n^*}{N} I\left(N \leq \frac{1}{2}n^*\right)$, and $I(A)$ stands for the indicator function of an event A.

We observe that $0 < J_1 < 2$ and a bounded random variable is uniformly integrable. Also, $J_1 \overset{P_{\mu,\sigma}}{\to} 1$ as $c \to 0$ in view of Theorem 1. Hence, $E_{\mu,\sigma}[J_1] = 1 + o(1)$ as $c \to 0$. Next, we handle the term $E_{\mu,\sigma}[J_2]$ and use 2 under (C5) and (C6) to express:

$$\begin{aligned}
E_{\mu,\sigma}[J_2] &\leq E_{\mu,\sigma}\left\{\frac{n^*}{n_{1c} I\left(N \leq \frac{1}{2}n^*\right)}\right\} \\
&= O(n^*)O(n^{*-1/(1+\lambda)})O(n^{*-r/2(1+\lambda)}) \\
&= O\left(n^{*\frac{2\lambda - r}{2(1+r)}}\right) \to 0, \text{ as } c \to 0,
\end{aligned}$$

as long as we pick some appropriate $r > \max\{2, 2\lambda\}$. Hence, (12) holds. $\square$

Lemma 4 *With* n^* *defined by (3), we have*

$$N^* \equiv (N - n^*)/N^{1/2} \overset{L}{\to} N(0, \delta^2), \text{ as } c \to 0. \tag{14}$$

Proof: We recall that $N/n^* \overset{P_{\mu,\sigma}}{\to} 1$ as $c \to 0$ under (C2). Having this settled, under (C3) and (C4), Anscombe's (1952) random central limit theorem would imply:

$$\begin{aligned}
& n^{*1/2}(\sigma^{-1}W_N - 1) \overset{L}{\to} N(0, \delta^2), \text{ and} \\
& n^{*1/2}(\sigma^{-1}W_{N-1} - 1) \overset{L}{\to} N(0, \delta^2) \text{ as } c \to 0,
\end{aligned} \tag{15}$$

with δ^2 coming from (C3). By Ghosh-Mukhopadhyay theorem (1975) and Slutsky's theorem, the result immediately follows from (15).

Lemma 5 *With* n^* *defined by (3), we have that* N^{*2} *is uniformly integrable for sufficiently small* $c \leq c_0$ *with some* $c_0 (> 0)$.

Proof: We proceed to prove this lemma in the spirit of Ghosh and Mukhopadhyay (1980) and Ghosh et al. (1997, Lemma 7.2.3, pp. 217-219) and shall first show that $(N - n^*)2/n^*$ is uniformly integrable in $c \leq c_0$ so that the desired result follows. Recall that $[u]$ denotes the largest integer that is smaller than $u (> 0)$.

We may write for any $b > b_0 + 1$, $b_0 = \left(\sigma\sqrt{A/c_1}\right)^{-1/2}$, where c_1 is some appropriate constant such that $c \leq c_1$. Then,

$$\begin{aligned}
& E_{\mu,\sigma}\left\{\frac{(N-n^*)^2}{n^*} I\left(\frac{(N-n^*)^2}{n^*} > b^2\right)\right\} \\
& = 2\int_b^\infty x P_{\mu,\sigma}\left(|N - n^*| > x\sqrt{n^*}\right) \mathrm{d}x.
\end{aligned} \tag{16}$$

Writing $k_1 = \left[n^* + x\sqrt{n^*}\right] + 1$ with $x \geq b$, we obtain:

$$\begin{aligned} & P_{\mu,\sigma}\left(N > n^* + x\sqrt{n^*}\right) \\ \leq & \leq P_{\mu,\sigma}\left(k_1 - 1 \leq W_{k_1-1}\sqrt{A/c}\right) \\ \leq & P_{\mu,\sigma}\left(\frac{W_{k_1-1}}{\sigma} \geq \frac{n^* + x\sqrt{n^*} - 1}{n^*}\right) \\ \leq & P_{\mu,\sigma}\left(\left|\frac{W_{k_1-1}}{\sigma} - 1\right| \geq \frac{x\sqrt{n^*} - 1}{n^*}\right) \\ \leq & \left(\frac{x}{\sqrt{n^*}} - \frac{1}{n^*}\right)^{-2r_1} E_{\mu,\sigma}\left\{\left|\frac{W_{k_1-1}}{\sigma} - 1\right|\right\}^{2r_1}. \end{aligned}$$

