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 estimation of the difference of locations of two negative exponential populations under a modified Linex loss function: Real data illustrations from cancer studies and reliability analysis

Nitis Mukhopadhyay & Sudeep R. Bapat

SEQUENTIAL ANALYSIS
2016, VOL. 35, NO. 3, 387–412
http://dx.doi.org/10.1080/07474946.2016.1206386

Taylor & Francis
Taylor & Francis Group

Multistage estimation of the difference of locations of two negative exponential populations under a modified Linex loss function: Real data illustrations from cancer studies and reliability analysis

Nitis Mukhopadhyay and Sudeep R. Bapat

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

ABSTRACT

We have designed modified two-stage and purely sequential strategies to estimate the difference of location parameters from two independent negative exponential populations having unknown but proportional scale parameters under a modified Linex loss function. This article extends one-sample methodologies of Mukhopadhyay and Bapat (2016, *Sequential Analysis*). Some preliminary results are established along the lines of Mukhopadhyay and Hamdy (1984, *Canadian Journal of Statistics*) and Mukhopadhyay and Darmanto (1988, *Sequential Analysis*). We have resorted to Mukhopadhyay and Duggan (1997, *Sankhya, Series A*) in developing asymptotic second-order properties for the modified two-stage methodology and to nonlinear renewal theory of Lai and Siegmund (1977; 1979, *Annals of Statistics*) and Woodroofe (1977, *Annals of Statistics*) in addressing analogous properties under the purely sequential methodology. Then, we supplement with extensive sets of data analysis via computer simulations validating that both modified two-stage and purely sequential methods perform very well. Both methodologies are also illustrated and implemented using real datasets from cancer studies and reliability analysis.

ARTICLE HISTORY
Received 22 February 2016
Revised 17 April 2016,
27 May 2016
Accepted 10 June 2016

KEYWORDS
Cancer research; first-order properties; Linex loss; location parameter; modified Linex; modified two-stage; negative exponential; nonlinear renewal theory; purely sequential; real data; reliability analysis; risk per unit cost; scale parameter; second-order properties; simulations; two-stage

MATHEMATICS SUBJECT CLASSIFICATIONS
62L12; 62L05; 62G20; 62F10; 62P10; 62P30.

1. Introduction

We develop sequential and two-stage procedures to estimate the difference of two independent negative exponential locations under a modified Linex loss function. Many papers have mainly devoted their attentions to normal populations, largely under a different form of the loss function. Some of these include Mukhopadhyay et al. (2010) and Aoshima et al. (2011). Some notable papers pertaining to two-sample estimation problems under negative exponential settings include Mukhopadhyay and Hamdy (1984) and Mukhopadhyay and Darmanto (1988).

Negative exponential distributions often play a central role in reliability, life testing, and clinical experiments. For an elaborate review, one may refer to Zelen (1966); Johnson and Kotz (1970); Lawless and Singhal (1980), Balakrishnan and Basu (1995), and other sources. We provide illustrations with real data in Section 7.

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

Sequential and two-stage procedures for estimating a negative exponential location have been discussed extensively in Basu (1971); Ghurye (1958); Mukhopadhyay (1974, 1980, 1984, 1988, 1995), Swanepoel and van Wyk (1982), and other sources. A sequential point estimation analog 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).

Point estimation of a negative exponential location parameter under a variant of the Linex loss function has been introduced in Mukhopadhyay and Bapat (2016). One may refer to Mukhopadhyay and Bapat (2016) for additional explanations and references. The present investigation expands their ideas by developing appropriate methodologies and associated properties in the context of two-sample comparisons.

1.1. Two-sample scenario and the goal

Let $X_1, \ldots, X_n, \ldots$ and $Y_1, \ldots, Y_n, \ldots$ be two independent and identically distributed (i.i.d.) sequences of random variables with the probability density functions (p.d.f.s) $f(x; \mu_1, \sigma)$ and $f(y; \mu_2, b\sigma)$ with $b(> 0)$ respectively, where we denote

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

where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$ are both *unknown parameters*. $I(.)$ denotes an indicator function of (.) which takes the value 1(or 0) when $t >$ (or $\leq$)μ.

Additionally, we assume that the Xs are independent of the Ys and b is *known*. The location parameter μ, if positive, is viewed as the minimum guarantee time or the threshold of the distribution. The parameter σ is called the scale. We record a pair (X_i, Y_i) at-a-time or a group of pairs of observations (X_i, Y_i) at-a-time as needed, $i = 1, \ldots, n, \ldots$. Our goal is to estimate the "difference" parameter, namely:

$$\Delta = \mu_1 - \mu_2 \tag{1.2}$$

1.2. Motivation for a scenario where b is known a'priori

In Figure 1, we show columns from Parthenon, Greece. In Figure 1(a), a single column is shown, whereas in Figure 1(b) a cluster of eight columns is shown. These columns are identically constructed. In large structures (for example, bridges), one will often see a column or a cluster of columns strategically placed to withstand stress. This is done to raise a structure that may not fall apart easily with a high probability under reasonable application of stress levels.

Now, visualize Figure 1(a), expose the column to a stress level (L), and observe the time (U) it takes the column to show up cracks. We may reasonably assume that U is governed by the p.d.f. $f(u; \mu_1, \sigma)$ defined by (1.1); that is, we will not observe a failure before time μ_1.

Next, we may visualize Figure 1(b), expose eight pillars to a stress level (L) each, and observe the time (V) it takes the cluster of columns to show fatal cracks. Consider independent $U_1, \ldots, U_8$ distributed identically as U. Then, the integrity of the cluster in Figure 1(b) may be described by $V = \min\{U_1, \ldots, U_8\} \sim f(v; \mu_2, \frac{1}{8}\sigma)$, the time of failure of the weakest column among all eight columns. We will not observe a failure of the cluster before time μ_2.

