Taylor & Francis
Taylor & Francis Group

Sequential Analysis
Design Methods and Applications

ISSN: 0747-4946 (Print) 1532-4176 (Online) Journal homepage: http://www.tandfonline.com/loi/lsqa20

Multistage point estimation methodologies for a negative exponential location under a modified linex loss function: Illustrations with infant mortality and bone marrow data

Nitis Mukhopadhyay & Sudeep R. Bapat

To cite this article: Nitis Mukhopadhyay & Sudeep R. Bapat (2016) Multistage point estimation methodologies for a negative exponential location under a modified linex loss function: Illustrations with infant mortality and bone marrow data, Sequential Analysis, 35:2, 175-206

To link to this article: http://dx.doi.org/10.1080/07474946.2016.1165532

Published online: 27 Jun 2016.

SEQUENTIAL ANALYSIS
2016, VOL. 35, NO. 2, 175–206
http://dx.doi.org/10.1080/07474946.2016.1165532

Taylor & Francis
Taylor & Francis Group

Multistage point estimation methodologies for a negative exponential location under a modified linex loss function: Illustrations with infant mortality and bone marrow data

Nitis Mukhopadhyay and Sudeep R. Bapat

Department of Statistics, University of Connecticut, Storrs, Connecticut, USA

ABSTRACT

We have designed Stein-type (Stein, 1945, *Annals of Mathematical Statistics*) two-stage, modified two-stage (Mukhopadhyay and Duggan, 1997, *Sankhya, Series A*), and purely sequential strategies (Chow and Robbins, 1965, *Annals of Mathematical Statistics*) to estimate an unknown location parameter of a negative exponential distribution having an unknown scale parameter under a newly defined and modified Linex loss function. We aim at controlling the associated risk function per unit cost by bounding it from above with a fixed preassigned positive number, ω, and we emphasize both asymptotic first-order and asymptotic second-order properties for the modified two-stage and purely sequential estimation strategies. In developing asymptotic second-order properties for the modified two-stage methodology, we have heavily relied upon basic ideas rooted in Mukhopadhyay and Duggan (1997). In developing asymptotic second-order properties for the purely sequential methodology, however, we have heavily relied upon nonlinear renewal theory (Lai and Siegmund, 1977, 1979, *Annals of Statistics*; Woodroofe, 1977, *Annals of Statistics*). Then, we take to extensive data analysis carried out via computer simulations when requisite sample sizes range from small to moderate to large. We find that the Stein-type two-stage estimation methodology oversamples significantly and yet the achieved risk is not close to preset goal ω. On the other hand, both modified two-stage and purely sequential estimation strategies perform remarkably well. We have validated their main theoretical first-order and second-order properties through simulated data. The latter methodologies have been illustrated and implemented using two real data sets from health studies, namely, infant mortality data and bone marrow data.

ARTICLE HISTORY
Received 1 October 2015
Revised 15 January 2016
Accepted 16 February 2016

KEYWORDS
Bone marrow data; cancer research; first-order properties; infant mortality data; Linex loss; location parameter; modified linex loss; modified two-stage; negative exponential; nonlinear renewal theory; purely sequential; real data; risk; risk per unit cost; scale parameter; second-order properties; simulations; two-stage

MATHEMATICS SUBJECT CLASSIFICATION
62L12; 62L05; 62G20; 62F10; 62P10; 62P30

1. Introduction

We develop location parameter estimation methodologies for a negative exponential distribution under a modified Linex loss function. We consider a negative exponential distribution having the following *probability density function* (p.d.f.):

$$f(x;\mu,\sigma) = \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right) I(x > \mu), \tag{1.1}$$

CONTACT Nitis Mukhopadhyay nitis.mukhopadhyay@uconn.edu Department of Statistics, University of Connecticut, Austin Building U4120, 215 Glenbrook Road, Storrs, CT 06269-4120, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsqa.
Recommended by Tumulesh K. S. Solanky

where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$ are *unknown parameters*. $I(.)$ denotes an indicator function of (.) that takes the value 1 (or 0) when $x > (\text{or} \leq)\mu$. This distribution is also known as a two-parameter exponential distribution. The parameter μ, if positive, may be interpreted as the minimum guarantee time or the threshold of the distribution in the sense that no failure will occur before μ. The parameter σ is called a scale.

The p.d.f. (1.1) has found its use in many reliability problems such as to depict failure times of electrical components and complex equipment. One may refer to Johnson and Kotz (1970), Bain (1978), and Balakrishnan and Basu (1995) for illustrations. Another area where it has been used is clinical trials such as in studying the behavior of tumor systems in animals and analysis of survival data in cancer research. One may refer to Zelen (1966). We provide illustrations with real data analysis in Section 8.

Before we go any further, we explain our notation clearly. Because both μ, σ are unknown, the parameter vector $\boldsymbol{\theta} = (\mu, \sigma)$ remains unknown. When we write $P(.)$ or $E(.)$, they should be interpreted as $P_{\boldsymbol{\theta}}(.)$ or $E_{\boldsymbol{\theta}}(.)$ respectively. In the same spirit, when we write $\xrightarrow{P}$ (convergence in probability) or w.p.1 (with probability one) or $\xrightarrow{\mathcal{L}}$ (convergence in law or distribution), they are all with respect to $P_{\boldsymbol{\theta}}$. We drop subscript $\boldsymbol{\theta}$ for simplicity.

We address methodologies for estimating μ under a *variant* of a customary Linex loss function defined as follows:

$$L_n \equiv L_n(\hat{\mu}_n, \mu) = \exp\left(\frac{a(\hat{\mu}_n - \mu)}{\sigma}\right) - \frac{a(\hat{\mu}_n - \mu)}{\sigma} - 1, \tag{1.2}$$

where $\hat{\mu}_n$ is meant to be a generic estimator of μ based on a random sample of size n. This modified Linex loss function (1.2) is different from the one that was first proposed by Varian (1975), namely,

$$\exp\big(a(\hat{\mu}_n - \mu)\big) - a(\hat{\mu}_n - \mu) - 1. \tag{1.3}$$

The Linex loss (1.3) was an appropriate function to be considered in cases of an asymmetric penalty due to bias. It was supposed to address estimation error by penalizing overestimation and underestimation unequally where overestimation is deemed more (less) serious than underestimation when $a > (<)0$.

Varian (1975) and Zellner (1986) popularized Linex loss (1.3) and brought it to the forefront of statistical science with interesting applications. We will explain in Section 2.1 why we have replaced (1.3) with (1.2) and then propose to work under this modified Linex loss (1.2). We may add that under a squared error loss function (or its variant), sequential inference problems devoted to negative exponential distributions can be found in Mukhopadhyay (1974, 1984, 1988, 1995) and other sources.

A sequential point estimation problem under a Linex loss function (1.3) was first developed by Chattopadhyay (1998) utilizing nonlinear renewal theory from Woodroofe (1977, 1982) and Lai and Siegmund (1977, 1979). Methodologies pertaining to Linex loss (1.3) were developed by Chattopadhyay (2000), Chattopadhyay et al. (2000, 2005), Takada (2000, 2006), Takada and Nagao (2004), Chattopadhyay et al. (2005), Chattopadhyay and Sengupta (2006), Zacks and Mukhopadhyay (2009), and others. Many researchers obtained sequential point estimation methods under the Linex loss (1.3) largely under a normal distribution. There is a substantial literature available on Bayes sequential estimation problems under Linex loss (1.3) that may be reviewed from Takada (2001), Jokiel-Rokita (2008), Hwang and Lee (2012), and other sources.

Second- and higher-order approximations were developed by Mukhopadhyay and Duggan (1997) and Mukhopadhyay (1999) under a variety of loss functions in the context of appropriately modified two-stage estimation strategies. For a review of two-stage, purely sequential, and other kinds of stopping rules, one may refer to Sen (1981), Woodroofe (1982), Siegmund (1985), Mukhopadhyay and Solanky (1994), Ghosh et al. (1997), Mukhopadhyay et al. (2004), Mukhopadhyay and de Silva (2009), Zacks (2009), and other sources.

In Section 2, we include some preliminaries plus an explanation for the modified Linex loss (1.2), formulation of the risk function, and its optimization by bounding it from above. In Section 3, we consider a Stein-type Stein (1945, 1949) two-stage procedure where one first gathers pilot data followed by the remaining requisite data determined by the final sample size. We derive some asymptotic properties for this two-stage estimation strategy (Theorem 3.1). Some technical details follow from Chow and Robbins (1965) and Mukhopadhyay (1984, 1988, 1995).

Section 4 introduces a modified two-stage procedure in which we assume a lower bound $\sigma_L(>0)$ for the standard deviation σ along the lines of Mukhopadhyay and Duggan (1997, 1999). We develop both first-order (Theorem 4.1) as well as second-order (Theorems 4.2–4.4) properties in the present scenario of a negative exponential setting under the modified Linex loss (1.2). Higher-order approximations may be reviewed from Isogai et al. (2011) and seen as natural extensions of Mukhopadhyay and Duggan (1997) and Mukhopadhyay (1999).

Section 5 develops a purely sequential methodology followed by its asymptotic first-order properties (Theorem 5.1). Then, we exploit nonlinear renewal theory from Woodroofe (1977, 1982) and Lai and Siegmund (1977, 1979) to obtain requisite second-order approximations (Theorems 5.2–5.3). For some of the technicalities and second-order properties, we have referred to Mukhopadhyay (1974, 1984, 1988), Lombard and Swanepoel (1978), Swanepoel and van Wyk (1982), and others.

In Section 6, we outline the proofs of selected major results. Section 7 highlights performances of the proposed estimation methodologies obtained with the help of computer simulations for a wide range of values of the sample sizes and a large variety of parameter configurations. Our presented data analyses are both extensive and thorough. Selected conclusions from the theorems studied in Sections 3–6 are critically examined and validated with data analysis in Section 7.

These are supplemented with illustrations (Section 8) using two real data sets from health studies: The first illustration (Section 8.1) uses infant mortality data available from Leinhardt and Wasserman (1979). Our second illustration (Section 8.2) uses bone marrow transplant data that came from a multicenter clinical trial with patients prepared for transplantation with a radiation-free conditioning regimen (Klein and Moeschberger, 2003). We end with a brief set of concluding thoughts (Section 8.3).

2. Modified Linex loss and some preliminaries

In this section, we first introduce an appropriately modified Linex loss function (1.2) and then calculate the associated risk function. But, why it is that we must modify the customary Linex loss function (1.3)? Let us explain.

Having recorded a random sample $X_1, \ldots, X_n$ of *independent and identically distributed* (i.i.d.) observations from a negative exponential distribution (1.1), one may customarily estimate μ by the maximum likelihood estimator $X_{n:1}$, the smallest order statistic. Suppose

that one begins with the original Linex loss function (1.3) from Varian (1975) and Zellner (1986):

$$\exp\left(a(X_{n:1}-\mu)\right)-a(X_{n:1}-\mu)-1, n\geq 1.$$

Now, the associated risk function requires evaluation of $E[\exp\left(a(X_{n:1}-\mu)\right)]$ which is alternatively expressed as $E\left[\exp\left(\frac{a\sigma}{n}Y\right)\right]$ where $Y\sim\exp(1)$, a standard exponential distribution. But, this particular term, $E\left[\exp\left(\frac{a\sigma}{n}Y\right)\right]$, will be finite provided that $n>a\sigma$. However, because the scale parameter σ remains unknown, there is no way for one to guarantee that a sample size n will certainly exceed $a\sigma$ when $a>0$. Even if we have some reasonable estimator $\hat{\sigma}$ of σ available and n exceeds $a\hat{\sigma}$, it will not guarantee that our sample size n will exceed $a\sigma$ when $a>0$. Moreover, under two-stage and purely sequential sampling, one will require $n>a\sigma$ for all fixed n which will be hard to guarantee if a is positive.

Remark 2.1. If "a" is assumed negative, however, then one may continue to work under the customary Linex loss:

$$\exp\left(a(X_{n:1}-\mu)\right)-a(X_{n:1}-\mu)-1, n\geq 1,$$

without any further modification. We leave out associated details for brevity. In what follows, however, we chart a new direction by providing a unified treatment under (2.1) whether a is positive or negative.

2.1. Modified Linex loss function

We focus on working with fixed a without distinguishing whether a is exclusively positive or negative. Hence, we develop sampling methodologies for estimating μ by $X_{n:1}$ under the *modified Linex loss* function formulated as follows. Instead of (1.2), we rewrite it:

$$L_n\equiv L_n(X_{n:1},\mu)=\exp\left(\frac{a(X_{n:1}-\mu)}{\sigma}\right)-\frac{a(X_{n:1}-\mu)}{\sigma}-1, \tag{2.1}$$

where a is a constant. We obtain the corresponding risk function by taking expectations across (2.1) as follows:

$$\begin{aligned}\text{Risk}_n\equiv E[L_n]&=E\left[\exp\left(\frac{a(X_{n:1}-\mu)}{\sigma}\right)-\frac{a(X_{n:1}-\mu)}{\sigma}-1\right]\\&=\left(1-\frac{a}{n}\right)^{-1}-\frac{a}{n}-1,\quad \text{for } n>a.\end{aligned} \tag{2.2}$$

Upon expanding (2.2), we clearly obtain

$$\text{Risk}_n=\frac{a^2}{n^2}+o\left(\frac{1}{n^2}\right). \tag{2.3}$$

2.2. Cost per unit sampling

Next, what we consider is the cost function, $\text{Cost}_n(>0)$, the exact nature of which ought to depend upon the problem on hand. But, it is reasonable to assume that the cost for each observation should go up (or down) as σ goes down (or up). With this understanding, we

propose a cost function of the following form:

$$\text{Cost}_n = cn\sigma^{-k} \text{ with fixed and known } c(>0), k(>0). \tag{2.4}$$

2.3. Proposed criterion: Bound the risk per unit cost from above

We wish to bound the associated "risk" from above where we interpret "risk" as the *risk per unit cost* (RPUC), namely,

$$\text{RPUC}_n \equiv \frac{\text{Risk}_n}{\text{Cost}_n} = \frac{a^2}{n^2}\frac{\sigma^k}{cn} + o\left(n^{-3}\right). \tag{2.5}$$

Thus, we fix a suitable number $\omega(>0)$ and require that $\text{RPUC}_n \le \omega$ for all μ, σ which leads us to determine the required optimal fixed sample size, had σ been known as follows:

$$n \ge \left(\frac{a^2}{c\omega}\right)^{1/3} \sigma^{k/3} = n^*, \text{ say.} \tag{2.6}$$

But, the magnitude of n^* is unknown even though its expression is known. Hence, we proceed to develop two-stage, modified two-stage, and purely sequential bounded risk estimation strategies in Sections 3–5.