Under (C6), we claim that there exists a $\lambda_1\,(>0)$ depending only on r_1 such that

$$\begin{aligned} & P_{\mu,\sigma}\left(N > n^* + x\sqrt{n^*}\right) \\ & \leq \lambda_1 (k_1 - 1)^{-r_1} n^{*r_1} \left(x - 1/\sqrt{n^*}\right)^{-2r_1}. \end{aligned} \tag{17}$$

Note that $(k_1 - 1)^{-r_1} n^{*r_1} < 1$, for $x \geq b > b_0 + 1 = \left(\sigma\sqrt{A/c_1}\right)^{-1/2} + 1$ and $n^* = \sigma\sqrt{A/c} \geq \sigma\sqrt{A/c_1}$. It follows that

$$\begin{aligned} & \int_b^\infty x P_{\mu,\sigma}\left(|N - n^*| > x\sqrt{n^*}\right) \mathrm{d}x \\ & \leq \lambda_1 \int_b^\infty x (x - b_0)^{-2r_1} \mathrm{d}x \\ & = \lambda_1 \frac{(b - b_0)^{1-2r_1}\left((3 - 2r_1) b - b_0\right)}{(1 - 2r_1)(2 - 2r_1)} \to 0 \end{aligned} \tag{18}$$

as $b \to \infty$ by choosing $r_1 > 1$ appropriately.

Next, note that if $\sqrt{n^*} \leq b$,

$$\int_b^\infty x P_{\mu,\sigma}\left(N - n^* < -x\sqrt{n^*}\right) \mathrm{d}x = 0.$$

If $\sqrt{n^*} > b$, there exists some $0 < \gamma < 1$ such that $(1-\gamma)\sqrt{n^*} > b$, when $c \leq c_2$, for some technically picked c_2. Then,

$$\begin{aligned} & \int_b^\infty x P_{\mu,\sigma}\left(N - n^* < -x\sqrt{n^*}\right) \mathrm{d}x \\ & \leq \int_b^{\sqrt{n^*}} x P_{\mu,\sigma}\left(N \leq \gamma n^*\right) \mathrm{d}x \\ & + \int_b^{(1-\gamma)\sqrt{n^*}} x P_{\mu,\sigma}\left(\gamma n^* < N < n^* - x\sqrt{n^*}\right) \mathrm{d}x. \end{aligned} \tag{19}$$

By Lemma 2, $P_{\mu,\sigma}(N \leq \gamma n^*) \leq \lambda_2 n^{*^{-\frac{r_2}{2(1+\lambda)}}}$, for some appropriate $r_2\,(> 2 + 2\lambda)$ and $\lambda_2\,(>0)$ depending on r_2 alone. Hence,

$$\begin{aligned} & \int_b^{\sqrt{n^*}} x P_{\mu,\sigma}(N \leq \gamma n^*) \mathrm{d}x \\ & \leq \lambda_2 b^{2 - \frac{r_2}{(1+\lambda)}} \to 0 \text{ as } b \to \infty. \end{aligned} \tag{20}$$

As for $b \leq x \leq (1-\gamma)\sqrt{n^*}$, write

$$k_2 = [\gamma n^*] + 1 \text{ and } k_3 = \left[n^* - x\sqrt{n^*}\right].$$

We have that

$$\begin{aligned} & P_{\mu,\sigma}\left(\gamma n^* < N < n^* - x\sqrt{n^*}\right) \\ & = P_{\mu,\sigma}\left(\bigcup_{n=k_2}^{k_3} \{N = n\}\right) \\ & \leq P_{\mu,\sigma}\left(\bigcup_{n=k_2}^{k_3} \left\{\frac{W_n}{\sigma} < \frac{n}{n^*}\right\}\right). \end{aligned} \tag{21}$$

Note that for a small c, say $c \leq c_3$, for some c_3,

$$\frac{n}{n^*} \leq \frac{k_3}{n^*} \leq \frac{n^* - x\sqrt{n^*}}{n^*} = 1 - \frac{x}{\sqrt{n^*}}.$$