(a) (b)

Figure 1. Columns from Parthenon, Greece: (a) One column by itself (U), (b) A cluster of eight identical columns (V). Photo courtesy: Nitis Mukhopadhyay.

Under a stress test conducted in a laboratory, we may want to estimate $\Delta = \mu_1 - \mu_2$ by observing independent pairs of observations $\{(X_i, Y_i); i = 1, \ldots, n, \ldots\}$ where we begin by first identifying X with U and Y with V, respectively. That is, X_i $(Y_i) =$ ith observation on U (V), $i = 1, \ldots, n, \ldots$. This setup clearly agrees with a two-sample scenario described in Section 1.1 involving a known multiplier, $b = \frac{1}{8}$. We emphasize that b needs to be specified a priori before data collection begins. Often b will be specified by the subject-matter expert(s) based on personal knowledge from previously run similar studies.

In a situation where it may be more reasonable to postulate that $X \sim f(x; \mu_1, \sigma)$ and $Y \sim f(y; g\mu_2, h\sigma)$ with both g, h positive and known, then obviously $Y^* \equiv g^{-1}Y \sim f(y^*; \mu_2, b\sigma)$ with $b = g^{-1}h$ positive and known. By exploiting one-to-one coding from Y to Y^* and then working with X, Y^* instead of X, Y would bring one back to our proposed two-sample estimation of $\Delta = \mu_1 - \mu_2$.

1.3. More review and layout of the article

A customary Linex loss function was first proposed by Varian (1975). A Linex loss is useful in situations where one assigns asymmetric penalty due to bias by weighing in overestimation and underestimation unequally. Interesting applications of this loss can be seen in Varian (1975) and Zellner (1986).

Before we go any further, we explain our notation clearly. Since μ_1, μ_2, σ are all unknown, the parameter vector $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma)$ 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 $\xrightarrow{\pounds}$ (convergence in law or distribution), they are all with respect to $P_{\boldsymbol{\theta}}$. We drop subscript $\boldsymbol{\theta}$ in the sequel for simplicity.

In Section 2, we include some preliminaries of a modified Linex loss function in the spirit of Mukhopadhyay and Bapat (2016), formulation of the risk function, and its optimization by bounding it from above. Section 3 introduces a modified two-stage procedure in which we assume a *known* lower bound $\sigma_L (> 0)$ for the unknown standard deviation σ along the lines of Mukhopadhyay and Duggan (1997, 1999). Both first- and second-order properties are developed for estimating Δ.

Section 4 introduces a purely sequential methodology in the present context for estimating Δ. For some of the technicalities and second-order properties, we have relied upon Mukhopadhyay (1974, 1984, 1988), Lombard and Swanepoel (1978), and Swanepoel and van Wyk (1982). Generally speaking, the broad literature on sequential estimation as well as many associated tools of trade may be reviewed from Mukhopadhyay and Solanky (1994), Ghosh et al. (1997) and Mukhopadhyay and de Silva (2009).

In Section 5, we outline proofs for some selected results. Section 6 presents extensive data analysis based upon computer simulations for a large variety of parameter configurations highlighting the performance of the proposed methodologies for a wide range of small, moderate, and large sample sizes. Selected conclusions from the theorems studied in Sections 3–4 are critically examined and validated with data analysis.

Section 7 shows analysis on two real datasets, one from cancer studies and the other from reliability analysis, both supporting the proposed methodologies. The first illustration (Section 7.1) uses parallel data sets related to survival times of cancer patients from Shanker et al. (2016). The second illustration (Section 7.2) uses data on lifetimes of steel components from the text of Lawless (1982). Section 7.3 gives brief conclusions.

2. Modified Linex loss and some preliminaries

In this section, we work under an appropriately modified Linex loss function in the spirit of Mukhopadhyay and Bapat (2016) and then calculate the associated risk function. For motivation, one may refer to that article.

2.1. Equal sample size and the estimators

Having recorded pairs of random samples $\{X_i, Y_i; i = 1, \ldots, n\}, n > 1$, we estimate the difference parameter $\Delta \equiv \mu_1 - \mu_2$ from (1.2) with the maximum likelihood estimator

$$\widehat{\Delta}_n \equiv X_{n:1} - Y_{n:1}$$

where $X_{n:1} = \min\{X_1, \ldots, X_n\}$ and $Y_{n:1} = \min\{Y_1, \ldots, Y_n\}$.

The minimum variance unbiased estimator of σ is given by

$$\hat{\sigma}_n \equiv W_n = \frac{1}{2}\left(U_n + b^{-1} V_n\right),$$

where we let

$$U_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X_{n:1}) \quad \text{and} \quad V_n = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - Y_{n:1}). \tag{2.1}$$

2.2. Modified Linex loss function

Let us begin with some fixed a, positive or negative. In the spirit of Mukhopadhyay and Bapat (2016), a *modified Linex loss* function in estimating Δ from (1.2) with $\widehat{\Delta}_n \equiv X_{n:1} - Y_{n:1}$ is formulated as follows:

$$L_n = \exp\left(a\left[\sigma^{-1}(\widehat{\Delta}_n - \Delta)\right]\right) - a\left[\sigma^{-1}(\widehat{\Delta}_n - \Delta)\right] - 1. \tag{2.2}$$

The corresponding risk function is obtained by taking expectations across (2.2), namely,

$$\text{Risk}_n \equiv E[L_n] = \left(1 - \frac{a}{n}\right)^{-1}\left(1 + \frac{ab}{n}\right)^{-1} - \frac{a}{n} + \frac{ab}{n} - 1. \tag{2.3}$$

Upon expanding (2.3), we clearly obtain

$$\text{Risk}_n = \frac{a^2(1 + b^2 - b)}{n^2} + o\left(\frac{1}{n^2}\right) \tag{2.4}$$

for $n > \max\{1, a, -ab\}$.