Now, having recorded data $X_1, \ldots, X_n$ of fixed size n, we denote $\hat{\sigma}_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X_{n:1})$ as our customary minimum variance unbiased estimator of σ with $n(>\max\{1, a\})$.

2.4. Evaluating the risk per unit cost when sample size is random

Clearly, the optimal fixed sample size n^* given by (2.6) needs to be estimated. We must begin with pilot data of appropriate size $m(>\max\{1, a\})$ and then move forward step by step with the help of implementing a two-stage, modified two-stage, or purely sequential sampling strategy to record more data subsequently beyond pilot data as needed.

Suppose that a final sample size, denoted by a random variable Q, is determined by an adaptive multistage sampling strategy. Then, the next theorem shows an exact analytical expression for the risk per unit cost associated with the terminal estimator $X_{Q:1} = \min\{X_1, \ldots, X_Q\}$ of μ once sampling has stopped. Its proof is outlined in Section 6.1.

Theorem 2.1. *Under a multistage estimation strategy, suppose that (i) the final sample size Q is an observable random variable that is finite w.p.1, and (ii) Q is determined in such a way that the event $Q = q$ is measurable with respect to $\{\hat{\sigma}_j; m \le j \le q\}$, for all fixed $q \ge m(>\max\{1, a\})$. Then, the expression for the risk per unit cost associated with the terminal estimator $X_{Q:1}$ is given by*

$$E\left[RPUC_Q\right] = \omega\frac{n^{*3}}{a^2}\left\{E\left[\frac{1}{Q}\left(1 - \frac{a}{Q}\right)^{-1}\right] - E\left(\frac{a}{Q^2}\right) - E\left(\frac{1}{Q}\right)\right\}. \tag{2.7}$$

3. Stein-type two-stage procedure

A Stein-type two-stage procedure along the lines of Stein (1945, 1949) may be logistically convenient to implement in certain situations because we may collect data in two batches. At the first stage, we collect pilot data $X_1, \ldots, X_m$ of size $m(>\max\{1, a\})$. Recall that

$\hat{\sigma}_m = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - X_{m:1})$ is an estimator of σ obtained from pilot data. We define the stopping rule as

$$N = \max\left\{m, \left\lfloor d_\omega \hat{\sigma}_m^{k/3} \right\rfloor + 1\right\} \text{ with } d_\omega = \left(a^2(c\omega)^{-1}\right)^{1/3}, \tag{3.1}$$

as an estimator of n^* found in (2.6) where

$$\lfloor u \rfloor \text{ denotes the largest integer less than } u(>0).$$

Now, if $N = m$, then we will not require any more data at the second stage. However, if $N > m$, we sample the difference $N - m$ at the second stage by recording an additional set of observations $X_{m+1}, \ldots, X_N$. From full data obtained by combining both stages, we propose to estimate μ by the smallest order statistic:

$$X_{N:1} = \min\{X_1, \ldots, X_N\}.$$

Along the lines of (2.1), the associated loss function will be:

$$L_N = \exp\left(\frac{a(X_{N:1} - \mu)}{\sigma}\right) - \frac{a(X_{N:1} - \mu)}{\sigma} - 1. \tag{3.2}$$

A major difference between (2.1) and (3.2) is that the sample size N used in (3.2) is a random variable unlike n.

Now, since $X_{n:1}$ and $I(N = n)$ are independent for all fixed $n \geq m$, using (2.7) with Q replaced by N from (3.1), we get

$$\omega^{-1}E\left[\text{RPUC}_N\right] = \frac{n^{*3}}{a^2}\left\{E\left[\frac{1}{N}\left(1 - \frac{a}{N}\right)^{-1}\right] - E\left(\frac{a}{N^2}\right) - E\left(\frac{1}{N}\right)\right\}. \tag{3.3}$$

Theorem 3.1. *For the two-stage methodology (3.1), for fixed values of μ, σ, c, k, a, m we have as $\omega \to 0$:*

(i) $N/n^* \xrightarrow{P} \left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}$; $n^*/N \xrightarrow{P} \left(\frac{\sigma}{\hat{\sigma}_m}\right)^{k/3}$ *if* $m > \max\{1, a\}$;

(ii) $E\left(N/n^*\right) \to E\left[\left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}\right] (> 1)$ *if* $m > \max\{a, 1, 1 - \frac{1}{3}k\}$;

where n^* comes from (2.6).

Theorem 3.1(ii) shows that the two-stage methodology (3.1) oversamples on an average compared with n^* even asymptotically. Such a feature has been observed in the past under numerous Stein-type two-stage estimation strategies. A proof of Theorem 3.1 will be sketched in Section 6.2.

4. Modified two-stage procedure

From Theorem 3.1, we note that the ratios $E\left(N/n^*\right)$ and $E\left(n^*/N\right)$ did not converge to 1 under the Stein-type two-stage methodology (3.1). We thus resort to a modified two-stage procedure along the lines of Mukhopadhyay and Duggan (1997).

A key idea is to introduce a lower bound σ_L such that $0 < \sigma_L < \sigma$ with σ_L known. Given this additional input, from the expression of n^* found in (2.6), we note that $n^* > \left(a^2(c\omega)^{-1}\right)^{1/3}\sigma_L^{k/3}$. Thus, the pilot size m may be chosen in such a way that $m \approx \left(a^2(c\omega)^{-1}\right)^{1/3}\sigma_L^{k/3}$.

With that spirit, Mukhopadhyay and Duggan (1997) introduced a modification for estimating a normal mean that allowed them to study second-order properties of the associated two-stage estimation strategy. Such a modification has been widely adopted in the literature. Thus, for the problem at hand, we follow along, fix an integer $m_0(> \max\{1, a\})$, and gather pilot data $X_1, \ldots, X_m$ of size m defined as follows:

$$m \equiv m(\omega) = \max\left\{m_0, \left\lfloor d_\omega \sigma_L^{k/3} \right\rfloor + 1\right\} \text{ with } d_\omega = \left(a^2(c\omega)^{-1}\right)^{1/3}, \tag{4.1}$$

where all of the constants remain as defined in Section 3.

Based on pilot data, we evaluate the statistic $\hat{\sigma}_m = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - X_{m:1})$, and then determine N as follows:

$$N = \max\left\{m, \left\lfloor d_\omega \hat{\sigma}_m^{k/3} \right\rfloor + 1\right\} \tag{4.2}$$

as an estimator of n^* defined in (2.6). Recall that $\lfloor u \rfloor$ denotes the largest integer less than $u(> 0)$ as before.

If $N = m$, we would not require any additional observations at the second stage. However, if $N > m$, we sample the difference $N - m$ at the second stage by recording an additional set of observations $X_{m+1}, \ldots, X_N$. From full data $X_1, \ldots, X_N$ obtained by combining both stages, we propose to estimate μ by the smallest order statistic:

$$X_{N:1} = \min\{X_1, \ldots, X_N\}.$$

4.1. First-order asymptotics

We begin with some first-order results to contrast with those stated in Section 3. One will surely note that there is no stated sufficient condition involving m in Theorem 4.1. This is so because in the present setup, we have $m \equiv m(\omega) \to \infty$ as $\omega \to 0$. A proof will be sketched in Section 6.3.

Theorem 4.1. *With m and N respectively defined in (4.1) and (4.2), for each fixed value of μ, σ, c, k, a we have as $\omega \to 0$:*
(i) $N/n^* \xrightarrow{P} 1$; $n^*/N \xrightarrow{P} 1$;
(ii) $E\left[(N/n^*)^t\right] \to 1, t = -1, 1$ [*asymptotic first − order efficiency*];
where n^* comes from (2.6).

We note that all expressions shown in Theorem 4.1 converge to 1 which is in direct contrast with those from Theorem 3.1. That is, the modified two-stage procedure (4.1)–(4.2) has more attractive first-order properties than those under customary Stein-type two-stage procedure (3.1). One may claim convergence of higher positive and negative moments of N/n^* in part (ii) by referring to Mukhopadhyay and Duggan (1997, 1999) and Mukhopadhyay (1999), but we leave them out for brevity.

4.2. Second-order asymptotics

Again we avoid giving any sufficient condition involving m in Theorems 4.2–4.4, stated in this section, because $m \equiv m(\omega) \to \infty$ as $\omega \to 0$. We now supplement with second-order

properties that become readily accessible along the lines of Mukhopadhyay and Duggan (1997, 1999).

The following expressions provide second-order expansions as bounds of $E\left[(N/n^*)^t\right]$:

$$\begin{aligned} 1+\left\{t\psi+\frac{1}{2}t(t-1)\sigma_0^2\right\}\frac{1}{n^*}+o\left(\frac{1}{n^*}\right) &\le E\left(\frac{N}{n^*}\right)^t \\ &\le 1+\left\{t\psi+t+\frac{1}{2}t(t-1)\sigma_0^2\right\}\frac{1}{n^*}+o\left(\frac{1}{n^*}\right), \text{ if } t>0; \text{ and} \\ 1+\left\{t\psi+t+\frac{1}{2}t(t-1)\sigma_0^2\right\}\frac{1}{n^*}+o\left(\frac{1}{n^*}\right) &\le E\left(\frac{N}{n^*}\right)^t \\ &\le 1+\left\{t\psi+\frac{1}{2}t(t-1)\sigma_0^2\right\}\frac{1}{n^*}+o\left(\frac{1}{n^*}\right), \text{ if } t<0, \end{aligned} \tag{4.3}$$

where

$$\psi=\left(\frac{c}{a^2}\right)^{1/3}\left\{\frac{k}{6}\left(\frac{k}{3}-1\right)\left(\frac{a^2}{c}\right)^{1/3}\right\}\left(\frac{\sigma}{\sigma_L}\right)^{k/3}, \text{ and } \sigma_0^2=\frac{k^2}{9}\left(\frac{\sigma}{\sigma_L}\right)^{k/3}. \tag{4.4}$$

A proof of (4.3) follows directly from a more generally stated Theorem 2.1 in Mukhopadhyay and Duggan (1999). Indeed, Theorem 4.1, part (ii) holds for all non-zero t that directly follows from (4.3). For completeness, however, we first show a bound for $E(N-n^*)$.

Theorem 4.2. *With m and N respectively defined in (4.1) and (4.2), for each fixed value of μ, σ, c, k, a we have as $\omega \to 0$:*

$$\psi+O(\omega^{1/2}) \le E(N-n^*) \le \psi+1+O(\omega^{1/2}) [\textit{asymptotic second-order efficiency}], \tag{4.5}$$

where ψ is defined in (4.4) and n^* comes from (2.6).

A proof can be constructed using (6.2) and the rest is omitted for brevity. Theorem 4.2 shows the second-order efficiency property of the modified two-stage procedure (4.1)–(4.2) in the sense of Ghosh and Mukhopadhyay (1981). It is possible, however, to show a sharper result, namely, $E(N-n^*)=\psi+O(\omega^{1/2})$, but we leave it out for brevity.

Next, we provide a result that obtains the asymptotic distribution of a standardized version of N along the lines of Ghosh and Mukhopadhyay (1975) and Mukhopadhyay and Duggan (1997, 1999). Its proof follows from Lemma 2.1, part (i) in Mukhopadhyay and Duggan (1999) and hence it is omitted.

Theorem 4.3. *With m and N respectively defined in (4.1) and (4.2), for each fixed value of μ, σ, c, k, a we have as $\omega \to 0$:*

$$U \equiv n^{*-1/2}(N-n^*) \xrightarrow{\mathcal{L}} N(0,\sigma_0^2), \tag{4.6}$$

where σ_0^2 is defined in (4.4) and n^* comes from (2.6).

Again, because $X_{n:1}$ and $I(N=n)$ are independent for all fixed $n \ge m$, the associated expression for $E\left[\text{RPUC}_N\right]$ will resemble (2.7) with Q replaced by N from (4.2). The following theorem evaluates the risk per unit cost up to second-order approximation. We outline its proof in Section 6.4.

Theorem 4.4. *With m and N respectively defined in (4.1) and (4.2), for each fixed value of μ, σ, c, k, a we have as $\omega \to 0$:*

$$1 + \frac{1}{n^*}(6\sigma_0^2 + a - 3\psi - 3) + o\left(\frac{1}{n^*}\right) \le \omega^{-1} E\left[RPUC_N\right] \le 1 + \frac{1}{n^*}(6\sigma_0^2 + a - 3\psi) + o\left(\frac{1}{n^*}\right) \quad (4.7)$$

[*asymptotic second-order risk efficiency*],

where ψ is as defined in (4.4) and n^* comes from (2.6).

5. A purely sequential procedure

Unlike a modified two-stage procedure where we record observations in two batches, a purely sequential procedure is more involved operationally, but it also provides more accurate inferences. We only take as many observations step by step as required depending on the rule of termination.

In this section, we develop a purely sequential methodology and make use of nonlinear renewal theory to provide second-order approximations for the average sample size and RPUC, the risk per unit cost.

We recall the expressions of n^* from (2.6). We again fix an integer $m(> \max\{1, a\})$ and obtain pilot data $X_1, \ldots, X_m$ of size m. We then proceed by recording one additional observation X at every stage as needed according to the following rule:

$$N = \inf\left\{n \ge m : n \ge d_\omega \hat{\sigma}_n^{k/3}\right\} \text{ with } d_\omega = \left(a^2 (c\omega)^{-1}\right)^{1/3}. \quad (5.1)$$

As before, this stopping variable N again estimates n^* from (2.6). From full data $X_1, \ldots, X_N$ so obtained upon termination, we propose to estimate μ by the smallest order statistic:

$$X_{N:1} = \min\{X_1, \ldots, X_N\}.$$

Again, because $X_{n:1}$ and $I(N = n)$ are independent for all fixed $n \ge m$, the associated expression for $E\left[\mathrm{RPUC}_N\right]$ will resemble (2.7) with Q replaced by N from (5.1). In what follows, if the condition $m(> \max\{1, a\})$ will suffice for a result to hold, so we will not mention it. We will, however, mention a more stringent condition on m if that is what is required for a particular result to hold.

5.1. First-order asymptotics

We begin with some useful first-order asymptotic properties summarized in the next theorem. A proof of this theorem will be outlined in Section 6.5.