Hence, it follows from (21) that

$$\begin{aligned} & P_{\mu,\sigma}\left(\gamma n^* < N < n^* - x\sqrt{n^*}\right) \\ & \leq P_{\mu,\sigma}\left(\frac{W_n}{\sigma} - 1 < -\frac{x}{\sqrt{n^*}}, \text{ for some } k_2 \leq n \leq k_3\right) \\ & \leq P_{\mu,\sigma}\left(\left|\frac{W_n}{\sigma} - 1\right| > \frac{x}{\sqrt{n^*}}, \text{ for some } k_2 \leq n \leq k_3\right) \\ & \leq P_{\mu,\sigma}\left(\max_{k_2 \leq n \leq k_3}\left|\frac{W_n}{\sigma} - 1\right| > \frac{x}{\sqrt{n^*}}\right) \\ & \leq \frac{E_{\mu,\sigma}\left|\frac{W_{k_2-1}}{\sigma} - 1\right|^{2r_3}}{\left(x/\sqrt{n^*}\right)^{2r_3}} \\ & \leq \lambda_3 k_2^{-r_3} x^{-2r_3} n^{*r_3} \leq \lambda_4 x^{-2r_3}, \end{aligned} \tag{22}$$

for some appropriate $\lambda_3\,(>0)$ and $\lambda_4\,(>0)$, both depending only on r_3. Choosing $r_3 > 1$, we get from (22) that

$$\begin{aligned} & \int_b^{(1-\gamma)\sqrt{n^*}} x P_{\mu,\sigma}\left(\gamma n^* < N < n^* - x\sqrt{n^*}\right) \mathrm{d}x \\ & \leq \lambda_4 \int_b^{(1-\gamma)\sqrt{n^*}} x^{1-2r_3} \mathrm{d}x \leq \frac{\lambda_4}{2 - 2r_3} b^{2-2r_3} \to 0 \end{aligned} \tag{23}$$

as $b \to \infty$.

Now choosing $c_0 = \min\{c_1, c_2, c_3\}$, with (18), (20) and (23), we can prove the uniform integrability of $(N - n^*)^2 / n^*$ in $c \le c_0$. To complete the proof of the lemma, observe that

$$\begin{aligned} & E_{\mu,\sigma}\left\{ \frac{(N-n^*)^2}{n^*} I\left(\frac{(N-n^*)^2}{N} > b^2 \right) \right. \\ & \quad \left. \times I\left(N > \frac{1}{2}n^* \right) \right\} \\ & \le 2E_{\mu,\sigma}\left\{ \frac{(N-n^*)^2}{n^*} I\left((N-n^*)^2 > \frac{1}{2}b^2 n^* \right) \right\} \\ & \quad \to 0 \text{ as } b \to \infty \text{ uniformly in } c \le c_0. \end{aligned} \tag{24}$$

Furthermore, choosing $b > n^*$, it follows that for $c \le c_0$,

$$\begin{aligned} & E_{\mu,\sigma}\left\{ \frac{(N-n^*)^2}{n^*} I\left(\frac{(N-n^*)^2}{N} > b^2 \right) \right. \\ & \quad \left. \times I\left(N \le \frac{1}{2}n^* \right) \right\} \\ & \le m^{-1} {n^*}^2 P_{\mu,\sigma}\left(N \le \frac{1}{2}n^* \right) \le \lambda_5 {n^*}^{2-r_4} \le \lambda_5 b^{2-r_4} \\ & \quad \to 0 \text{ as } b \to \infty, \end{aligned} \tag{25}$$

with $r_4 (> 2)$ chosen appropriately and some $\lambda_5\ (> 0)$ depending only on r_4. In view of (24) and (25), the intended result holds. □

Following from Lemma 4 and Lemma 5, we now proceed with the *asymptotic second-order regret* property of the purely sequential MRPE methodology from (4)-(5), given in the theorem below.

Theorem 6 (Asymptotic Second-Order Regret) *Define Regret to be $\omega(c) = R_N(c) - R_{n^*}(c)$. We have as $c \to 0$:*

$$\omega(c) = \delta^2 c + o(c), \tag{26}$$

with δ^2 coming from (C3).

4 Illustrations

In this Section, we provide possible applications with W_n in the stopping rule (4) substituted by *Sample standrad deviation*, Gini's Mean Difference (GMD), and *Mean Absolute Deviance* (MAD), respectively. Corresponding asymptotically second-order regret purely sequential point estimation methodologies are proposed thereby.