2.3. Cost per unit sampling

Next, we consider a cost function, $\text{Cost}_n(> 0)$. The exact form of this function must depend upon the underlying practical scenario. It may be acceptable to assume that the cost for each observation should go up (or down) as σ goes down (or up). With this in mind, we propose a cost function of the following form:

$$\text{Cost}_n = cn\sigma^{-k} \text{ with fixed and known } c(> 0), k(> 0). \tag{2.5}$$

One may alternatively think that Cost_n should have been $c(2n)\sigma^{-k}$ instead of $cn\sigma^{-k}$ because after all we record $2n$ observations. In addition, the cost per unit observation on X-data could be surely different from the cost per unit observation on Y-data giving rise to an expression such as $(c_1 n + c_2 n)\sigma^{-k}$ instead of $cn\sigma^{-k}$. However, Cost_n from (2.5) is general enough because in our formulation, the choice of c remains generic, covering easily those other possibilities.

2.4. Proposed criterion: Bounded risk per unit cost

Adapting the basic formulation from Mukhopadhyay and Bapat (2016), we bound the associated "risk" where we interpret risk as the *risk per unit cost* (RPUC), namely,

$$RPUC_n \equiv \frac{\text{Risk}_n}{\text{Cost}_n} = \frac{a^2(1 + b^2 - b)}{n^2}\frac{\sigma^k}{cn} + o\left(n^{-3}\right). \tag{2.6}$$

We require that $\text{RPUC}_n \leq \omega$ for all parameter vector $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma)$ where $\omega(> 0)$ is a fixed constant, which leads us to determine the required optimal fixed sample size had σ been known as follows:

$$n \geq \left(\frac{a^2(1 + b^2 - b)}{c\omega}\right)^{1/3} \sigma^{k/3} = n^*, \text{ say}. \tag{2.7}$$

However, the magnitude of n^* is unknown even though its expression is known. Hence, we proceed to develop modified two-stage and purely sequential bounded risk estimation strategies in Sections 3–4.

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

The idea is to estimate the optimal fixed sample size n^* given by (2.7). We begin with pilot data of appropriate size $m(> \max\{1, a, -ab\})$ from both X, Y and then move forward step by step with the help of implementing a modified two-stage or a purely sequential sampling strategy to record more data subsequently as needed beyond the pilot stage.

Suppose that a final sample size, denoted by a random variable Q, is determined by an adaptive multistage sampling strategy. Then, the next result shows an exact analytical expression for the risk per unit cost associated with the terminal estimator $\widehat{\Delta}_Q = X_{Q:1} - Y_{Q:1}$ for Δ once sampling is terminated.

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 $q \ge m(> \max\{1, a, -ab\})$. Then, the expression for the risk per unit cost associated with the terminal estimator $\widehat{\Delta}_Q$ is given by*

$$E\left[RPUC_Q\right] = \omega \frac{n^{*3}}{a^2(1 + b^2 - b)} \left\{ E\left[\frac{1}{Q}\left(1 - \frac{a}{Q}\right)^{-1}\left(1 + \frac{ab}{Q}\right)^{-1}\right] - E\left(\frac{a - ab}{Q^2}\right) - E\left(\frac{1}{Q}\right)\right\}. \quad (2.8)$$

Its proof is similar along the lines of Section 6.1 of Mukhopadhyay and Bapat (2016) and hence it is omitted for brevity.

3. A modified two-stage procedure

The customary two-stage procedure developed by Stein (1945, 1949) has some advantages and some disadvantages that were elaborated in Mukhopadhyay and Bapat (2016). We thus resort to a modified two-stage procedure along the lines of Mukhopadhyay and Duggan (1997) to achieve attractive second-order properties.

The 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 (2.7), we note that

$$n^* > \left(a^2(1 + b^2 - b)(c\omega)^{-1}\right)^{1/3} \sigma_L^{k/3}.$$

Thus, we fix an integer $m_0(> \max\{1, a, -ab\})$ and choose the pilot size m in such a way that $m \approx \left(a^2(1 + b^2 - b)(c\omega)^{-1}\right)^{1/3} \sigma_L^{k/3}$. We formally define m as follows:

$$m \equiv m(\omega) = \max\left\{m_0, \left\lfloor d_\omega \sigma_L^{k/3} \right\rfloor + 1\right\} \text{ with } d_\omega = \left(a^2(1 + b^2 - b)(c\omega)^{-1}\right)^{1/3}, \quad (3.1)$$

and gather pilot data $\{X_i, Y_i; i = 1, \ldots, m\}$. Here and elsewhere, $\lfloor s \rfloor$ denotes the largest integer less than $s(> 0)$.

From pilot data, we obtain the statistic $\hat{\sigma}_m \equiv W_m = \frac{1}{2}\left(U_m + b^{-1}V_m\right)$ where U_m and V_m are defined via (2.1) and then determine the terminal sample size N as follows:

$$N \equiv N(\omega) = \max\left\{m, \left\lfloor d_\omega W_m^{k/3} \right\rfloor + 1\right\} \quad (3.2)$$

which is an estimator of n^* defined in (2.7).

If $N = m$, we would not require additional observations at the second stage. However, if $N > m$, then we sample the difference $N - m$ at the second stage by recording an additional set of pairs of observations $\{(X_i, Y_i), i = m+1, \ldots, N\}$. From full data $\{(X_i, Y_i), i = 1, \ldots, N\}$ obtained by combining both stages, we propose to estimate Δ by the difference of the smallest order statistics, namely:

$$\widehat{\Delta}_N \equiv X_{N:1} - Y_{N:1} = \min\{X_1, \ldots, X_N\} - \min\{Y_1, \ldots, Y_N\}.$$

Along the lines of (2.2), the associated loss function will be

$$L_N = \exp\left(a\left[\sigma^{-1}(\widehat{\Delta}_N - \Delta)\right]\right) - a\left[\sigma^{-1}(\widehat{\Delta}_N - \Delta)\right] - 1, \tag{3.3}$$

A major difference between (2.2) and (3.3) is that the sample size N used in (3.3) is a random variable unlike n.