Theorem 5.1. *For N defined in (5.1), for each fixed value of μ, σ, a, k, c we have as $\omega \to 0$:*

(i) $N/n^* \xrightarrow{P} 1$;
(ii) $E\left[(N/n^*)^t\right] \to 1$ *for* $t > 0$ *if* $m(> \max\{2, a\})$[*asymptotic first-order efficiency*];
(iii) $E\left[(n^*/N)^t\right] \to 1$ *for* $t > 0$ *if* $m > \max\{1 + \frac{1}{3}kt, a\}$;
(iv) $\omega^{-1} E\left[RPUC_N\right] \to 1$ *if* $m > \max\{1 + \frac{4}{3}k, a\}$[*asymptotic first-order risk efficiency*];

(v) $V \equiv n^{*-1/2}(N-n^*) \xrightarrow{\mathcal{L}} N(0,\sigma_1^2)$ *with* $\sigma_1^2 = \frac{k^2}{9}$,
where n^* *comes from (2.6).*

We note that the expressions shown in Theorem 5.1, parts (i)–(iii) converge to 1, which is again in direct contrast with those from Theorem 3.1. That is, parallel to the results from Theorem 4.1 under the modified two-stage procedure (4.1)–(4.2), the purely sequential procedure (5.1) also has attractive asymptotic first-order properties compared to those under customary Stein-type two-stage procedure (3.1).

However, one major difference between the properties of the modified two-stage procedure (4.1)–(4.2) and the purely sequential procedure (5.1) already stands out. We recall that Theorem 4.3 concluded that

$$U \equiv n^{*-1/2}(N-n^*) \xrightarrow{\mathcal{L}} N(0,\sigma_0^2) \text{ with } \sigma_0^2 = \frac{k^2}{9}\left(\frac{\sigma}{\sigma_L}\right)^{k/3}.$$

In contrast, Theorem 5.1, part (v) shows that

$$V \equiv n^{*-1/2}(N-n^*) \xrightarrow{\mathcal{L}} N(0,\sigma_1^2) \text{ with } \sigma_1^2 = \frac{k^2}{9}.$$

In other words, asymptotic normal distribution for the standardized stopping variable V certainly has a smaller or tighter variance, namely, σ_1^2 under the purely sequential strategy, which gives it an edge over the modified two-stage strategy.

5.2. Second-order asymptotics

Observe that N from (5.1) can be rewritten as $J+1$ w.p.1 where

$$\begin{aligned} J &= \inf\left\{n \geq m-1 : \Sigma_{i=1}^n Z_i \leq n^{*-3/k} n^{\frac{3}{k}+1}\left(1+\frac{1}{n}\right)^{\frac{3}{k}}\right\} \\ &= \inf\left\{n \geq m-1 : \Sigma_{i=1}^n Z_i \leq h^* n^\delta L(n)\right\}, \end{aligned} \tag{5.2}$$

where the Z_is are i.i.d. exp(1) random variables. This has been established inside the proof of Theorem 5.1, part (iii) that is laid out in Section 6.5.

Next, we refer to nonlinear renewal theory, originally developed by Woodroofe (1977) and Lai and Siegmund (1977, 1979). We match (5.2) with the representation laid out in Mukhopadhyay (1988) and (Mukhopadhyay and Solanky, 1994, Section 2.4.2) as follows:

$$\delta = \frac{3}{k}+1,\ h^* = n^{*-3/k},\ L(n) = 1+\frac{3}{kn}+o\left(\frac{1}{n}\right),\ \text{so that } L_0 = \frac{3}{k}, \tag{5.3}$$

and also note:

$$\theta = 1, \tau^2 = 1, \beta^* = \frac{k}{3}, n_0^* = n^* \quad \text{and} \quad p = \frac{k^2}{9}. \tag{5.4}$$

Condition (2.5) from Mukhopadhyay (1988) or, equivalently (2.4.8) from Mukhopadhyay and Solanky (1994), is satisfied with $B=1$ and $b=1$. We define two special entities:

$$\begin{aligned} \nu \equiv \nu_k &= \frac{k}{6}\left(\frac{9}{k^2}+1\right) - \frac{1}{2}\Sigma_{n=1}^\infty n^{-1} E\left\{\max\left(0, \chi_{2n}^2 - 2n\left(\frac{3}{k}+1\right)\right)\right\}, \\ \eta \equiv \eta_k &= \frac{1}{3}k\nu - \left(\frac{1}{18}k^2 + \frac{1}{6}k + 1\right). \end{aligned} \tag{5.5}$$

Table 1. Selected values of $\nu \equiv \nu_k$ and $\eta \equiv \eta_k$ from (5.5), $k = 1(1)6$.

	k					
	1	2	3	4	5	6
ν_k	1.646	0.963	0.745	0.638	0.577	0.542
η_k	−0.673	−0.913	−1.254	−1.703	−2.259	−2.915

along the lines of (2.4.9)–(2.4.10) in Mukhopadhyay and Solanky (1994). Table 1 shows a few values of $\nu \equiv \nu_k$ and $\eta \equiv \eta_k$ for a couple of combinations of input for k.

To conclude this section, we now specify asymptotic second-order expansions of both positive and negative moments of $\frac{N}{n^*}$ (Theorem 5.2) as well as an asymptotic second-order expansion of the risk per unit cost (Theorem 5.3) under the purely sequential setting (5.1).

Theorem 5.2. *For N defined in (5.1), for each fixed value of μ, σ, a, k, c and every non-zero real number t, we have as $\omega \to 0$:*

$$E\left[\left(N/n^*\right)^t\right] = 1 + \left\{t\eta_k + t + \frac{1}{2}t(t-1)p\right\} n^{*-1} + o\left(n^{*-1}\right)$$
$$[\textit{asymptotic second-order efficiency}], \tag{5.6}$$

when (i) $m > \frac{(3-t)k}{3} + 1$ for $t \in (-\infty, 2) - \{-1, 1\}$; (ii) $m > \frac{k}{3} + 1$ for $t = 1$ and $t \geq 2$; and (iii) $m > \frac{2k}{3} + 1$ for $t = -1$, where n^*, p, and η_k come from (2.6), (5.4), and (5.5), respectively.

Theorem 5.3. *For N defined in (5.1), for each fixed value of μ, σ, a, k, c, we have the following second-order expansion of the risk per unit cost as $\omega \to 0$:*

$$\omega^{-1}E\left[RPUC_N\right] = 1 + (6p + a - 3\eta_k - 3)n^{*-1} + o\left(n^{*-1}\right)$$
$$[\textit{asymptotic second-order risk efficiency}], \tag{5.7}$$

when $m > \frac{7k}{3} + 1$ with n^*, p, and η_k coming from (2.6), (5.4), and (5.5), respectively.

Very brief outlines of proofs of Theorems 5.2–5.3 are sketched in Section 6.6. We have simply remarked how to connect what we want to prove here with established details from Mukhopadhyay (1988) and Mukhopadhyay and Solanky (1994).

6. Selected technicalities

In this section, we sketch some of the proofs of our main results from Sections 2–5. Often some intermediate steps are left out for brevity.

6.1. Proof of Theorem 2.1

From (2.1), the associated loss function is given by

$$L_Q = \exp\left(\frac{a(X_{Q:1} - \mu)}{\sigma}\right) - \frac{a(X_{Q:1} - \mu)}{\sigma} - 1.$$

Now, we recall risk per unit cost from (2.5) and proceed to evaluate $E\left[\mathrm{RPUC}_Q\right]$ as follows:

$$E\left[\frac{L_Q}{\mathrm{Cost}_Q}\right] = \sum_{m\le q<\infty} E\left[\frac{L_Q}{\mathrm{Cost}_Q} \mid Q=q\right] P(Q=q)$$
$$= \sum_{m\le q<\infty} E\left[\frac{L_q}{\mathrm{Cost}_q} \mid Q=q\right] P(Q=q).$$

But, under the conditions (i) and (ii) stated in Theorem 2.1, we may rewrite:

$$E\left[\mathrm{RPUC}_Q\right]$$
$$= \sum_{m\le q<\infty} E\left[\frac{L_q}{\mathrm{Cost}_q}\right] P(Q=q)$$
$$= \sum_{m\le q<\infty} \frac{E[L_q]}{\mathrm{Cost}_q} P(Q=q)$$
$$= \sum_{m\le q<\infty} \frac{\mathrm{Risk}_q}{\mathrm{Cost}_q} P(Q=q)$$
$$= \sum_{m\le q<\infty} \left\{\left(1-\frac{a}{q}\right)^{-1} - \frac{a}{q} - 1\right\} (cq\sigma^{-k})^{-1},$$

utilizing previous expressions of Risk_q and Cost_q from (2.2) and (2.4), respectively, while substituting q for n.

From (2.6), we see that $\frac{\sigma^k}{c} = \frac{n^{*3}\omega}{a^2}$ and thus we obtain

$$E\left[\mathrm{RPUC}_Q\right] = \omega \frac{n^{*3}}{a^2} E\left[\frac{1}{Q}\left(1-\frac{a}{Q}\right)^{-1} - \frac{a}{Q^2} - \frac{1}{Q}\right].$$

This is (2.7).

6.2. Proof of Theorem 3.1

From (3.1) we have the following basic inequality:

$$d_\omega \hat{\sigma}_m^{k/3} \le N \le m + d_\omega \hat{\sigma}_m^{k/3} \text{ w.p.1},$$

which implies:

$$n^{*-1} d_\omega \hat{\sigma}_m^{k/3} \le n^{*-1} N \le n^{*-1} m + n^{*-1} d_\omega \hat{\sigma}_m^{k/3} \text{ w.p.1.} \tag{6.1}$$

Now, as $\omega \to 0$, we have $n^* \to \infty$, $\frac{m}{n^*} \to 0$, so that $\frac{N}{n^*} \to \left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}$ w.p.1. The rest of part (i) follows immediately.

Part (ii) follows by taking expectations throughout (6.1) and then taking limits as $\omega \to 0$. Observe that $\frac{2(m-1)}{\sigma}\hat{\sigma}_m \sim \chi^2_{2(m-1)}$, which implies:

$$E\left[\hat{\sigma}_m^{k/3}\right] = \left(\frac{\sigma}{m-1}\right)^{k/3} \Gamma(m-1+k/3)\{\Gamma(m-1)\}^{-1}.$$

The proof that $E\left[\left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}\right] > 1$ follows from Mukhopadhyay and Hilton (1986).

6.3. Proof of Theorem 4.1

Part (i): Utilizing (6.1), we have the following basic inequality:

$$d_\omega \hat{\sigma}_m^{k/3} \leq N \leq mI(N=m) + d_\omega \hat{\sigma}_m^{k/3} + 1 \text{ w.p.1.} \tag{6.2}$$

Along the lines of (Mukhopadhyay and Duggan, 1999, Lemma 2.1), we know that $P(N = m) = O(\kappa^{m-1})$ where $\kappa = \frac{\sigma_L}{\sigma} \exp\left\{1 - \frac{\sigma_L}{\sigma}\right\}$ so that $\kappa \in (0,1)$. Thus, we can claim that $I(N=m) \xrightarrow{P} 0$ as $\omega \to 0$. Now, after dividing (6.2) throughout by n^*, we get

$$\left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3} \leq \frac{N}{n^*} \leq \frac{m}{n^*} I(N=m) + \left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3} + \frac{1}{n^*} \text{ w.p.1.} \tag{6.3}$$

But, $\hat{\sigma}_m \xrightarrow{P} \sigma$, $\frac{m}{n^*} = O(1)$, and $n^* \longrightarrow \infty$ as $\omega \to 0$. Part (i) is immediate from (6.3).
Part (ii): Taking expectations throughout (6.3), we have

$$E\left[\left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}\right] \leq E\left(\frac{N}{n^*}\right) \leq \frac{m}{n^*} P(N=m) + E\left[\left(\frac{\hat{\sigma}_m}{\sigma}\right)^{k/3}\right] + \frac{1}{n^*} \tag{6.4}$$

Clearly,

$$E\left[\hat{\sigma}_m^{k/3}\right] = \sigma^{k/3}\left\{1 + \frac{k}{3}\left(\frac{k}{3} - 1\right)(2m)^{-1} + O(m^{-2})\right\}.$$

Thus, a proof in the case of $E\left(\frac{N}{n^*}\right)$ is complete in view of (6.4). Next, in order to handle $E\left(\frac{n^*}{N}\right)$, we first note that $\liminf_{\omega\to 0} E\left(\frac{n^*}{N}\right) \geq 1$ by Fatou's lemma. Also, we have

$$\limsup_{\omega\to 0} E\left(\frac{n^*}{N}\right) \leq \limsup_{\omega\to 0} E\left(\frac{\sigma}{\hat{\sigma}_m}\right)^{k/3} = 1.$$

Now, the proof is complete.

6.4. Proof of Theorem 4.4

We may recall (2.7) and then rewrite

$$\begin{aligned}
&\omega^{-1} E\left[\text{RPUC}_N\right] \\
&= \frac{n^{*^3}}{a^2}\left\{E\left[\frac{1}{N}\left(1 - \frac{a}{N}\right)^{-1}\right] - E\left(\frac{a}{N^2}\right) - E\left(\frac{1}{N}\right)\right\} \\
&= E\left[\left(\frac{N}{n^*}\right)^{-3}\right] + an^{*-1} E\left[\left(\frac{N}{n^*}\right)^{-4}\right] + a^2 n^{*-2} E\left[U_N\right],
\end{aligned} \tag{6.5}$$

where U_N is a remainder term which is $O_P\left(\left(\frac{N}{n^*}\right)^{-5}\right)$.

One can show that $E[U_N] = O(1)$. Next, we may combine (6.5) and (4.3)–(4.4) with $t = -3, -4$ to express:

$$\omega^{-1} E\left[\text{RPUC}_N\right] \geq 1 + \frac{1}{n^*}(6\sigma_0^2 + a - 3\psi - 3) + o\left(\frac{1}{n^*}\right),$$

and

$$\omega^{-1}E\left[\mathrm{RPUC}_N\right] \le 1 + \frac{1}{n^*}(6\sigma_0^2 + a - 3\psi) + o\left(\frac{1}{n^*}\right).$$

This completes the proof.