(a) *Sample standard deviation*. A customarily used unbiased estimator for the unknown normal mean is the sample variance, denoted by S_n^2, where

$$S_n^2 = (n-1)^{-1} \Sigma_{i=1}^n (X_i - \overline{X}_n)^2.$$

Naturally, the sample standard deviation S_n is used to estimate the population standard deviation σ. Hence, we consider to substitute W_n with S_n and propose the sequential methodology with the following stopping boundary condition:

$$\begin{aligned} N_0 &\equiv N_0(c) \\ &= \inf\left\{ n \ge m : n \ge \sqrt{A/c}\left(S_n + n^{-\lambda} \right) \right\}. \end{aligned} \tag{27}$$

Clearly, conditions (C1)-(C7) are all satisfied in terms of S_n, so we develop the regret approximation:

$$\omega_0(c) = \frac{1}{2}c + o(c) \text{ as } c \to 0, \tag{28}$$

as a fact of that

$$\sqrt{n}\,(S_n/\sigma - 1) \xrightarrow{L} N\left(0, \frac{1}{2}\right) \text{ as } n \to \infty.$$

One may see Mukhopadhyay and de Silva (2009) and other sources for reference.

(b) *Gini's Mean Difference* (GMD). As a robust estimator of the population standard deviation σ, GMD is originally developed in Gini (1914, 1921), defined as follows:

$$g_n = \binom{n}{2}^{-1} \Sigma\Sigma_{1 \le i < j \le n} |X_i - X_j|. \tag{29}$$

Under the normal assumption, we can accordingly construct the unbiased and consistent estimator based on GMD denoted by G_n, where

$$G_n = \frac{\sqrt{\pi}}{2} g_n. \tag{30}$$

One should notice that G_n is indeed a U-Statistic. As a result, conditions (C1)-(C7) automatically hold. See Hoeffding (1948, 1961), Sen and Ghosh (1981), Lee (1990), Mukhopadhyay and Hu (2017, 2018), Hu and Mukhopadhyay (2019) and etc. for more details. Thus, we consider the following stopping rule:

$$\begin{aligned} N_1 &\equiv N_1(c) \\ &= \inf\left\{ n \ge m : n \ge \sqrt{A/c}\left(G_n + n^{-\lambda} \right) \right\}. \end{aligned} \tag{31}$$

With the fact that

$$\sqrt{n}\left(G_n/\sigma-1\right) \xrightarrow{L} N\left(0, \frac{\pi+6\sqrt{3}-12}{3}\right)$$

as $n \to \infty$, we conclude the asymptotic second-order regret below:

$$\omega_2(c) = \frac{\pi+6\sqrt{3}-12}{3}c + o\left(c\right) \text{ as } c \to 0. \quad (32)$$

(c) *Mean Absolute Deviance* (MAD). MAD is another robust estimator for σ as a counterpart of the traditional sample standard deviation. Denoted by m_n, MAD is defined as follows:

$$m_n = n^{-1}\Sigma_{i=1}^{n}\left|X_i - \overline{X}_n\right|. \quad (33)$$

Again we construct the unbiased and consistent estimator for σ given below:

$$M_n = \sqrt{\frac{\pi n}{2\left(n-1\right)}}m_n. \quad (34)$$

To verify (C1)-(C7), one may refer to Babu and Rao (1992), Mukhopadhyay and Hu (2017, 2018) and Hu and Mukhopadhyay (2019). Hence, we substitute W_n with M_n and develop the purely sequential MRPE methodology with the stopping rule given by

$$\begin{aligned} N_2 &\equiv N_2(c) \\ &= \inf\left\{n \geq m : n \geq \sqrt{A/c}\left(M_n + n^{-\lambda}\right)\right\}. \end{aligned} \quad (35)$$

Based on the result that

$$\sqrt{n}\left(M_n/\sigma-1\right) \xrightarrow{L} N\left(0, \frac{\pi-2}{2}\right) \text{ as } n \to \infty,$$

we claim that the corresponding regret approximation is

$$\omega_2\left(c\right) = \frac{\pi-2}{2}c + o\left(c\right) \text{ as } c \to 0. \quad (36)$$

5 Simulations

In the spirit of Mukhopadhyay and Hu (2017), we implement the purely sequential minimum risk point estimation methodologies based on various stopping rules given by (27), (31), and (35) respectively in the normal case. To be more specific, we generate pseudorandom samples from a $N\left(5,4\right)$ population; and we also fix the weight function $A = 100$, the pilot sample size $m = 10$, and $\lambda = 2$, while selecting a wide range of values of c including $0.16, 0.04, 0.01$, and 0.025 so that the optimal sample sizes n^* can be determined to be $50, 100, 200$, and 400 accordingly by (3). Throughout the section, we are using the following system codes to indicate each specific methodology under implementation:

$i = 0$: S_n-based stopping rule (27);
$i = 1$: G_n-based stopping rule (31);
$i = 2$: M_n-based stopping rule (35).