Now, since $\widehat{\Delta}_n$ and $I(N = n)$ are independent for all fixed $n \geq m$, using (2.8) with Q replaced by N from (3.2), we get

$$\begin{aligned}\omega^{-1}E\left[\text{RPUC}_N\right] = \frac{n^{*3}}{a^2(1+b^2-b)} &\left\{E\left[\frac{1}{N}\left(1-\frac{a}{N}\right)^{-1}\left(1+\frac{ab}{N}\right)^{-1}\right]\right. \\ &\left. -E\left(\frac{a-ab}{N^2}\right) - E\left(\frac{1}{N}\right)\right\}. \end{aligned}\tag{3.4}$$

3.1. First-order asymptotics

We begin with some attractive first-order results involving N. One will surely note that there is no stated sufficient condition involving m in Theorem 3.1. This is so because in the present setup, we have $m \equiv m(\omega) \to \infty$ as $\omega \to 0$.

Theorem 3.1. *With m and N respectively defined in (3.1) and (3.2), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ 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 (2.7).

Its proof follows in similar lines as in Section 6.3 of Mukhopadhyay and Bapat (2016) and we thus leave it out for brevity. One clearly sees that all of the expressions in Theorem 3.1 converge to 1, which looks very encouraging because we can expect N and n^* to be in the same ball park. Indeed, 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.

3.2. Second-order asymptotics

The second order-asymptotics are similar to what is presented in equation (4.3) of Mukhopadhyay and Bapat (2016), with N replaced from (3.2), and we leave them out for brevity. One should note that all such asymptotics are readily accessible along the lines of Mukhopadhyay and Duggan (1997, 1999).

Next, let us define the expressions for ψ and σ_0^2 as follows:

$$\psi = \left(\frac{c}{a^2(1+b^2-b)}\right)^{1/3}\left\{\frac{k}{12}\left(\frac{k}{3}-1\right)\left(\frac{a^2(1+b^2-b)}{c}\right)^{1/3}\right\}\left(\frac{\sigma}{\sigma_L}\right)^{k/3} \quad \text{and}$$

$$\sigma_0^2 = \frac{k^2}{18}\left(\frac{\sigma}{\sigma_L}\right)^{k/3}. \tag{3.5}$$

For completeness, now we exhibit both lower and upper bounds for $E(N-n^*)$.

Theorem 3.2. *With m and N respectively defined in (3.1) and (3.2), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ we have as $\omega \to 0$*

$$\psi + O(\omega^{1/2}) \le E(N-n^*) \le \psi + 1 + O(\omega^{1/2}) \text{ [asymptotic second} - \text{order efficiency]}, \tag{3.6}$$

where ψ is defined in (3.5) and n^* comes from (2.7).

Theorem 3.2 shows the second-order efficiency property of the modified two-stage procedure (3.1)–(3.2) in the sense of Ghosh and Mukhopadhyay (1981). Next, we look at a result which obtains the asymptotic distribution of a standardized version of N along the lines of Ghosh and Mukhopadhyay (1975) and Mukhopadhyay and Duggan (1997, 1999).

Theorem 3.3. *With m and N respectively defined in (3.1) and (3.2), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ we have as $\omega \to 0$:*

$$U^* \equiv n^{*-1/2}(N-n^*) \xrightarrow{\mathcal{L}} N(0, \sigma_0^2), \tag{3.7}$$

where σ_0^2 is defined in (3.5) and n^* comes from (2.7).

Its proof follows from Lemma 2.1, part (i) in Mukhopadhyay and Duggan (1999) and hence it is omitted. The following theorem evaluates the risk per unit cost up to second-order approximation. The proof follows along the lines of section 6.4 in Mukhopadhyay and Bapat (2016) and is thus left out for brevity.

Theorem 3.4. *With m and N respectively defined in (3.1) and (3.2), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ we have as $\omega \to 0$*

$$1 + \frac{1}{n^*}\left(6\sigma_0^2 - 3\psi - 3 + a\left(\frac{b^2-b^3-b+1}{1+b^2-b}\right)\right) + o\left(\frac{1}{n^*}\right) \le \omega^{-1}E\left[RPUC_N\right]$$
$$\le 1 + \frac{1}{n^*}\left(6\sigma_0^2 - 3\psi + a\left(\frac{b^2-b^3-b+1}{1+b^2-b}\right)\right) + o\left(\frac{1}{n^*}\right)$$
$$\text{[asymptotic second} - \text{order risk efficiency]} \tag{3.8}$$

where ψ, σ_0^2 are as defined in (3.5) and n^* comes from (2.7).

4. A purely sequential procedure

A purely sequential procedure provides tighter results compared to those (Theorem 3.4) for the modified two-stage procedure. This happens perhaps because this methodology allows us to take as many pairs of observations sequentially step by step as required depending on the rule of termination.

In this section, we will pursue a purely sequential sampling strategy along the lines of Mukhopadhyay and Bapat (2016) to address our two-sample problem. We will 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 (2.7). We again fix an integer $m(> \max\{1, a, -ab\})$ and obtain pilot data $\{X_i, Y_i; i = 1, \ldots, m\}$. We then proceed by recording one additional pair (X, Y) at every step as needed determined by the following rule:

$$N \equiv N(\omega) = \inf\left\{n \geq m : n \geq d_\omega W_n^{k/3}\right\} \text{ with } d_\omega = \left(a^2(1 + b^2 - b)(c\omega)^{-1}\right)^{1/3}. \quad (4.1)$$

Again, this stopping variable N estimates n^*. From full data $\{(X_i, Y_i), i = 1, \ldots, N\}$ gathered upon termination of sampling, we propose to estimate Δ by the difference of the smallest order statistics, namely:

$$\widehat{\Delta}_N \equiv X_{N:1} - Y_{N:1} = \min\{X_1, \ldots, X_N\} - \min\{Y_1, \ldots, Y_N\}.$$

Along the lines of (2.2), the associated loss function will be

$$L_N = \exp\left(a\left[\sigma^{-1}(\widehat{\Delta}_N - \Delta)\right]\right) - a\left[\sigma^{-1}(\widehat{\Delta}_N - \Delta)\right] - 1.$$

Since $\widehat{\Delta}_n$ and $I(N = n)$ are independent for all fixed $n \geq m$, the associated expression for $E\,[\text{RPUC}_N]$ will again resemble (2.8) with Q replaced by N from (4.1).