6.5. Proof of Theorem 5.1

Part (i): Follows from using Lemma 1 of Chow and Robbins (1965).
Part (ii): With $m \ge 3$ we can claim that $\frac{N-1}{N-2} \le 2$ w.p.1, $U^* = \sup_{n\ge 2}\left\{\frac{1}{n}\Sigma_{i=1}^{n}(X_i - \mu)\right\}$, for sufficiently small $\omega(>0)$ so that $n^* > m$, observe the following inequality (w.p.1):

$$N \le m + d_\omega \hat{\sigma}_{N-1}^{k/3} \le m + 2^{k/3} d_\omega \left\{\frac{1}{N-2}\Sigma_{i=1}^{N-1}(X_i - \mu)\right\}^{k/3} \le m + 2^{k/3} d_\omega U^*,$$

which implies

$$\frac{N}{n^*} \le 1 + 2^{k/3}\sigma^{-k/3}U^*. \tag{6.6}$$

Now, by Wiener's (1939) ergodic theorem, it follows that $E[U^{*t}]$ is finite for all fixed positive number t. The right-hand side of (6.6) is also free from ω so that we can claim uniform integrability of all positive powers of $\frac{N}{n^*}$. Then, appealing to the dominated convergence theorem and part (i) complete a proof of part (ii).
Part (iii): Let us denote

$$S_n^* = \Sigma_{i=2}^{n}(n - i + 1)(X_{n:i} - X_{n:i-1}),$$

so that we have $\hat{\sigma}_n = S_n^*/(n-1)$. Let $Y_1, Y_2, \ldots$ be i.i.d. exp(σ) random variables and let $S_n = \Sigma_{i=1}^{n-1} Y_i$. Now, using the embedding ideas from Lombard and Swanepoel (1978) and Swanepoel and van Wyk (1982), we can claim that the distribution of $\{S_n^* : n \le n_0\}$ is identical to that of $\{S_n : n \le n_0\}$ for all n_0. Thus, N given by 5.1 is equivalently expressed as

$$\begin{aligned} N &\equiv \inf\left\{n \ge m : n \ge d_\omega\left(\frac{1}{n-1}\Sigma_{i=1}^{n-1}Y_i\right)^{k/3}\right\} \\ &= \inf\left\{n \ge m : \left(\frac{n}{n^*}\right)^{3/k}(n-1) \ge \Sigma_{i=1}^{n-1}Z_i\right\}, \end{aligned} \tag{6.7}$$

where Z_is are i.i.d. exp(1); that is, the standard exponential random variables.

Then, N from (6.7) can be written as $J+1$ w.p.1 with J defined in (5.2). Using Lemma 2.3 from Woodroofe (1977) or Theorem 2.4.8, part (i) of Mukhopadhyay and Solanky (1994) with $b = 1$, we can claim:

$$P(J \le \frac{1}{2}n^*) = O(n^{*-\frac{3}{k}(m-1)}). \tag{6.8}$$

Next, with fixed $t > 0$, because $N = J+1$ w.p.1, we have $0 < \left(\frac{n^*}{N}\right)^t \le \left(\frac{n^*}{J}\right)^t$ so that $\left(\frac{n^*}{N}\right)^t$ will be uniformly integrable if we show that

$$\left(\frac{n^*}{J}\right)^t \text{ is uniformly integrable.} \tag{6.9}$$

Now, we may write

$$\left(\frac{n^*}{J}\right)^t I(J > \frac{1}{2}n^*) < 2^t,$$

so that $\left(\frac{n^*}{J}\right)^t I(J > \frac{1}{2}n^*)$ must be uniformly integrable. But, $\left(\frac{n^*}{J}\right)^t I(J > \frac{1}{2}n^*) \xrightarrow{P} 1$ and, hence, we must have

$$E\left[\left(\frac{n^*}{J}\right)^t I(J > \frac{1}{2}n^*)\right] = 1 + o(1). \tag{6.10}$$

Additionally, in view of (6.8), we also note the following:

$$E\left[\left(\frac{n^*}{J}\right)^t I(J \leq \frac{1}{2}n^*)\right] \leq \left(\frac{n^*}{m-1}\right)^t P(J \leq \frac{1}{2}n^*) = O(n^{*-\frac{3}{k}(m-1)+t}), \tag{6.11}$$

which is $o(1)$ if $m > 1 + \frac{1}{3}kt$.

Combining (6.10)–(6.11), we note that (6.9) will follow so that we can claim

$$E\left[\left(\frac{n^*}{J}\right)^t\right] = 1 + o(1) \text{ if } m > 1 + \frac{1}{3}kt, \quad \text{with } t > 0, \tag{6.12}$$

which is part (iii).

Part (iv): We may go back to (6.5) and express

$$\omega^{-1}E\left[\text{RPUC}_N\right] = E\left[\left(\frac{N}{n^*}\right)^{-3}\right] + an^{*-1}E\left[V_N\right], \tag{6.13}$$

where V_N is a remainder term, which is $O_P\left(\left(\frac{N}{n^*}\right)^{-4}\right)$.

Clearly, in view of (6.12), $E\left[\left(\frac{n^*}{J}\right)^3\right] = 1 + o(1)$ if $m > 1 + k$ and one can show that $E\left[V_N\right] = O(1)$ if $m > 1 + \frac{4}{3}k$. Then, part (iv) follows from (6.13).

Part (v): This result follows directly from an application of Ghosh and Mukhopadhyay's (1975) theorem. One may also refer to Theorem 2.4.3 or Theorem 2.4.8, part (ii) in Mukhopadhyay and Solanky (1994).

6.6. Outlines of proofs of Theorems 5.2–5.3

Theorem 5.2 follows along the lines of Theorem 2.4.8, part (iv) and from its established applications found in Mukhopadhyay and Solanky (1994).

For a proof of Theorem 5.3, we recall that the associated expression for $E\left[\text{RPUC}_N\right]$ will resemble (2.7) with Q replaced by N from (5.1). Then, one will proceed with an expansion of $\omega^{-1}E\left[\text{RPUC}_N\right]$ similar to that in (6.5) and exploit Theorem 5.1, part (iii) with $t = 5$ and Theorem 5.2 with $t = -3, -4$. Additional details are left out for brevity.

7. Data analysis: Simulations

Thus far we have developed theoretical properties for three proposed estimation strategies in Sections 3–5. Section 6 gave outlines of some selected derivations. Now, it is time to implement the methodologies and investigate how those estimation strategies may perform when sample sizes are small (20, 30) to moderate (50, 100, 150) to large (300, 400, 500). Computer simulations help in this investigation. All simulations are carried out with (R Core Team, 2014) codes based on 10, 000 (= H, say) replications run under each configuration and each methodology.

Under each procedure, we generated pseudorandom observations from the distribution (1.1) with $\mu = 5, \sigma = 10$. Then, we fixed certain values of a, c, k and appropriate rounded values for n^*, thereby solving for a corresponding value of the risk bound, ω. In other words, a set of preassigned values for a, c, k, ω will have the associated n^* values as shown in our tables (column 1).

We fix a pilot sample size, namely, m in the contexts of Stein-type two-stage methodology (3.1) and purely sequential methodology (5.1). In the context of modified two-stage methodology (4.1)–(4.2), we fix a positive lower bound σ_L for σ and a number m_0, thereby determining m from (4.1). While implementing a methodology to determine the final sample size (N) and a terminal estimator ($X_{N:1}$) of μ, we pretended that we did not know μ, σ, and n^* values.

Now, let us set the notations that we will use in the tables to follow. We focus on implementing a particular estimation methodology under a fixed configuration of all necessary input (e.g., $a, c, k, \omega, m, m_0, \sigma_L$ as appropriate). We ran the ith replication by beginning with m pilot observations and then eventually ending sampling by recording the final sample size $N = n_i$, terminal estimator $X_{N:1} = x_{n_i:1}$, and the achieved risk per unit cost:

$$\text{RPUC}_{n_i} = \omega \frac{n^{*3}}{a^2} \left\{ \frac{1}{n_i} \left(1 - \frac{a}{n_i} \right)^{-1} - \frac{a}{n_i^2} - \frac{1}{n_i} \right\} = r_i, \text{ say,} \tag{7.1}$$

$i = 1, \ldots, H$. The basic notations are itemized next where H is kept fixed at 10,000:

$\overline{n} = H^{-1} \Sigma_{i=1}^{H} n_i$	Estimate of $E(N)$ or n^*;
$s_{\overline{n}} = \sqrt{(H^2 - H)^{-1} \Sigma_{i=1}^{H} (n_i - \overline{n})^2}$	Estimated standard error of $\overline{n}$;
$\overline{x}_{\min} = H^{-1} \Sigma_{i=1}^{H} x_{n_i:1}$	Estimate of μ;
$s_{\overline{x}_{\min}} = \sqrt{(H^2 - H)^{-1} \Sigma_{i=1}^{H} (x_{n_i:1} - \overline{x}_{\min})^2}$	Estimated standard error of $\overline{x}_{\min}$;
r_i	RPUC_{n_i} from (7.1)
$\overline{r} = H^{-1} \Sigma_{i=1}^{H} r_i$ with r_i from (7.1)	Risk estimator
$s_{\overline{r}} = \sqrt{(H^2 - H)^{-1} \Sigma_{i=1}^{H} (r_i - \overline{r})^2}$	Estimated standard error of $\overline{r}$;
$\overline{z} = \overline{r}/\omega$	Estimated risk efficiency to be; compared with 1;
$s_{\overline{z}} = s_{\overline{r}}/\omega$	Estimated standard error of $\overline{z}$;

Now, we are in a position to summarize observed performances of the proposed estimation methodologies laid down in Sections 3–5. We have many sets of tables and results obtained

from simulations under additional configurations. For brevity, however, we outline only a small subset of our findings.

7.1. Stein-type two-stage procedure (3.1)

First we present the performances of Stein-type two-stage estimation methodology (3.1) in Table 2 for

$$n^* = 30, 100, 300, 500 \text{and}$$
$$(k, m) = (1, 3), (2, 4), (3, 5), (4, 6), (5, 7).$$

Table 2 specifies μ, σ, a, c. Each block shows (k, m), n^* (column 1), ω (column 2), the estimated (from 10,000 simulations) values $\bar{x}_{\min}$, $s_{\bar{x}_{\min}}$ (column 3), values $\bar{n}$, $s_{\bar{n}}$ (column 4), the ratio $\bar{n}/n^*$ (column 5), and values $\bar{z}$, $s_{\bar{z}}$ (column 6).

All $\bar{x}_{\min}$ values appear rather close to $5(= \mu)$ with very small estimated standard error values $s_{\bar{x}_{\min}}$, for sample sizes over 300. For smaller (k, m) values, namely, $(1, 3)$, $(2, 4)$, it appears that $\bar{n}$ slightly underestimates n^*, but these performances reverse for larger choices of (k, m). This is consistent with Theorem 3.1, part (ii). The last column shows that the two-stage estimation methodology (3.1) is not successful in delivering a risk bound approximately under (or close to) our preset goal ω.

Table 2. Simulation results from 10,000 replications for Stein-type two-stage procedure (3.1) with $\mu = 5$, $\sigma = 10, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}(s_{\bar{x}_{\min}})$	$\bar{n}(s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}(s_{\bar{z}})$
		$\boldsymbol{k = 1, m = 3}$			
30	3.7×10^{-3}	5.3720(0.0040)	28.84(0.0704)	0.9615	1.9767(0.0454)
100	1×10^{-4}	5.1120(0.0012)	94.95(0.2319)	0.9495	1.9568(0.0341)
300	3.7×10^{-6}	5.0385(0.0004)	284.88(0.7064)	0.9496	1.9635(0.0320)
500	8×10^{-7}	5.0227(0.0002)	473.89(1.1778)	0.9477	1.9872(0.0367)
		$\boldsymbol{k = 2, m = 4}$			
30	3.7×10^{-2}	5.3978(0.0050)	29.51(0.1130)	0.9837	4.2397, (0.2016)
100	1×10^{-3}	5.1227(0.0015)	97.74(0.3790)	0.9774	3.9148(0.1488)
300	3.7×10^{-5}	5.0413(0.0004)	289.75(1.1285)	0.9659	4.5254(0.5000)
500	8×10^{-6}	5.0249(0.0003)	482.53(1.8796)	0.9650	3.9985(0.1588)
		$\boldsymbol{k = 3, m = 5}$			
30	3.7×10^{-1}	5.4293(0.0057)	30.38(0.1497)	1.0127	7.0348(0.2574)
100	1×10^{-2}	5.1341(0.0018)	100.55(0.5064)	1.0055	9.3451, (0.9627)
300	3.7×10^{-4}	5.0440(0.0006)	302.75(1.5099)	1.0091	9.7991(1.0965)
500	8×10^{-5}	5.0265(0.0003)	501.48(2.4959)	1.0029	15.2665(6.1989)
		$\boldsymbol{k = 4, m = 6}$			
30	3.7	5.4622(0.0064)	31.53(0.1874)	1.0510	8.8788(0.2425)
100	0.1	5.3720(0.0040)	105.02(0.6233)	1.0502	16.2954(1.5576)
300	3.7×10^{-3}	5.0480(0.0008)	311.15(1.8387)	1.0371	21.7228(4.6001)
500	8×10^{-4}	5.0282(0.0004)	521.66(3.1057)	1.0433	14.8328(1.4949)
		$\mathbf{k = 5, m = 7}$			
30	37.03	5.4663(0.0065)	32.64(0.2190)	1.0880	9.3771(0.2085)
100	1	5.1456(0.0023)	110.94(0.7588)	1.1094	21.6822(1.5331)
300	3.7×10^{-2}	5.0510(0.0008)	324.97(2.2472)	1.0832	41.4347(9.9591)
500	8×10^{-3}	5.0306(0.0005)	555.25(3.8969)	1.1105	30.0233(4.5692)

7.2. Modified two-stage procedure (4.1)–(4.2)

Now, we move to summarize performances for the modified two-stage estimation methodology (4.1)–(4.2) in Table 3 for

$$n^* = 20, 30, 50, 100, 150, 300, 400, 500 \text{ and}$$
$$(k, m_0) = (1, 3), (2, 4), (3, 5), (4, 6), (5, 7).$$

In this methodology, we need a positive and known lower bound $\sigma_L(= 3)$ for true σ but σ remains unknown. The pilot size m was determined from (4.1) but that m is not shown in Table 3. The estimation methodology (4.2) was implemented as described. Table 3 specifies $\mu, \sigma, \sigma_L, a, c$, and each block shows (k, m_0), n^* (column 1), ω (column 2), the estimated (from 10,000 simulations) values $\bar{x}_{\min}$, $s_{\bar{x}_{\min}}$ (column 3), values $\bar{n}$, $s_{\bar{n}}$ (column 4), the ratio $\bar{n}/n^*$ (column 5), and values $\bar{z}$, $s_{\bar{z}}$ (column 6).