Table 1: Simulations from $N(5,4)$ with $A = 100$, $m = 10$ and $\lambda = 2$ under 1000 runs

n^*	$100c$	i	$\overline{n}$	$s\left(\overline{n}\right)$	$\widehat{\xi}$	$\widehat{\omega}/c$
50	16	0	50.01	0.167	0.988	0.593
		1	50.31	0.170	0.988	0.612
		2	50.26	0.178	0.987	0.666
100	4	0	99.95	0.241	0.993	0.600
		1	100.33	0.235	0.994	0.561
		2	100.33	0.249	0.994	0.636
200	1	0	200.01	0.332	0.997	0.561
		1	200.25	0.338	0.997	0.580
		2	200.03	0.356	0.996	0.644
400	0.25	0	399.93	0.459	0.998	0.531
		1	400.28	0.451	0.998	0.514
		2	400.23	0.487	0.998	0.599

Table 1 presents the simulated performance under 1000 independent replications of runs. As reflected in the table, the average of estimated sample sizes are close to the optimal sample size. Furthermore, all the sequential methodologies (27), (31), and (35) enjoy the asymptotic second-order regret, and the estimated regrets provided in the last column are comparable to the theoretical values given in (28), (32), and (36), respectively for $i = 0, 1, 2$.

6 Saving Sampling Operations

Not surprisingly, the purely sequential sampling methodologies require a lot of sampling operations, which may put a damper. As is pointed out in Hu and Zhuang (2019), we can accelerate the proposed new class of purely sequential MRPE methodologies without sacrificing the first- and second-order properties in two directions: (i) to sample $k(\geq 2)$ observations instead of one observation at-a-time successively; (ii) to proceed purely sequential sampling to determine only a proportion $\rho(0 < \rho < 1)$ of the desired final sample, followed by a batch of the remaining observations gathered in one step. Combining these ideas, we expect to save roughly $100(1-k^{-1}\rho)\%$ of sampling operations.

Therefore, we proposed a class of accelerated sequential MRPE methodologies accordingly. Given the pilot sample size $m \geq 2$, and the prefixed quantities $0 < \rho \leq 1$ and $k \geq 1$, we develop the following stopping rule:

$$\begin{aligned} T_1 \equiv & T_1(c) = \inf\{n \geq 0 : m + kn \\ & \geq \rho\sqrt{A/c}\left(W_{m+kn} + (m+kn)^{-\lambda}\right)\}, \end{aligned} \tag{37}$$

and the final sample size is given by

$$T_2 \equiv T_2(c) = [\rho^{-1}(m + kT) + 1], \tag{38}$$

where $[u]$ continues to represent the largest integer smaller than $u(> 0)$.

Apparently, $P_{\mu,\sigma}(T_2 < \infty) = 1$ and $T_2 \uparrow \infty$ w.p.1 as $c \downarrow 0$. Now, we are in a position to provide the following crucial results, as summarized in Theorem 7.

Theorem 7 *Define Risk Efficiency and Regret associated with the accelerated sequential MRPE methodologies (37)-(38) to be $\xi^*(c) = R_{T_2}(c)/R_{n^*}(c)$ and $\omega^*(c) = R_{T_2}(c) - R_{n^*}(c)$, respectively. We have*

(i) Asymptotic First-Order Efficiency:

$$\lim_{c \to 0} E_{\mu,\sigma}[T_2/n^*] = 1. \tag{39}$$

(ii) Asymptotic First-Order Risk Efficiency:

$$\lim_{c \to 0} \xi^*(c) = 1. \tag{40}$$

(iii) Asymptotic Second-Order Regret:

$$\omega^*(c) = \rho^{-1}\delta^2 c + o(c) \text{ as } c \to 0, \tag{41}$$

where δ^2 coming from (C3).

The proof of Theorem 7 can be done in a similar way as we proved Theorems 1-6. We leave out many details for brevity. It is worth mentioning that it should surprise no one that $\omega^*(c)$ in (41) is larger than $\omega(c)$ in (26). The increased regret under accelerated sequential sampling is a straight result of saving sampling operations. One should balance the operational convenience and the increased regret for practical purposes.