4.1. First-order asymptotics

The first-order asymptotics that we are now presenting resemble Theorem 5.1 in Mukhopadhyay and Bapat (2016). The proof of this theorem follows along same lines as in their Section 6.5 and hence we show a partial derivation in Section 5.1.

Theorem 4.1. *For N defined in (4 .1), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ we have as $\omega \to 0$.*

(i) $N/n^* \xrightarrow{P} 1; n^*/N \xrightarrow{P} 1;$
(ii) $E\left[(N/n^*)^t\right] \to 1$ *for* $t > 0$ *if* $m(> \max\{2, a, -ab\})$ [*asymptotic first-order efficiency*];
(iii) $E\left[(n^*/N)^t\right] \to 1$ *for* $t > 0$ *if* $m > \max\{1 + \frac{1}{3}kt, a, -ab\}$;
(iv) $\omega^{-1}E\,[RPUC_N] \to 1$ *if* $m > \max\{1+\frac{4}{3}k, a, -ab\}$ [*asymptotic first-order risk efficiency*];
(v) $V^* \equiv n^{*-1/2}(N - n^*) \xrightarrow{\mathcal{L}} N(0, \sigma_1^2)$ *with* $\sigma_1^2 = \frac{k^2}{18}$
where n^* *comes from (2.7).*

The expressions shown in Theorem 4.1, parts (i)–(iv) converge to 1, which are parallel to the results from Theorem 3.1 under the modified two-stage procedure (3.1)–(3.2). In other words, the purely sequential procedure (4.1) also has attractive asymptotic first-order properties.

Now, in order to contrast the asymptotic normality properties of standardized N of the modified two-stage procedure (3.1)–(3.2) and the purely sequential procedure (4.1), first recall from Theorem 3.3 that

$$U^* \equiv n^{*-1/2}(N - n^*) \xrightarrow{\mathcal{L}} N(0, \sigma_0^2) \quad \text{with } \sigma_0^2 = \frac{k^2}{18}\left(\frac{\sigma}{\sigma_L}\right)^{k/3}.$$

As opposed to that, Theorem 4.1, part (v) shows that

$$V^* \equiv n^{*-1/2}(N - n^*) \xrightarrow{\mathcal{L}} N(0, \sigma_1^2) \quad \text{with } \sigma_1^2 = \frac{k^2}{18}.$$

Thus, we clearly observe that the asymptotic normal distribution for the standardized stopping variable V^* definitely has a smaller asymptotic variance (σ_1^2) compared to that (σ_0^2) for U^*. This gives an edge of the purely sequential strategy over the modified two-stage strategy.

4.2. Second-order asymptotics

Going back to applications of nonlinear renewal theoretic results provided in Mukhopadhyay and Bapat (2016) for a one-sample problem, we express N from (4.1) as $R + 1$ w.p.1 where

$$\begin{aligned} R &= \inf\left\{n \geq m - 1 : \Sigma_{i=1}^{n} Z_i \leq n^{*-3/k} n^{\frac{3}{k}+1}(1 + \frac{1}{n})^{\frac{3}{k}}\right\} \\ &= \inf\left\{n \geq m - 1 : \Sigma_{i=1}^{n} Z_i \leq h^* n^{\delta} L(n)\right\}, \end{aligned} \tag{4.2}$$

with the Z_is being i.i.d. gamma$(2, \frac{1}{2})$ random variables. A brief explanation is laid out inside the proof of Theorem 4.1, part (iii) in Section 5.1.

Using Woodroofe (1977), Lai and Siegmund (1977, 1979), and especially the representations laid out in Mukhopadhyay (1988) and (Mukhopadhyay and Solanky, 1994, Section 2.4.2), clearly R from (4.2) matches with (2.4.7) in Mukhopadhyay and Solanky (1994) where

$$\delta = \frac{3}{k} + 1, \quad h^* = n^{*-3/k}, \quad L(n) = 1 + \frac{3}{kn} + o\left(\frac{1}{n}\right), \quad \text{so that } L_0 = \frac{3}{k}, \tag{4.3}$$

and also note

$$\theta = 1, \quad \tau^2 = \frac{1}{2}, \quad \beta^* = \frac{k}{3}, \quad n_0^* = n^* \quad \text{and} \quad p = \frac{k^2}{18}. \tag{4.4}$$

Condition (2.5) from Mukhopadhyay (1988) or, equivalently, (2.4.8) from Mukhopadhyay and Solanky (1994) is satisfied with $B = 2$ and $b = 1$. We define two special entities:

$$\begin{aligned} \nu \equiv \nu_k &= \frac{k}{6}\left(\frac{9}{k^2} + 1\right) - \frac{1}{4}\Sigma_{n=1}^{\infty} n^{-1} E\left\{\max\left(0, \chi_{4n}^2 - 4n\left(\frac{3}{k} + 1\right)\right)\right\}, \quad \text{and} \\ \eta \equiv \eta_k &= \frac{1}{3}k\nu - \left(\frac{1}{36}k^2 + \frac{1}{12}k + 1\right). \end{aligned} \tag{4.5}$$

along the lines of (2.4.9)–(2.4.10) in Mukhopadhyay and Solanky (1994) and (5.5) in Mukhopadhyay and Bapat (2016). Table 1 illustrates a few values of $\nu \equiv \nu_k$ and $\eta \equiv \eta_k$ with $k = 1, 2, 3, 4, 5, 6$.