All $\bar{x}_{\min}$ values appear closer to $5(= \mu)$ with very small estimated standard error values $s_{\bar{x}_{\min}}$ for sample size 150 or over. For all (k, m) values, it appears that $\bar{n}$ estimates n^* very accurately across the board. These features are consistent with Theorem 4.1, parts (i)–(ii). The last column shows that the modified two-stage estimation methodology (4.1)–(4.2) is very successful for $(k, m) = (1, 3)$ for all n^* values under consideration in delivering a risk bound of approximately our preset goal ω. However, in the case of $k = 2, 3, 4, 5$, a similar level of success in delivering a risk bound of approximately our preset goal ω is observed as n^* successively exceeded 50, 100, 150, 300, respectively.

The entries in Table 4 were obtained similar to those in Table 3 with one major difference. Table 4 used another positive and known lower bound $\sigma_L(= 5)$ for true σ but σ continued to remain unknown. Again, the pilot size m was determined from (4.1) but m is not shown in Table 4. Comparing columns 5 from Tables 3 and 4, it is clear $\bar{n}/n^*$ is nearer to 1 in Table 4 . Also, entries found in the last column of Table 4 look more favorable then those in Table 3. This feature should be expected because the specified positive and known lower bound $\sigma_L(= 5)$ is nearer to $\sigma(= 10)$ than $\sigma_L(= 3)$ is.

Table 5 corresponds to $\sigma_L = 11$, whereas true σ was 10. That is, Table 5 shows the performances of the modified two-stage estimation methodology (4.1)–(4.2) if σ_L is misspecified in that it just barely goes over true σ. It is clear that the modified two-stage estimation methodology holds up rather well under mild misspecification of σ_L.

In Tables 6–8 we provide the values of ψ found in (4.4) corresponding to the configurations highlighted in Tables 3–5 respectively. Importance of ψ comes from the fact that $E(N - n^*)$ values could be expected to lie inside the corresponding interval $[\psi, \psi + 1]$ for large n^* values in view of Theorem 4.2. Thus, one could expect $\bar{n} - n^*$ values to lie inside the corresponding interval $[\psi, \psi + 1]$ for large n^* values. In Tables 6–7, we find that nearly all $\bar{n} - n^*$ values lie very close to (or inside) the corresponding interval $[\psi, \psi + 1]$ along with the bold entries missing the boat by a wider margin. But, the entries in Table 8 correspond to the case of misspecifying σ_L considered by Table 5 and hence those entries in Table 8 are supposed to be completely out of line with regard to any sense of practicality of second-order approximation under misspecification of σ_L.

We now provide Figures 1–2 to validate empirically the normality result described in (4.6). We considered four scenarios, namely, $\sigma_L = 3$, $k = 1$, $n^* = 30$ (Figure 1(a)); $\sigma_L = 3$, $k = 4$, $n^* = 500$ (Figure 1(b)); $\sigma_L = 5$, $k = 2$, $n^* = 100$ (Figure 2(a)); and $\sigma_L = 5$, $k = 5$, $n^* = 500$ (Figure 2(b)). Under a specific configuration, we recorded observed values:

$$N = n_i, i = 1, \ldots, H(= 10,000)$$

Table 3. Simulation results from 10,000 replications for the modified two-stage procedure (4.1)–(4.2) with $\mu = 5, \sigma = 10, \sigma_L = 3, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}(s_{\bar{x}_{\min}})$	$\bar{n}(s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}, (s_{\bar{z}})$
		$\boldsymbol{k = 1, m_0 = 3}$			
20	1.25×10^{-2}	5.5036(0.0050)	20.34(0.0187)	1.0174	1.0553(0.0032)
30	3.7×10^{-3}	5.3308(0.0032)	30.32(0.0225)	1.0109	1.0364(0.0024)
50	8×10^{-4}	5.2008(0.0020)	50.36(0.0291)	1.0072	1.0193(0.0018)
100	1×10^{-4}	5.0992(0.0009)	100.33(0.0414)	1.0033	1.0103(0.0012)
150	2.96×10^{-5}	5.0657(0.0006)	150.35(0.0499)	1.0023	1.0062(0.0010)
300	3.7×10^{-6}	5.0333(0.0003)	300.31(0.0707)	1.0010	1.0035(0.0007)
400	1.56×10^{-6}	5.0252(0.0002)	400.47(0.0813)	1.0011	1.0014(0.0006)
500	8×10^{-7}	5.0199(0.0001)	500.20(0.0922)	1.0004	1.0028(0.0005)
		$\boldsymbol{k = 2, m_0 = 4}$			
20	1.25×10^{-1}	5.5316(0.0056)	20.24(0.0466)	1.0123	1.4659(0.0139)
30	3.7×10^{-2}	5.3410(0.0035)	30.35(0.0554)	1.0119	1.2423(0.0082)
50	8×10^{-3}	5.1995(0.0020)	50.31(0.0712)	1.0062	1.1359(0.0053)
100	1×10^{-3}	5.1003(0.0010)	100.22(0.0995)	1.0022	1.0661(0.0033)
150	2×10^{-4}	5.0673(0.0006)	150.25(0.1203)	1.0016	1.0416(0.0025)
300	3.7×10^{-5}	5.0330(0.0003)	300.26(0.1731)	1.0008	1.0210(0.0017)
400	1.56×10^{-5}	5.0248(0.0002)	400.17(0.1978)	1.0004	1.0160(0.0015)
500	8×10^{-6}	5.0198(0.0002)	500.69(0.2250)	1.0013	1.0100(0.0013)
		$\boldsymbol{k = 3, m_0 = 5}$			
20	1.25	5.5887(0.0072)	20.54(0.0891)	1.0274	1.5358(0.0690)
30	3.7×10^{-1}	5.3713(0.0041)	30.65(0.1069)	1.0217	1.3024(0.0387)
50	8×10^{-2}	5.2194(0.0023)	50.28(0.1325)	1.0057	1.2447(0.0178)
100	1×10^{-2}	5.1021(0.0010)	100.27(0.1862)	1.0027	1.1249(0.0081)
150	2×10^{-3}	5.0672(0.0006)	150.69(0.2258)	1.0046	1.1048(0.0055)
300	3.7×10^{-4}	5.0336(0.0003)	300.54(0.3150)	1.0018	1.0825(0.0035)
400	1.56×10^{-4}	5.0247(0.0002)	400.47(0.3628)	1.0011	1.0274(0.0029)
500	8×10^{-5}	5.0204(0.0002)	500.37(0.4111)	1.0007	1.0087(0.0018)
		$\boldsymbol{k = 4, m_0 = 6}$			
20	12.5	5.6329(0.0082)	21.42(0.1259)	1.0713	1.7548(0.1084)
30	3.7	5.5033(0.0050)	31.63(0.1684)	1.0546	1.6420(0.0908)
50	0.8	5.2266(0.0027)	51.85(0.2166)	1.0370	1.3259(0.0668)
100	0.1	5.1084(0.0010)	101.97(0.3037)	1.0197	1.1544(0.0208)
150	0.02	5.0693(0.0007)	151.53(0.3659)	1.0102	1.1124(0.0128)
300	3.7×10^{-3}	5.0336(0.0003)	302.45(0.5159)	1.0081	1.0987(0.0066)
400	1.56×10^{-3}	5.0247(0.0002)	400.40(0.5934)	1.0010	1.0487(0.0055)
500	8×10^{-4}	5.0200(0.0002)	500.16(0.6661)	1.0003	1.0215(0.0048)
		$\boldsymbol{k = 5, m_0 = 7}$			
20	125	5.6304(0.0082)	22.52(0.1467)	1.1263	2.1548(0.0837)
30	37.03	5.4653(0.0064)	33.22(0.2242)	1.1075	1.7487(0.0758)
50	8	5.2869(0.0044)	54.77(0.3838)	1.1154	1.5655(0.0644)
100	1	5.1201(0.0015)	104.29(0.4803)	1.0429	1.3241(0.0597)
150	0.2	5.0334(0.0008)	155.43(0.5801)	1.0362	1.2015(0.0514)
300	0.037	5.0358(0.0003)	305.21(0.8044)	1.0173	1.1198(0.0425)
400	0.015	5.0260(0.0002)	405.39(0.9324)	1.0134	1.0489(0.0398)
500	0.008	5.0204(0.0002)	507.11(1.0365)	1.0142	1.0341(0.0304)

and thus calculated 10,000 associated standardized $(n_i - n^*)/\sqrt{n^*}$ values. Under each specific configuration, such 10,000 observed $u_i \equiv (n_i - n^*)/\sqrt{n^*}$ values provide the empirical distribution of the standardized sample size (dashed curve in red). We superimpose on it the appropriate theoretical $N(0, \sigma_0^2)$ distributions (solid curve in blue) where $\sigma_0^2 = \frac{1}{9}(\frac{\sigma}{\sigma_L})^{k/3}k^2$ coming from (4.4).

Table 4. Simulation results from 10, 000 replications for the modified two-stage procedure (4.1)–(4.2) with $\mu = 5, \sigma = 10, \sigma_L = 5, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}(s_{\bar{x}_{\min}})$	$\bar{n}(s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}(s_{\bar{z}})$
		$\boldsymbol{k = 1, m_0 = 3}$			
30	3.7×10^{-3}	5.3290(0.0033)	30.34(0.0210)	1.0116	1.0292(0.0022)
100	1×10^{-4}	5.1008(0.0010)	100.36(0.0373)	1.0036	1.0073(0.0011)
300	3.7×10^{-6}	5.0336(0.0003)	300.45(0.0656)	1.0015	1.0016(0.0006)
500	8×10^{-7}	5.0198(0.0001)	500.35(0.0824)	1.0007	1.0015(0.0004)
		$\boldsymbol{k = 2, m_0 = 4}$			
30	3.7×10^{-2}	5.3312(0.0033)	30.32(0.0469)	1.0109	1.1678(0.0061)
100	1×10^{-3}	5.1010(0.0010)	100.33(0.0841)	1.0033	1.0445(0.0027)
300	3.7×10^{-5}	5.0340(0.0003)	300.46(0.1468)	1.0015	1.0132(0.0015)
500	8×10^{-6}	5.0199(0.0002)	500.10(0.1886)	1.0002	1.0099(0.0011)
		$\boldsymbol{k = 3, m_0 = 5}$			
30	3.7×10^{-1}	5.3546(0.0037)	30.51(0.0805)	1.0173	1.5370(0.0148)
100	1×10^{-2}	5.1009(0.0010)	100.54(0.1421)	1.0054	1.1252(0.0052)
300	3.7×10^{-4}	5.0340(0.0003)	300.37(0.2447)	1.0012	1.0405(0.0026)
500	8×10^{-5}	5.0195(0.0001)	500.86(0.3158)	1.0017	1.0210(0.0019)
		$\boldsymbol{k = 4, m_0 = 6}$			
30	3.7	5.3681(0.0041)	31.20(0.1222)	1.0401	1.5429(0.0343)
100	0.1	5.1019(0.0010)	101.35(0.2167)	1.0135	1.2921(0.0099)
300	3.7×10^{-3}	5.0338(0.0003)	301.47(0.3687)	1.0049	1.0837(0.0042)
500	8×10^{-4}	5.0201(0.0002)	501.40(0.4701)	1.0022	1.0496(0.0030)
		$\boldsymbol{k = 5, m_0 = 7}$			
30	37.03	5.4061(0.0050)	32.52(0.1771)	1.0841	1.7654(0.0710)
100	1	5.1074(0.0011)	102.38(0.3005)	1.0238	1.5306(0.0202)
300	3.7×10^{-2}	5.0341(0.0003)	302.29(0.5197)	1.0076	1.1724(0.0065)
500	8×10^{-3}	5.0200(0.0002)	501.74(0.6724)	1.0034	1.1062(0.0047)

The two curves as shown appear slightly off from each other in Figure 1(a) which is meant for small $n^*(= 30)$, however we recall that a visual asymptotic match should be expected for larger n^* (Theorem 4.3). Indeed, Figures 1(b) and 2(a)–2(b) show clearly that the empirical distribution curve and the theoretical distribution curve nearly lie on each other when $n^*(= 100, 500)$ is large.

On 10, 000 observed values of u, we also performed the customary Kolmogorov-Smirnov (K-S) test for normality in the case of each dataset that generated Figures 1(a)–1(b) and Figures 2(a)–2(b). We summarize the observed values of K-S test statistic (D) under the null hypothesis of normality with associated p-values.

Case	Parameter configuration	K-S stat D	p-value
1a	$\sigma_L = 3, k = 1, n^* = 30$	0.56149	0.8770
1b	$\sigma_L = 3, k = 4, n^* = 500$	0.51959	0.9608
2a	$\sigma_L = 5, k = 2, n^* = 100$	0.51191	0.9762
2b	$\sigma_L = 5, k = 5, n^* = 500$	0.51170	0.9766

The p-values shown under cases 1b, 2a, 2b are sizably larger than the p-value shown under case 1a and all four p-values are much larger than 0.05. Our earlier sentiments supported by visual examinations of Figures 1(a)–1(b) and Figures 2(a)–2(b) are clearly validated by the

Table 5. Simulation results from 10,000 replications for the modified two-stage procedure (4.1)–(4.2) with $\mu = 5, \sigma = 10, \sigma_L = 11, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}(s_{\bar{x}_{\min}})$	$\bar{n}(s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}(s_{\bar{z}})$
		$\boldsymbol{k = 1, m_0 = 3}$			
30	3.7×10^{-3}	5.3199(0.0031)	31.43(0.0085)	1.0478	0.9014(0.0015)
100	1×10^{-4}	5.0976(0.0009)	104.22(0.0076)	1.0422	0.8921(0.0011)
300	3.7×10^{-6}	5.0320(0.0031)	310.10(0.0067)	1.0336	0.9083(0.0005)
500	8×10^{-7}	5.0192(0.0001)	517.03(0.0040)	1.0340	0.9061(0.0004)
		$\boldsymbol{k = 2, m_0 = 4}$			
30	3.7×10^{-2}	5.3078(0.0030)	32.78(0.0162)	1.0926	0.8005(0.0009)
100	1×10^{-3}	5.0935(0.0009)	107.52(0.0169)	1.0752	0.8130(0.0005)
300	3.7×10^{-5}	5.0311(0.0003)	320.20(0.0127)	1.0673	0.8250(0.0002)
500	8×10^{-6}	5.0190(0.0001)	533.07(0.0083)	1.0664	0.8267(0.0001)
		$\boldsymbol{k = 3, m_0 = 5}$			
30	3.7×10^{-1}	5.2935(0.0029)	34.11(0.0249)	1.1372	0.7179(0.0011)
100	1×10^{-2}	5.0903(0.0009)	110.90(0.0278)	1.1090	0.7421(0.0008)
300	3.7×10^{-4}	5.0301(0.0002)	330.26(0.0180)	1.1008	0.7519(0.0004)
500	8×10^{-5}	5.0179(0.0001)	550.11(0.0133)	1.1002	0.7522(0.0001)
		$\boldsymbol{k = 4, m_0 = 6}$			
30	3.7	5.2792(0.0028)	36.21(0.0305)	1.2070	0.6026(0.0010)
100	0.1	5.0876(0.0008)	115.09(0.0367)	1.1509	0.6648(0.0007)
300	3.7×10^{-3}	5.0291(0.0002)	341.34(0.0240)	1.1378	0.6810(0.0005)
500	8×10^{-4}	5.0175(0.0001)	568.08(0.0120)	1.1361	0.6830(0.0001)
		$\boldsymbol{k = 5, m_0 = 7}$			
30	37.03	5.2649(0.0026)	37.70(0.0411)	1.2566	0.5426(0.0014)
100	1	5.0829(0.0008)	119.24(0.0430)	1.1924	0.5985(0.0008)
300	3.7×10^{-2}	5.0288(0.0002)	352.42(0.0300)	1.1747	0.6188(0.0004)
500	8×10^{-3}	5.0168(0.0001)	587.12(0.0166)	1.1742	0.6187(0.0002)

Table 6. Values of $\bar{n} - n^*$, ψ and $\psi + 1$ for each k used in Table 3.