References:

[1] F. J. Anscombe, Large-Sample Theory of Sequential Estimation, *Proceedings of Cambridge Philosophical Society.* 48, 1952, pp. 600-607.

[2] F. J. Anscombe, Sequential Estimation, *Journal of Royal Statistical Society, Series B.* 15, 1953, pp. 1-29.

[3] G. J. Babu and C. R. Rao, Expansions for Statistics Involving the Mean Absolute Deviations, *Annals of the Institute of Statistical Mathematics.* 44, 1992, pp. 387-403.

[4] B. Chattopadhyay and N. Mukhopadhyay, Two-Stage Fixed-Width Confidence Intervals for a Normal Mean in the Presence of Suspect Outliers, *Sequential Analysis.* 32, 2013, pp. 134-157.

[5] Y. S. Chow and H. Robbins, On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean, *Annals of Mathematical Statistics.* 36, 1965, pp. 457-462.

[6] M. Ghosh and N. Mukhopadhyay, Sequential Point Estimation of the Difference of Two Normal Means, *Annals of Statistics.* 8, 1980, pp. 221-225.

[7] M. Ghosh, N. Mukhopadhyay and P. K. Sen, *Sequential Estimation*, Wiley–New York 1997.

[8] C. Gini, *L'Ammontare la Composizione della Ricchezza delle Nazioni*, Boca–Torino 1914.

[9] C. Gini, Measurement of Inequality of Incomes, *Economic Journal.* 31, 1921, pp. 124-126.

[10] W. Hoeffding, A Class of Statistics with Asymptotically Normal Distribution, *Annals of Mathematical Statistics.* 19, 1948, pp. 293-325.

[11] W. Hoeffding, The Strong Law of Large Numbers for U-Statistics, Institute of Statistics Mimeo Series #302, University of North Carolina, Chapel Hill.

[12] J. Hu and N. Mukhopadhyay, Second-order Asymptotics in a Class of Purely Sequential Minimum Risk Point Estimation (MRPE) Methodologies, *Japanese Journal of Statistics and Data Science.* 2, 2019, pp. 81-104.

[13] J. Hu and Y. Zhuang, A Generalized Sequential Sampling Design with Fewer Sampling Operations, *Submitted.* 2019.

[14] A. J. Lee, *U-statistics. Theory and Practice*, Dekker–New York 1990.

[15] N. Mukhopadhyay, Steins Two-Stage Procedure and Exact Consistency, *Scandinavian Actuarial Journal.* 1982, pp. 110-122.

[16] N. Mukhopadhyay and B. M. de Silva *Sequential Methods and Their Applications*, CRC–Boca Raton 2009.

[17] N. Mukhopadhyay and J. Hu, Confidence Intervals and Point Estimators for a Normal Mean under Purely Sequential Strategies Involving Gini's Mean Difference and Mean Absolute Deviation, *Sequential Analysis.* 36, 2017, pp. 210-239.

[18] N. Mukhopadhyay and J. Hu, Gini's Mean Difference and Mean Absolute Deviation Based Two-Stage Estimation for a Normal Mean with Known Lower Bound of Variance, *Sequential Analysis.* 37, 2018, pp. 204-221.

[19] W. D. Ray, Sequential Confidence Intervals for the Mean of a Normal Population with Unknown Variance, *Journal of Royal Statistical Society, Series B.* 19, 1957, pp. 133-143.

[20] H. Robbins, Sequential Estimation of the Mean of a Normal Population, *Probability and Statistics, H Cramer volume, Ulf Grenander, ed.* 1959, pp. 235-245.

[21] P. K. Sen and M. Ghosh, Sequential Point Estimation of Estimable Parameters Based on U-Statistics, *Sankhyā, Series A.* 43, 1981, pp. 331-344.

[22] N. Starr, On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean, *Annals of Mathematical Statistics.* 37, 1966, pp. 1173-1185.

[23] N. Starr and M. Woodroofe, Remarks on Sequential Point Estimation, *Proceedings of the National Academy of Sciences.* 63, 1969, pp. 285-288.

[24] M. Woodroofe, Second Order Approximations for Sequential Point and Interval Estimation, *Annals of Statistics.* 5, 1977, pp. 984-995.