Table 1. Selected values of $\nu \equiv \nu_k$ and $\eta \equiv \eta_k$ from (4.5), $k = 1(1)6$.

	k					
	1	2	3	4	5	6
ν_k	3.082	1.638	1.171	0.942	0.806	0.716
η_k	−0.084	−0.185	−0.328	−0.521	−0.768	−1.068

To conclude this section, we now specify asymptotic second-order expansions for both positive and negative moments of $\frac{N}{n^*}$ (Theorem 4.2) as well as an asymptotic second-order expansion of the risk per unit cost (Theorem 4.3) under the purely sequential setting (4.1). Recall that we must also have $m(> \max\{1, a, -ab\})$ satisfied in the background for Theorems 4.2–4.3 to hold. Brief outlines of proofs of Theorems 4.2–4.3 are shown in Section 5.2.

Theorem 4.2. *For N defined in (4.1), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, b, a$ and every non-zero real number t, we have as $\omega \to 0$:*

$$E\left[\left(N/n^*\right)^t\right] = 1 + \left\{t\eta_k + t + \frac{1}{2}t(t-1)p\right\} n^{*-1} + o\left(n^{*-1}\right)$$

$$[\textit{asymptotic second-order efficiency}], \tag{4.6}$$

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 $t = -1$, with n^*, p, and η_k coming from (2.7), (4.4), and (4.5) respectively.

Theorem 4.3. *For N defined in (4.1), for each fixed value of $\mu_1, \mu_2, \sigma, c, k, a$, we have the following second-order expansion of the risk per unit cost as $\omega \to 0$:*

$$\omega^{-1}E\left[RPUC_N\right] = 1 + \left(6p - 3\eta_k - 3 + a\left(\frac{b^2 - b^3 - b + 1}{1 + b^2 - b}\right)\right) n^{*-1} + o\left(n^{*-1}\right)$$

$$[\textit{asymptotic second-order risk efficiency}] \tag{4.7}$$

when $m > \max\left\{\frac{7k}{3} + 1, a, -ab\right\}$ with n^*, p, and η_k coming from (2.7), (4.4), and (4.5), respectively.

5. Selected technicalities

Since the technicalities are largely similar to those in Mukhopadhyay and Bapat (2016), we show some steps selectively. Intermediate steps are kept out for brevity.

5.1. Proof of Theorem 4.1

Part (i): Follows from Lemma 1 of Chow and Robbins (1965).
Part (ii): With $m \geq 3$ we can claim that $\frac{N-1}{N-2} \leq 2$ w.p.1, and let

$$H^* = \sup_{n \geq 2}\left\{\frac{1}{n}\Sigma_{i=1}^n(X_i - \mu_1) + \frac{1}{bn}\Sigma_{i=1}^n(Y_i - \mu_2)\right\},$$

for sufficiently small $\omega(>0)$ so that $n^* > m$, observe the following inequality (w.p.1):

$$N \le m + d_\omega W_{N-1}^{k/3} \le m + 2^{k/3} d_\omega \left\{ \frac{1}{N-2} \Sigma_{i=1}^{N-1}(X_i - \mu_1) + \frac{1}{b(N-2)} \Sigma_{i=1}^{N-1}(Y_i - \mu_2) \right\}^{k/3}$$
$$\le m + 2^{k/3} d_\omega H^*,$$

which implies (w.p.1):

$$\frac{N}{n^*} \le 1 + 2^{k/3}\sigma^{-k/3}H^*. \tag{5.1}$$

Now, by Wiener's (1939) ergodic theorem, it follows that $E[H^{*t}]$ is finite for all fixed positive numbers t. The right-hand side of (5.1) is also free from ω so that we can claim uniform integrability of all positive powers of $\frac{N}{n^*}$. Then, by appealing to the dominated convergence theorem and part (i), we claim part (ii).

Part (iii): Let us denote

$$S_{x,n}^* = \Sigma_{i=2}^n (n-i+1)(X_{n:i} - X_{n:i-1}) \quad \text{and} \quad S_{y,n}^* = \Sigma_{i=2}^n (n-i+1)(Y_{n:i} - Y_{n:i-1}),$$

so that we have $\hat{\sigma}_n = \frac{1}{2(n-1)}\left(S_{x,n}^* + \frac{1}{b}S_{y,n}^*\right)$. Let $L_1, L_2, \ldots$ and $M_1, M_2, \ldots$ be i.i.d. $exp(\sigma)$ and $exp(b\sigma)$ random variables respectively, Ls and Ms being independent, and denote $S_{x,n} = \Sigma_{i=1}^{n-1} L_i$,. $S_{y,n} = \Sigma_{i=1}^{n-1} M_i$.

Then, utilizing the embedding ideas from Lombard and Swanepoel (1978) and Swanepoel and van Wyk (1982), we can claim that the distribution of $\{S_{x,n}^*, S_{y,n}^*; n \le n_0\}$ is identical to that of $\{S_{x,n}, S_{y,n}; n \le n_0\}$ for all n_0. Thus, N given by (4.1) can be equivalently expressed as:

$$N \equiv \inf\left\{ n \ge m : n \ge d_\omega \left(\frac{1}{2(n-1)} \left(\Sigma_{i=1}^{n-1} L_i + \frac{\Sigma_{i=1}^{n-1} M_i}{b} \right) \right)^{k/3} \right\}$$
$$= \inf\left\{ n \ge m : \left(\frac{n}{n^*}\right)^{3/k} (n-1) \ge \Sigma_{i=1}^{n-1} Z_i \right\}, \tag{5.2}$$

where Z_is are i.i.d. gamma$(2, \frac{1}{2})$ random variables.