$n^* \backslash k$	1	2	3	4	5
20	0.34	0.24	0.54	1.42	**2.52**
30	0.32	0.35	0.65	1.63	**3.22**
50	0.36	0.31	0.28	1.85	4.77
100	0.33	0.22	0.27	1.97	4.29
150	0.35	0.25	0.69	1.53	5.06
300	0.31	0.26	0.22	**2.45**	**5.43**
400	0.47	0.17	0.47	**0.40**	**5.21**
500	0.20	0.69	0.37	**0.16**	**5.39**
ψ	−0.1659	−0.2479	0	1.1065	4.1323
$\psi + 1$	0.8341	0.7521	1	2.1065	5.1323

Table 7. Values of $\bar{n} - n^*$, ψ and $\psi + 1$ for each k used in Table 4.

$n^* \backslash k$	1	2	3	4	5
30	0.34	0.32	0.51	1.20	2.52
100	0.36	0.33	0.54	1.35	2.38
300	0.45	0.46	0.37	1.47	2.29
500	0.35	0.10	0.86	1.10	**1.74**
ψ	−0.1399	−0.1763	0	0.5599	1.7637
$\psi + 1$	0.8601	0.8237	1	1.5599	2.7637

Table 8. Values of $\bar{n} - n^*$, ψ and $\psi + 1$ for each k used in Table 5.

$n^* \backslash k$	1	2	3	4	5
30	**1.43**	**2.78**	**4.11**	**6.21**	**7.70**
100	**4.22**	**7.52**	**10.90**	**15.09**	**19.24**
300	**10.10**	**20.20**	**30.26**	**41.34**	**52.42**
500	**17.03**	**33.07**	**50.11**	**68.08**	**87.12**
ψ	−0.1076	−0.1042	0	0.1957	0.4739
$\psi + 1$	0.8924	0.8958	1	1.1957	1.4739

K-S test of normality under each scenario. That is, we are reasonably assured of a good fit between the observed values of u and a normal curve with a high level of confidence for all practical purposes.

Case	Parameter configuration	K-S stat D	p-value
3a	$m = 5, k = 1, n^* = 30$	0.56679	0.8664
3b	$m = 12, k = 4, n^* = 500$	0.51066	0.9787
4a	$m = 8, k = 2, n^* = 100$	0.50022	0.9996
4b	$m = 15, k = 5, n^* = 500$	0.51380	0.9724

7.3. Purely sequential procedure (5.1)

In this section, we summarize performances for the purely sequential estimation methodology (5.1) in Table 9 for

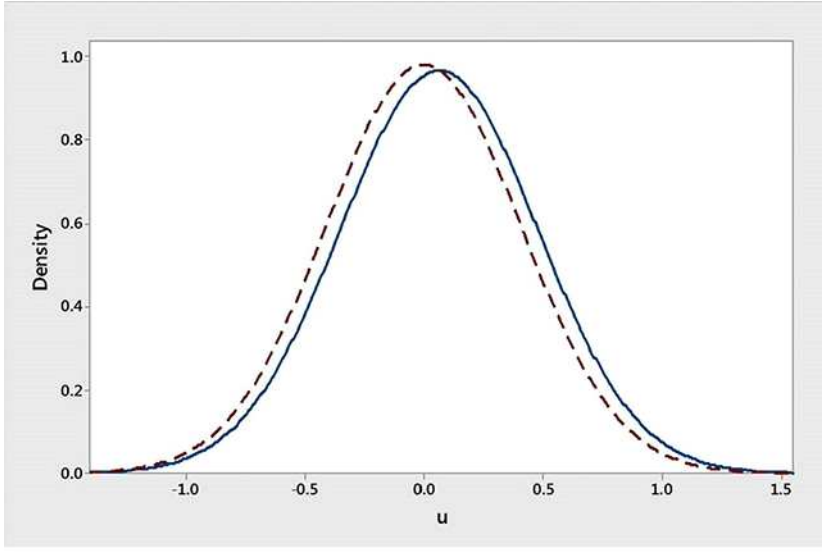
$$n^* = 20, 30, 50, 100, 150, 300, 400, 500 \text{ and}$$
$$(k, m) = (1, 5), (2, 8), (3, 9), (4, 12), (5, 15).$$

The estimation methodology (5.1) was implemented as described. Table 9 specifies μ, σ, a, c and each block shows (k, m), n^* (column 1), ω (column 2), the estimated (from 10,000 simulations) values $\overline{x}_{\min}$, $s_{\overline{x}_{\min}}$ (column 3), values $\overline{n}$, $s_{\overline{n}}$ (column 4), the ratio $\overline{n}/n^*$ (column 5), and values $\overline{z}$, $s_{\overline{z}}$ (column 6). An explanation of column 7 comes later.

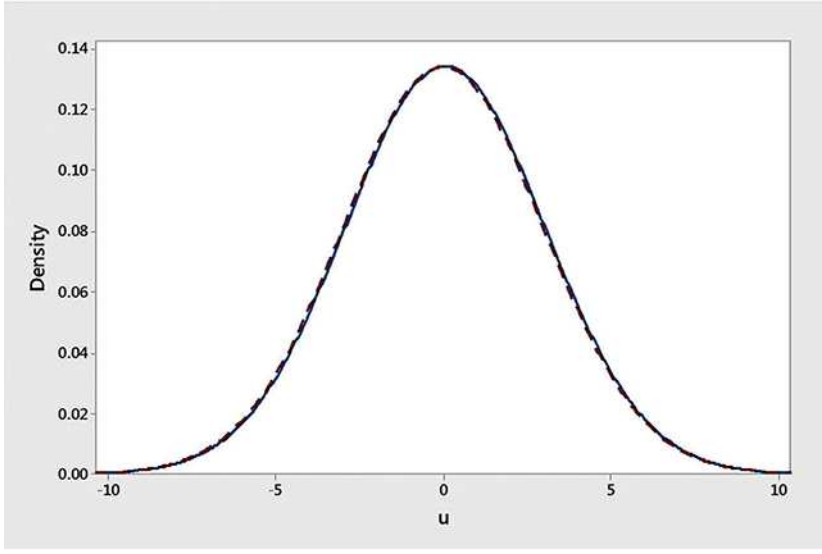
All $\overline{x}_{\min}$ values appear closer to $5(= \mu)$ with very small estimated standard error values $s_{\overline{x}_{\min}}$ for sample size 100 or over. For all (k, m) values, it appears that $\overline{n}$ estimates n^* very accurately across the board with slight occasional over- or underestimation. These features are consistent with Theorem 5.1, parts (i)–(ii). Column 6 shows that the purely sequential estimation methodology (5.1) is overall very successful for $(k, m) = (1, 5)$ for all n^* values under consideration in delivering a risk bound of approximately our preset goal ω. However, in the case of $k = 2, 3, 4, 5$, a similar level of success in delivering a risk bound of approximately our preset goal ω is observed as n^* successively exceeded 50,100, 150,300, respectively. It is our feeling that the purely sequential procedure provides estimates for the risk per unit cost, which are generally closer to 1 than the two-stage methodologies (3.1) or (4.1)–(4.2).

Theorem 5.3 showed that $\omega^{-1} E\,[\text{RPUC}_N]$ should be close to $1 + \varepsilon$ where we have

$$\varepsilon = (6p + a - 3\eta_k - 3)n^{*-1} \tag{7.2}$$



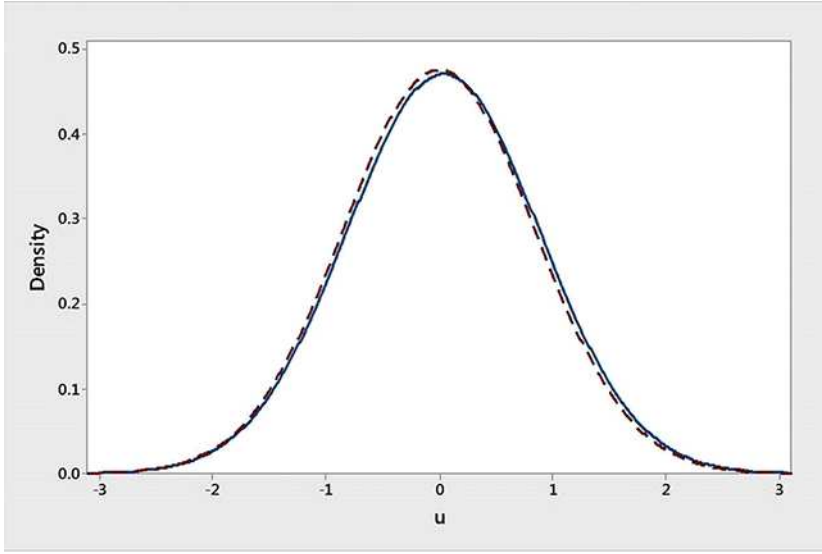
(a)



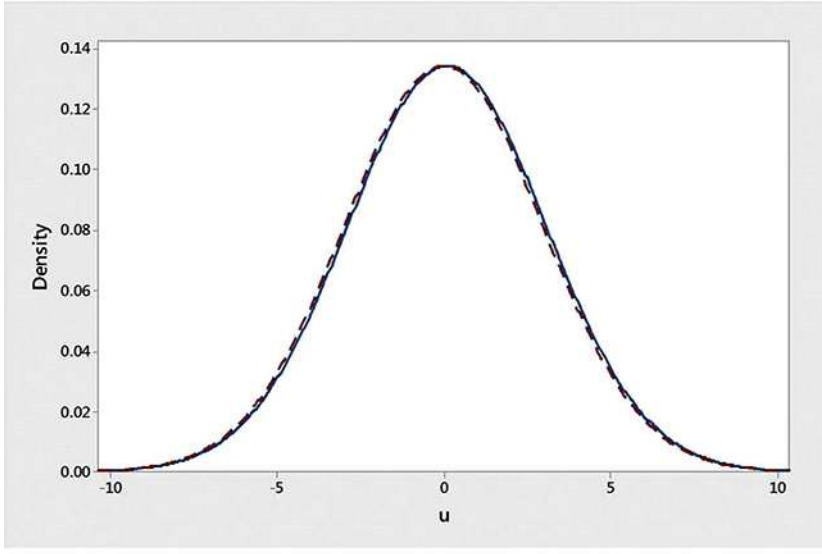
(b)

Figure 1. Plots of normality curves for the modified two-stage procedure (4.1)–(4.2) as validation for (4.6). The dashed curve (red) and the solid curve (blue) respectively correspond to the empirical distribution of the standardized sample size and the $N(0, \sigma_0^2)$ distribution with $\sigma_0^2 = \frac{1}{9}(\frac{\sigma}{\sigma_L})^{k/3}k^2$ coming from (4.4): **(a)** $\sigma_L = 3, k = 1, n^* = 30$; **(b)** $\sigma_L = 3, k = 4, n^* = 500$.

from (5.7) when n^* is large. η_k was defined by (5.5) and it was tabulated in Table 1. Column 7 in Table 9 shows these ε values under each configuration. Upon comparing $\overline{z}$ values from column 6 with ε values from column 7, it appears that the second-order approximation is a little slow to take hold under some configurations. It appears, however, that for all practical purposes, the $\overline{z}$ values are described fairly well by $1 + \varepsilon$, more or less all across the board.



(a)



(b)

Figure 2. Plots of normality curves for the modified two-stage procedure (4.1)–(4.2) as validation for (4.6). The dashed curve (red) and the solid curve (blue) respectively correspond to the empirical distribution of the standardized sample size and the $N(0, \sigma_0^2)$ distribution with $\sigma_0^2 = \frac{1}{9}(\frac{\sigma}{\sigma_L})^{k/3}k^2$ coming from (4.4): **(a)** $\sigma_L = 5, k = 2, n^* = 100$; **(b)** $\sigma_L = 5, k = 5, n^* = 500$.