Now, N from (5.2) can be written as $R+1$ w.p.1 with R defined in (4.2). 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

$$P(R \le \frac{1}{2}n^*) = O(n^{*-\frac{3}{k}(m-1)}). \tag{5.3}$$

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

$$\left(\frac{n^*}{R}\right)^t \text{ is uniformly integrable.} \tag{5.4}$$

Now, we may write (w.p.1):

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

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

$$E\left[\left(\frac{n^*}{R}\right)^t I\left(R > \frac{1}{2}n^*\right)\right] = 1 + o(1). \tag{5.5}$$

Additionally, in view of (5.3), we note the following:

$$E\left[\left(\frac{n^*}{R}\right)^t I\left(R \le \frac{1}{2}n^*\right)\right] \le \left(\frac{n^*}{m-1}\right)^t P\left(R \le \frac{1}{2}n^*\right) = O\left(n^{*-\frac{3}{k}(m-1)+t}\right), \tag{5.6}$$

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

Combining (5.5)–(5.6), clearly (5.4) follows so that we can claim

$$E\left[\left(\frac{n^*}{R}\right)^t\right] = 1 + o(1) \textit{ if } m > 1 + \frac{1}{3}kt, \text{with } t > 0, \tag{5.7}$$

which is part (iii).

Part (*iv*): The proof is similar to that in Section 6.5 of Mukhopadhyay and Bapat (2016).

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).

5.2. *Outlines of proofs of Theorems 4.2–4.3*

Theorem 4.2 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 4.3, we recall that the associated expression for $E[\text{RPUC}_N]$ will resemble (2.8) with Q replaced by N from (4.1). Then, one will proceed with an expansion of $\omega^{-1}E[\text{RPUC}_N]$ similar to that in Section 6.4 of Mukhopadhyay and Bapat (2016) and exploit Theorem 4.2 with $t = -3, -4$. Additional details are kept out for brevity.

6. Data analysis: Simulations

After developing some interesting theoretical results associated with the two proposed estimation methodologies in Sections 3–4, we now move on to implement these strategies via computer simulations. We examine how these estimation strategies may perform when sample sizes are small (20) to moderate (50, 100, 150) to large (300, 500). All simulations are carried out with R codes based on 10,000 ($= H$, say) replications under each configuration and methodology.

Under each procedure, we generated pseudo random observations from the distribution (1.1) with $\mu_1 = 8, \sigma_1 = \sigma = 10$ and $\mu_2 = 5, \sigma_2 = b\sigma = 10b$ where we fixed b so that σ_2 became an appropriate multiple of σ. Moreover, we also fixed certain values of a, c, k, and n^*, thereby solving for a corresponding value of the risk-bound, ω. Thus, a set of preassigned values for $b, a, c, k, \omega, \sigma$ will have the associated n^* values as shown in column 1 of our Tables 2–4, 6, and 8–11.

In the context of the modified two-stage methodology (3.1)–(3.2), we fixed a positive lower bound σ_L for σ and a number m_0, thereby determining m from (3.1). On the other hand, we fixed a pilot sample size m in the context of the purely sequential methodology (4.1). While implementing a methodology to determine the final sample size (N) and a terminal estimator $(X_{N:1} - Y_{N:1})$ of $\mu_1 - \mu_2$, we pretended that we did not know μ_1, μ_2, σ, and n^* values.

Now, we specify a set of notation used in the tables. Under a fixed configuration with all necessary input (e.g., $a, c, k, \omega, m, m_0, \sigma_L$ as appropriate), we focus on implementing a particular estimation methodology. We ran the ith replication by beginning with m pilot observations and then eventually ended sampling by recording a final sample size $N = n_i$, terminal estimator $x_{n_i:1} - y_{n_i:1}$, and the achieved risk per unit cost:

$$\text{RPUC}_{n_i} = \omega \frac{n^{*3}}{a^2(1+b^2-b)} \left\{ \frac{1}{n_i}\left(1-\frac{a}{n_i}\right)^{-1}\left(1+\frac{ab}{n_i}\right)^{-1} - \left(\frac{a-ab}{n_i^2}\right) - \left(\frac{1}{n_i}\right)\right\} \quad (6.1)$$

$= r_i$, say,

$i = 1, \ldots, H(= 10{,}000)$.

$\overline{n} = H^{-1}\Sigma_{i=1}^{H} n_i$	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 $\overline{n}$;
$\overline{x}_{\min} = H^{-1}\Sigma_{i=1}^{H} x_{n_i:1}, \overline{y}_{\min} = H^{-1}\Sigma_{i=1}^{H} y_{n_i:1}$	Estimates of μ_1 and, μ_2 respectively;
$\overline{\Delta} = H^{-1}\Sigma_{i=1}^{H}(x_{n_i:1} - y_{n_i:1})$	Estimate of Δ, that is $\mu_1 - \mu_2$;
$s_{\overline{x}_{\min}} = \sqrt{(H^2-H)^{-1}\Sigma_{i=1}^{H}(x_{n_i:1} - \overline{x}_{\min})^2}$	Estimated standard error of $\overline{x}_{\min}$;
$s_{\overline{y}_{\min}} = \sqrt{(H^2-H)^{-1}\Sigma_{i=1}^{H}(y_{n_i:1} - \overline{y}_{\min})^2}$	Estimated standard error of $\overline{y}_{\min}$;
$s_{\overline{\Delta}} = \sqrt{(H^2-H)^{-1}\Sigma_{i=1}^{H}(x_{n_i:1} - y_{n_i:1} - \overline{\Delta})^2}$	Estimated standard error of $\overline{\Delta}$;
r_i	RPUC_{n_i} from (6.1);
$\overline{r} = H^{-1}\Sigma_{i=1}^{H} r_i$ with r_i from (6.1)	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 $\overline{z}$;

We now summarize the observed performances of the proposed estimation methodologies laid down in Sections 3–4. We obtained a large set of tables and results summarizing extensive simulations run under a variety of configurations. For brevity, we outline a small subset of our findings.