We now provide Figures 3–4 to validate empirically the normality result described in Theorem 5.1, part (v). We considered four scenarios, namely $m = 5, k = 1, n^* = 30$ (Figure 3(a)); $m = 12, k = 4, n^* = 500$ (Figure 3(b)); $m = 8, k = 2, n^* = 100$ (Figure 4(a)); and $m = 15, k = 5, n^* = 500$ (Figure 4(b)). Under a specific configuration, we again recorded observed values:

$$N = n_i, i = 1, \ldots, H(= 10,000)$$

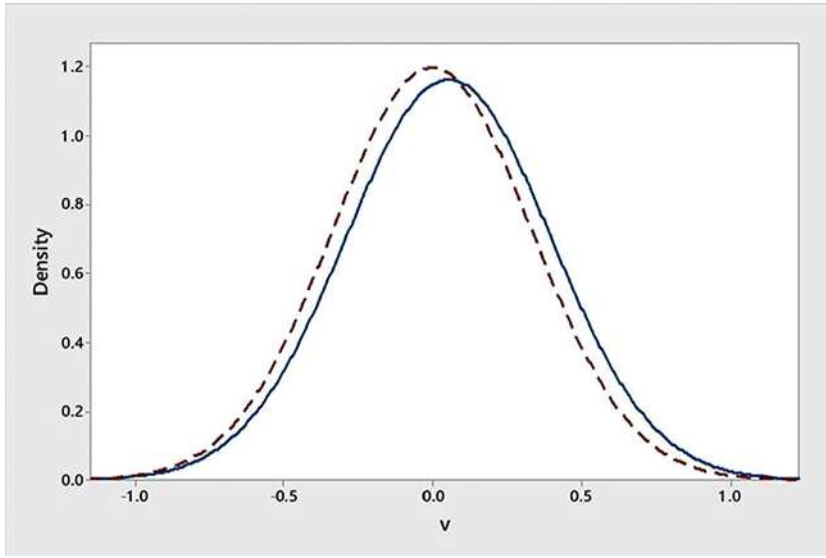
Table 9. Simulation results from 10,000 replications of the purely sequential procedure (5.1) with $\mu = 5$, $\sigma = 10, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}(s_{\bar{x}_{\min}})$	$\bar{n}(s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}(s_{\bar{z}})$	ε in (7.2)
			$\boldsymbol{k} = 1, \boldsymbol{m} = 5$			
20	1.25×10^{-2}	5.5021(0.0050)	20.31(0.0153)	1.0157	1.0421(0.0253)	0.0343
30	3.70×10^{-3}	5.3282(0.0032)	30.31(0.0188)	1.0106	1.0265(0.0243)	0.0228
50	8.00×10^{-4}	5.1999(0.0020)	50.33(0.0241)	1.0066	1.0145(0.0241)	0.0137
100	1.00×10^{-4}	5.0996(0.0009)	100.29(0.0334)	1.0029	1.0079(0.0238)	0.0068
150	2.96×10^{-5}	5.0662(0.0006)	150.38(0.0410)	1.0025	1.0044(0.0226)	0.0045
300	3.70×10^{-6}	5.0334(0.0003)	300.22(0.0582)	1.0007	1.0033(0.0216)	0.0022
400	1.56×10^{-6}	5.0251(0.0002)	400.34(0.0670)	1.0008	1.0015(0.0240)	0.0017
500	8.00×10^{-7}	5.0199(0.0002)	500.42(0.0743)	1.0008	1.0007(0.0228)	0.0013
			$\boldsymbol{k} = 2, \boldsymbol{m} = 8$			
20	1.25×10^{-1}	5.5147(0.0054)	19.98(0.0316)	0.9995	1.2963(0.0381)	0.1703
30	3.70×10^{-2}	5.3443(0.0034)	29.99(0.0386)	0.9997	1.1714(0.0323)	0.1135
50	8.00×10^{-3}	5.2041(0.0020)	50.13(0.0475)	1.0027	1.0734(0.0242)	0.0681
100	1.00×10^{-3}	5.0993(0.0009)	100.03(0.0669)	1.0003	1.0377(0.0227)	0.0340
150	2.00×10^{-4}	5.0671(0.0006)	149.90(0.0842)	0.9993	1.0284(0.0224)	0.0227
300	3.70×10^{-5}	5.0332(0.0003)	300.03(0.1159)	1.0001	1.0120(0.0235)	0.0113
400	1.56×10^{-5}	5.0249(0.0002)	399.96(0.1344)	0.9999	1.0096(0.0220)	0.0085
500	8.00×10^{-6}	5.0201(0.0002)	500.31(0.1483)	1.0006	1.0054(0.0212)	0.0068
			$\boldsymbol{k} = 3, \boldsymbol{m} = 9$			
20	1.25	5.5499(0.0059)	19.66(0.0469)	0.9831	1.3965(0.0880)	0.3882
30	3.70×10^{-1}	5.3574(0.0037)	29.46(0.0585)	0.9822	1.2464(0.0478)	0.2588
50	8.00×10^{-2}	5.2072(0.0021)	49.70(0.0726)	0.9941	1.1724(0.0088)	0.1552
100	1.00×10^{-2}	5.1014(0.0010)	99.78(0.1018)	0.9941	1.0988(0.0515)	0.0776
150	2.00×10^{-3}	5.0663(0.0006)	149.81(0.1240)	0.9987	1.0548(0.0400)	0.0517
300	3.70×10^{-4}	5.0338(0.0003)	299.75(0.1745)	0.9991	1.0270(0.0489)	0.0258
400	1.56×10^{-4}	5.0249(0.0002)	399.82(0.2013)	0.9995	1.0195(0.0375)	0.0194
500	8.00×10^{-5}	5.0203(0.0002)	499.85(0.2252)	0.9997	1.0153(0.0236)	0.0155
			$\boldsymbol{k} = 4, \boldsymbol{m} = 12$			
20	12.5	5.5555(0.0060)	19.66(0.0567)	0.9831	1.7051(0.0608)	0.6889
30	3.70	5.3823(0.0042)	29.10(0.0782)	0.9700	1.4793(0.1318)	0.4592
50	0.80	5.2198(0.0023)	49.06(0.0999)	0.9812	1.2920(0.0858)	0.2755
100	0.10	5.1018(0.0010)	99.46(0.1357)	0.9946	1.1498(0.0283)	0.1377
150	0.02	5.0665(0.0006)	149.13(0.1649)	0.9942	1.1108(0.0253)	0.0918
300	3.70×10^{-3}	5.0339(0.0003)	299.19(0.2305)	0.9973	1.0496(0.0236)	0.0459
400	1.56×10^{-3}	5.0252(0.0002)	399.10(0.2681)	0.9977	1.0378(0.0241)	0.0344
500	8.00×10^{-4}	5.0201(0.0002)	499.68(0.2984)	0.9993	1.0298(0.0234)	0.0275
			$\boldsymbol{k} = 5, \boldsymbol{m} = 15$			
20	125	5.5263(0.0055)	20.53(0.0557)	1.0266	2.0541(0.0431)	1.0722
30	37.03	5.3872(0.0042)	28.78(0.0890)	0.9596	1.7024(0.2270)	0.7148
50	8.00	5.2215(0.0024)	48.21(0.1246)	0.9642	1.4345(0.2552)	0.4288
100	1.00	5.1061(0.0011)	98.38(0.1738)	0.9838	1.2409(0.3578)	0.2144
150	0.20	5.0687(0.0007)	148.77(0.2078)	0.9918	1.1531(0.0529)	0.1429
300	0.037	5.0336(0.0003)	298.65(0.2936)	0.9955	1.0815(0.0422)	0.0714
400	0.015	5.0256(0.0002)	398.38(0.3362)	0.9959	1.0613(0.0410)	0.0536
500	0.008	5.0204(0.0002)	499.83(0.3733)	0.9996	1.0442(0.0394)	0.0428

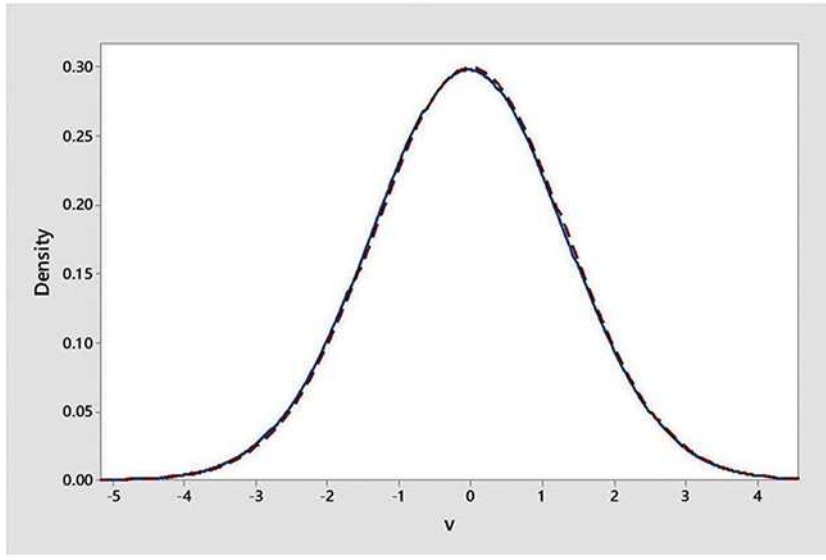
and thus calculated 10,000 associated standardized $(n_i - n^*)/\sqrt{n^*}$ values. Under each specific configuration, such 10,000 observed $v_i \equiv (n_i - n^*)/\sqrt{n^*}$ values provided the empirical distribution of the standardized sample size (dashed curve in red). We superimpose on it the

appropriate theoretical $N(0, \sigma_1^2)$ distributions (solid curve in blue) where $\sigma_1^2 = \frac{1}{9}k^2$ coming from Theorem 5.1, part (v).

The two curves as shown appear slightly off from each other in Figure 3(a) which is meant for small $n^*(= 30)$, however we recall that a visual asymptotic match should be expected for larger n^* (Theorem 5.1, part (v)). Indeed, Figures 3(b) and 4(a)–4(b) show clearly that the

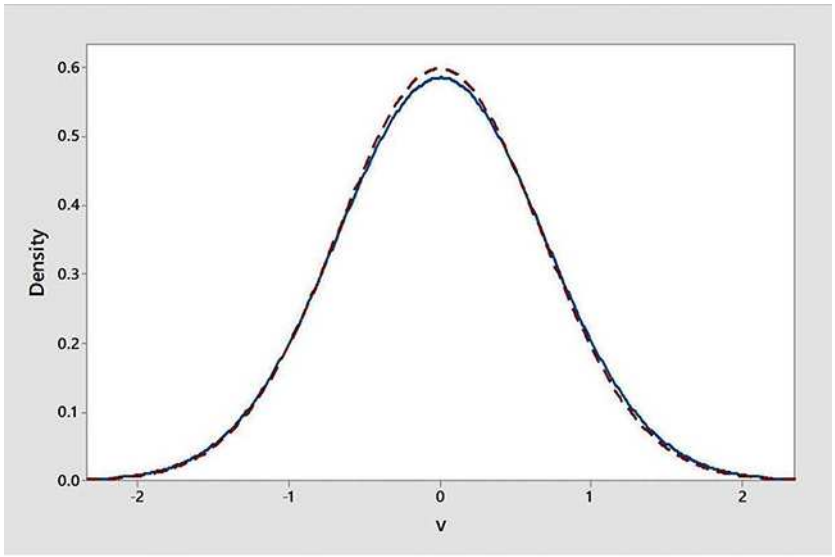


(a)

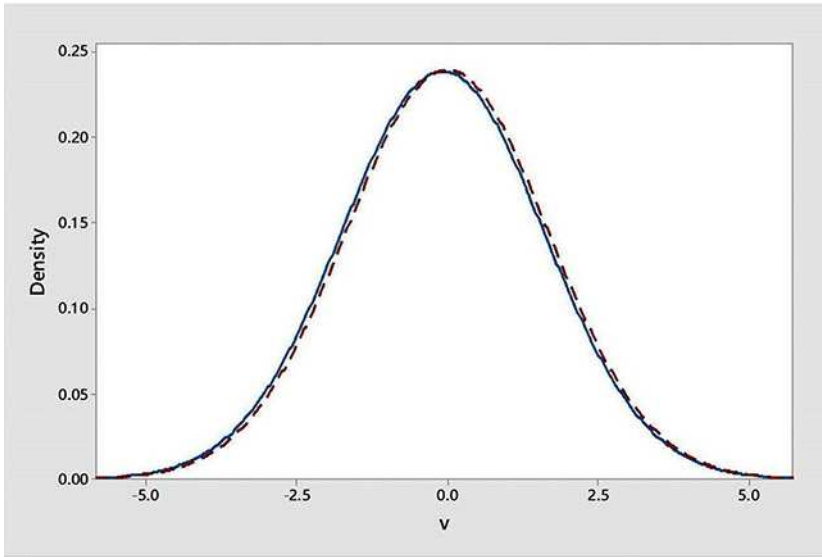


(b)

Figure 3. Plots of normality curves for the purely sequential procedure (5.1) as validation of Theorem 5.1, part (v). The dashed curve (red) and the solid curve (blue) respectively correspond to the empirical distribution of the standardized sample size and the $N(0, \sigma_1^2)$ distribution with $\sigma_1^2 = \frac{1}{9}k^2$ coming from Theorem 5.1, part (v): **(a)** $k = 1, m = 5,\ n^* = 30$; **(b)** $k = 4, m = 12,\ n^* = 500$.



(a)



(b)

Figure 4. Plots of normality curves for the purely sequential procedure (5.1) as validation of Theorem 5.1, part (v). The dashed curve (red) and the solid curve (blue) respectively correspond to the empirical distribution of the standardized sample size and the $N(0, \sigma_1^2)$ distribution with $\sigma_1^2 = \frac{1}{9}k^2$ coming from Theorem 5.1, part (v): **(a)** $k = 2, m = 8,\ n^* = 100$; **(b)** $k = 5, m = 15,\ n^* = 500$.

empirical distribution curve and the theoretical distribution curve nearly lie on each other when $n^*(= 100, 500)$ is large.

On 10,000 observed values of v, we also performed the customary K-S test for normality in the case of each data set that generated Figures 3(a)–3(b) and Figures 4(a)–4(b). We summarize the observed values of K-S test statistic (D) under the null hypothesis of normality with associated p-values.

The p-values shown under cases 3b, 4a, 4b are sizably larger than the p-value shown under case 3a and all four p-values are much larger than 0.05. Our earlier sentiments supported by visual examinations of Figures 3(a)–3(b) and Figures 4(a)–4(b) are clearly validated by the K-S test of normality under each scenario. That is, we are reasonably assured of a good fit between the observed values of v and a normal curve with a high level of confidence for all practical purposes.

8. Data analysis: Illustrations using real data

The modified two-stage estimation methodology (4.1)–(4.2) and the purely sequential estimation methodology (5.1) will now be illustrated using two real data sets from health studies.

The first illustration (Section 8.1) uses infant mortality data from Leinhardt and Wasserman (1979). R documentation for this data may be seen from the website:

$$https : //vincentarelbundock.github.io/Rdatasets/doc/car/Leinhardt.html. \tag{8.1}$$

Our second illustration (Section 8.2) uses bone marrow transplant data from the text Survival Analysis, by Klein and Moeschberger (2003). The data came from a multicenter clinical trial with patients prepared for transplantation with a radiation-free conditioning regimen that consisted of allogeneic marrow transplants for patients with AML.

8.1. Infant mortality data

The data on infant mortality rates were published in *The New York Times* in 1975. These data, named after Leinhardt, were adapted in Leinhardt and Wasserman (1979). Equation (8.1) shows a link for R documentation and a description of data pertaining to 105 nations around 1970.

Because infant mortality rates for Iran, Nepal, Laos, and Haiti were not available, we omitted them from our illustration. The variable of interest (X) was the infant mortality rate per 1,000 live births. The data consisted of $101(= 105 - 4)$ rows and we checked that the negative exponential model (1.1) fitted well.

Treating this dataset as the universe, we first found $\hat{\mu} = 9.6$ and $\hat{\sigma} = 80.24$ from full data. Next, we implemented both modified two-stage and purely sequential estimation procedures drawing observations from the full set of data as needed. It is emphasized, however, that implementation of sampling strategies did not exploit the numbers $\hat{\mu} = 9.6$ and $\hat{\sigma} = 80.24$.

We carried out a single run under both procedures for estimating μ. Tables 10–11 provide the results from implementing the stopping rules from (4.1)–(4.2) and (5.1), respectively corresponding to certain fixed values of c, k, a and the preset risk bound ω, chosen arbitrarily. We also fixed m or m_0 as needed. These numbers are shown in the tables.

Table 10 summarizes the results from the modified two-stage procedure. We assumed a lower bound, $\sigma_L = 40$, for the otherwise unknown scale parameter σ. Table 11 summarizes the results from the purely sequential procedure.

Under both methodologies, we notice that the terminal estimated values of μ are not too far away from corresponding $\hat{\mu} = 9.6$ that was obtained from full data. One other comment is in order: The numbers shown in the first column under n^* are computed using (2.6) after replacing σ with the number $\hat{\sigma} = 80.24$ obtained from full data. Again, in running the estimation methodologies, we did not exploit the number $\hat{\sigma} = 80.24$.

Table 10. Analysis of infant mortality rate data using modified two stage procedure (4.1)–(4.2) with $a = 1$, $c = 0.1, \sigma_L = 40$.

n^*	m_0	k	ω	$\hat{\mu}: x_{n:1}$	n	n/n^*	z
40	4	1	0.01253	9.8	41	1.025	0.9518
50	4		0.00642	9.6	50	1.000	1.0204
50	5	2	0.51507	10.1	52	1.040	0.9064
60	5		0.29807	9.6	57	0.950	1.2523

Table 11. Analysis of infant mortality rate data using purely sequential procedure (5.1) with $a = 1, c = 0.1$.

n^*	m	k	ω	$\hat{\mu}: x_{n:1}$	n	n/n^*	z	ε in (7.2)
40	5	1	0.01253	9.6	41	1.025	1.0245	0.0171
50	5		0.00641	9.7	48	0.960	1.0188	0.0137
50	7	2	0.51507	10.1	54	1.080	1.1073	0.0681
60	7		0.29807	9.8	63	1.050	1.1547	0.0567

We have provided n^* values just so that one is able to gauge whether the observed n-values look reasonable. The ratio n/n^* appears reasonably close to 1, which should be desired. The value z, that is, the ratio of achieved risk per unit cost and preset goal ω appears reasonably under (or close) to 1.

8.2. Bone marrow data

We looked at the data from a multicenter clinical trial involving 137 bone marrow patients prepared for transplant with a radiation-free conditioning regimen that consisted of allogeneic marrow transplants for patients with AML. The data set is explained in Section 1.3 (chapter 1) and enumerated in table D.1 (appendix D), of Klein and Moeschberger's (2003) textbook. For illustrative purposes, we were interested in estimating the minimum waiting time (in days) to death (or time on study time) for these patients. We got rid of two lowest waiting times (1 or 2 days) because they were suspect possible outliers. Other aspects of implementation of our modified two-stage and purely sequential estimation procedures and analysis remained similar to what we explained in Section 8.1.

We checked that the negative exponential model (1.1) fitted well to this data. Treating this data set as our universe, we first found $\hat{\mu} = 10.0$ days and $\hat{\sigma} = 841.57$ days. Next, we implemented both modified two-stage and purely sequential estimation procedures on full data. It is emphasized, however, that implementation of the sampling strategies did not exploit the numbers $\hat{\mu}$ or $\hat{\sigma}$.

We carried out a single run under both procedures for estimating μ. Tables 12–13 provide the results of implementing the stopping rules from (4.1)–(4.2) and (5.1), respectively, corresponding to certain fixed values of c, k, a and the preset risk bound ω chosen arbitrarily. We also fixed m or m_0 as needed. These numbers are shown in the tables.

Table 12 summarizes the results from the modified two-stage procedure. We assumed a lower bound, $\sigma_L = 500$, for the otherwise unknown scale parameter σ. Table 13 summarizes the results for the purely sequential procedure.

Here again, we have provided n^* values just so that one is able to gauge whether the observed n-values look reasonable. The ratio n/n^* appears reasonably close to 1, which should

Table 12. Analysis of time to death (or time on study) in bone marrow data using modified two-stage procedure (4.1)–(4.2) with $a = 1, \sigma_L = 500$.

n^*	m_0	c	k	ω	$\hat{\mu}: x_{n:1}$	n	n/n^*	z
60	4	0.1	1	0.03896	14	57	0.950	1.1871
70	4			0.02453	10	75	1.071	0.9240
80	5	0.4	2	3.45820	15	81	1.013	0.9754
90	5			2.42880	10	89	0.989	1.0458

Table 13. Analysis of time to death (or time on study) in bone marrow data using purely sequential procedure (5.1) with $a = 1$.

n^*	m	c	k	ω	$\hat{\mu}: x_{n:1}$	n	n/n^*	z	ε in (7.2)
60	4	0.1	1	0.03896	16	61	1.017	1.0256	0.0114
70	4			0.02453	10	70	1.000	1.0187	0.0098
80	5	0.4	2	3.45820	11	77	0.963	1.1362	0.0425
90	5			2.42880	10	88	0.978	1.0820	0.0378

be desired. The value z—that is, the ratio of achieved risk per unit cost and preset goal ω—appears reasonably under (or close) to 1.

8.3. Concluding thoughts

We sense more variabilities in Tables 10–11 or Tables 12–13 in comparison with what was reported earlier in Tables 3, 4, and 9, but that should be expected. We may remind ourselves that each row in Tables 10–11 or Tables 12–13 corresponds to a single run of a particular sampling strategy; that is, while implementing either the modified two-stage or the purely sequential estimation methodology. On the other hand, each row in Tables 3, 4, and 10 corresponded to averages obtained from 10,000 identical runs from simulations implementing either the modified two-stage or the purely sequential estimation methodology.

A quick glance at comparing the entries within Tables 10–11 or within Tables 12–13 reveals that the purely sequential estimation strategy appears to have an edge over the modified two-stage estimation strategy. At another level, however, one should realize that a modified two-stage strategy is logistically simpler to implement than a purely sequential estimation strategy. Indeed, both procedures are fully expected to perform very well under comparable experimental circumstances. A practitioner, however, may consider employing one of the two procedures (4.1)–(4.2) or (5.1) that will provide an acceptable level of comfort in running an experiment as one balances it with additional factors deemed locally important, namely, feasibility, efficiency, accuracy, operational convenience, and cost.

Acknowledgments

Acknowledgments Professor Stanley Wasserman and Professor John Klein gave us permission to use infant mortality data from Leinhardt and Wasserman (1979) and to use bone marrow data from Klein-Moeschberger (2003) book respectively. We express our sincere gratitude to these two colleagues. We also thank the Associate Editor and a referee for their encouraging comments.

References

Bain, L. J. (1978). *Statistical Analysis of Reliability and Life Testing Models*, New York: Dekker.

Balakrishnan, N. and Basu, A. P. (1995). *The Exponential Distribution: Theory, Methods and Application*, edited volume, Amsterdam: Gordon and Breach.

Chattopadhyay, S. (1998). Sequential Estimation of Normal Mean under Asymmetric Loss Function with a Shrinkage Stopping Rule, *Metrika* 48: 53–59.

Chattopadhyay, S. (2000). Sequential Estimation of Exponential Location Parameter Using an Asymmetric Loss Function, *Communications in Statistics - Theory & Methods* 29: 783–795.

Chattopadhyay, S., Chaturvedi, A., and Sengupta, R. N. (2000). Sequential Estimation of a Linear Function of Normal Means under Asymmetric Loss Function, *Metrika* 52: 225–235.

Chattopadhyay, S., Datta, S., and Sengupta, R. N. (2005). Asymmetric Penalized Prediction Using Adaptive Sampling Procedures, *Sequential Analysis* 24: 23–43.

Chattopadhyay, S. and Sengupta, R. N. (2006). Three-Stage and Accelerated Sequential Point Estimation of the Normal Mean Using LINEX Loss Function, *Statistics* 40: 39–49.

Chow, Y. S. and Robbins, H. (1965). On the Asymptotic Theory of Fixed Width Sequential Confidence Intervals for the Mean, *Annals of Mathematical Statistics* 36: 457–462.

Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic Normality of Stopping Times in Sequential Analysis, Unpublished Report.

Ghosh, M. and Mukhopadhyay, N. (1981). Consistency and Asymptotic Efficiency of Two-Stage and Sequential Procedures, *Sankhya, Series A* 43: 220–227.

Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*, New York: Wiley.

Hwang, L.-C. and Lee, C.-H. (2012). Bayes Sequential Estimation for One-Parameter Exponential Family under Asymmetric LINEX Loss Function, *Sequential Analysis* 31: 3–21.

Isogai, E., Kobayashi, K., and Uno, C. (2011). Higher Order Approximations by a Two-Stage Procedure for a Negative Exponential Distribution, *Journal of Statistical Planning and Inference* 141: 3304–3312.

Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distributions-2*, New York: Wiley.

Jokiel-Rokita, A. (2008). A Sequential Estimation Procedure for the Parameter of an Exponential Distribution Under Asymmetric Loss Function, *Statistics and Probability Letters* 78: 3091–3095.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis*, second edition, New York: Springer.

Lai, T. L. and Siegmund, D. (1977). A Nonlinear Renewal Theory with Applications to Sequential Analysis I, *Annals of Statistics* 5: 946–954.

Lai, T. L. and Siegmund, D. (1979). A Nonlinear Renewal Theory with Applications to Sequential Analysis II, *Annals of Statistics* 7: 60–76.

Leinhardt, S. and Wasserman, S. S. (1979). Exploratory Data Analysis: An Introduction to Selected Methods, in *Sociological Methodology*, K. Schuessler, ed., pp. 311–365, San Francisco: Jossey-Bass.

Lombard, F. and Swanepoel, J. W. H. (1978). On Finite and Infinite Confidence Sequences, *South African Statistical Journal* 12: 1–24.

Mukhopadhyay, N. (1974). Sequential Estimation of Location Parameter in Exponential Distributions, *Calcutta Statistical Association Bulletin* 23: 85–95.

Mukhopadhyay, N. (1984). Sequential and Two Stage Procedures for Selecting the Better Exponential Population Covering the Case of Unknown and Unequal Scale Parameters, *Journal of Statistical Planning and Inference* 9: 33–43.

Mukhopadhyay, N. (1988). Sequential Estimation Problems for Negative Exponential Populations, *Communications in Statistics-Theory & Methods, Series A* 17: 2471–2506.

Mukhopadhyay, N. (1995). Two-Stage and Multi-Stage Estimation, in *The Exponential Distribution: Theory, Methods and Application*, N. Balakrishnan and A. P. Basu, eds., pp. 429–452, Amsterdam: Gordon and Breach.

Mukhopadhyay, N. (1999). Higher Than Second Order Approximations Via Two-Stage Sampling, *Sankhya, Series A* 61: 254–269.

Mukhopadhyay, N., Datta, S., and Chattopadhyay, S. (2004). *Applied Sequential Methodologies,* edited volume, New York: Dekker.

Mukhopadhyay, N. and de Silva, B. M. (2009). *Sequential Methods and Their Applications*, Boca Raton: CRC.

Mukhopadhyay, N. and Duggan, W. T. (1997). Can a Two-Stage Procedure Enjoy Second Order Properties? *Sankhya, Series A* 59: 435–448.

Mukhopadhyay, N. and Duggan, W. T. (1999). On a Two-Stage Procedure Having Second Order Properties with Applications, *Annals of Institute of Statistical Mathematics* 51: 621–636.

Mukhopadhyay, N. and Hilton, G. F. (1986). Two Stage and Sequential Procedures for Estimating the Location Parameter of a Negative Exponential Distribution, *South African Statistical Journal* 20: 117–136.

Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedures*, New York: Dekker.

The New York Times, 28 September, 1975, p. E-3, Table 3.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Sen, P. K. (1981). *Sequential Nonparametrics*, New York: Wiley.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals,* New York: Springer.

Stein, C. (1945). A Two Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance, *Annals of Mathematical Statistics* 16: 243–258.

Stein, C. (1949). Some Problems in Sequential Estimation (abstract), *Econometrica* 17: 77–78.

Swanepoel, J. W. H. and van Wyk, J. W. J. (1982). Fixed Width Confidence Intervals for the Location Parameter of an Exponential Distribution, *Communications in Statistics - Theory & Methods, Series A* 11: 1279–1289.

Takada, Y. (2000). Sequential Point Estimation of Normal Mean under Linex Loss Function, *Metrika* 52: 163–171.

Takada, Y. (2001). Bayes Sequential Estimation of Poisson Mean under a LINEX Loss Function, *Sequential Analysis* 20: 55–64.

Takada, Y. (2006). Multistage Estimation Procedures with Bounded Risk for the Normal Mean Under LINEX Loss Function, *Sequential Analysis* 25: 227–239.

Takada, Y. and Nagao, H. (2004). Asymptotic Improvement of the Sample Mean Vector for Sequential Point Estimation of a Multivariate Normal Mean with a Linex Loss Function, *Sientiae Mathematicae Japonicae Online* 10: 71–79.

Varian, H. R. (1975). A Bayesian Approach to Real Estate Assessment, in *Studies in Bayesian Econometrics and Statistics*, L. J. Savage, S. E. Fienberg, and A. Zellner, eds., pp. 195–208, North Holland: Elsevier.

Wiener, N. (1939). The Ergodic Theorem, *Duke Mathematical Journal* 5: 1–18.

Woodroofe, M. (1977). Second Order Approximation for Sequential Point and Interval Estimation, *Annals of Statistics* 5: 984–995.

Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, CBMS 39, Philadelphia: SIAM.

Zacks, S. (2009). *Stage-Wise Adaptive Designs*, New York: Wiley.

Zacks, S. and Mukhopadhyay, N. (2009). On Exact and Asymptotic Properties of Two-Stage and Sequential Estimation of the Normal Mean under Linex Loss, *Communications in Statistics - Theory & Methods* 38: 2992–3014.

Zelen, M. (1966). Application of Exponential Models to Problems in Cancer Research, *Journal of Royal Statistical Society* 129: 368–398.

Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Functions, *Journal of American Statistical Association* 81: 446–451.