6.1. *Modified two-stage procedure (3.1)–(3.2)*

We now summarize performances for the modified two-stage estimation methodology (3.1)–(3.2) in Table 2 for

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

Table 2. Simulation results from 10,000 replications for the modified two-stage procedure (3.1)–(3.2) with $\mu_1 = 8, \mu_2 = 5, \sigma = 10, \sigma_L = 3, b = 1, a = 1, c = 0.1$.

n^*	ω	$\bar{x}_{\min}$ / $s_{\bar{x}_{\min}}$	$\bar{y}_{\min}$ / $s_{\bar{y}_{\min}}$	$\overline{\Delta}$ / $s_{\overline{\Delta}}$	$\bar{n}, (s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}, (s_{\bar{z}})$
				$k = 1, m_0 = 5$			
20	1.25×10^{-2}	8.4929 0.0049	5.4901 0.0049	3.0027 0.0069	20.38,(0.0132)	1.0194	0.9710,(0.0019)
50	8×10^{-4}	8.1973 0.0019	5.1974 0.0020	2.9998 0.0028	50.40,(0.0209)	1.0081	0.9865,(0.0012)
100	1×10^{-4}	8.0987 0.0010	5.1012 0.0010	2.9974 0.0014	100.45,(0.0297)	1.0045	0.9917,(0.0008)
300	3.7×10^{-6}	8.0334 0.0003	5.0336 0.0003	2.9998 0.0004	300.45,(0.0496)	1.0015	0.9971,(0.0004)
500	8×10^{-7}	8.0197 0.0001	5.0198 0.0001	2.9998 0.0002	500.43,(0.0645)	1.0008	0.9983,(0.0003)
				$k = 2, m_0 = 6$			
20	1.25×10^{-1}	8.5065 0.0050	5.5084 0.0051	2.9980 0.0071	20.40,(0.0334)	1.0202	1.1188,(0.0063)
50	8×10^{-3}	8.1998 0.0020	5.2001 0.0019	2.9996 0.0028	50.32,(0.0502)	1.0064	1.0424,(0.0032)
100	1×10^{-3}	8.1017 0.0010	5.0998 0.0010	3.0018 0.0014	100.42,(0.0719)	1.0042	1.0187,(0.0022)
300	3.7×10^{-5}	8.0335 0.0003	5.0336 0.0003	2.9998 0.0004	300.65,(0.1213)	1.0021	1.0033,(0.0012)
500	8×10^{-6}	8.0202 0.0002	5.0195 0.0001	3.0006 0.0002	500.20,(0.1560)	1.0004	1.0046,(0.0009)
				$k = 3, m_0 = 7$			
20	1.25	8.5184 0.0055	5.5333 0.0058	2.9850 0.0078	20.48,(0.0571)	1.0242	1.5479,(0.0181)
50	0.08	8.2056 0.0021	5.2066 0.0021	2.9990 0.0029	50.65,(0.0947)	1.0131	1.1970,(0.0079)
100	0.01	8.1045 0.0010	5.1024 0.0010	3.0020 0.0014	100.24,(0.1308)	1.0024	1.1020,(0.0046)
300	3.7×10^{-4}	8.0332 0.0003	5.0340 0.0003	2.9992 0.0004	300.46,(0.2243)	1.0015	1.0292,(0.0023)
500	8×10^{-5}	8.0201 0.0002	5.0198 0.0001	3.0002 0.0002	500.46,(0.2896)	1.0009	1.0175,(0.0017)

In this methodology, we need a positive and known lower bound $\sigma_L(= 3)$ for true σ, but σ remains unknown in practice. The pilot size m was determined from (3.1) but is not shown in Table 2. The estimation methodology (3.2) was implemented as described. Table 2 specifies $\mu_1, \mu_2, \sigma, \sigma_L, b, a, c$, 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), $\bar{y}_{\min}$, $s_{\bar{y}_{\min}}$ (column 4), $\overline{\Delta}(= \bar{x}_{\min} - \bar{y}_{\min})$, $s_{\overline{\Delta}}$ (column 5), values $\bar{n}$, $s_{\bar{n}}$ (column 6), the ratio $\bar{n}/n^*$ (column 7), and values $\bar{z}$, $s_{\bar{z}}$ (column 8).

All $\bar{x}_{\min}$ and $\bar{y}_{\min}$values appear close to $8(= \mu_1)$ and $5(= \mu_2)$, respectively, with very small estimated standard error values $s_{\bar{x}_{\min}}$, $s_{\bar{y}_{\min}}$for sample size 100 or over. We observe that $\overline{\Delta}$ accurately estimates the true difference Δ in locations, which is 3. For all (k, m_0) values, we see that $\bar{n}$ estimates n^* very accurately across the board. These features are consistent with Theorem 3.1, parts (i)–(ii).

The last column shows that the modified two-stage estimation methodology (3.1)–(3.2) is very successful in the case of all selected values of k for all n^* under consideration in delivering a risk-bound approximately preset ω. More specifically, in the case $k = 1$, we note that $\bar{z} < 1$;

that is, $\bar{r} < \omega$, for all chosen n^* values, which is remarkable. On the other hand, $\bar{z}$ values hang around 1 for all chosen n^* and k values, which is certainly encouraging.

The entries in Table 3 are similar to those in Table 2. Table 3 used a different lower bound $\sigma_L(= 5)$ for true σ. We highlight performances for $n^* = 20, 150, 500$. Again, the pilot size m was determined from (3.1) but m is not shown in Table 3. There are two sets of two blocks corresponding to different values of $b(= 2, 3)$ and $k(= 1, 2)$. We find that $\overline{n}/n^*$ is nearer to 1 in Table 3 compared to those in Table 2. Also, entries found in the last column of Table 3 look more attractive to those in Table 2. This feature should be expected since the specified positive and known lower bound $\sigma_L = 5$ is closer to $\sigma = 10$ than $\sigma_L = 3$ is.

In Table 4, we provide the values of ψ found in (3.5) corresponding to the configurations highlighted in Table 2. We expect that $E(N-n^*)$ values should lie inside the interval $[\psi, \psi+1]$ for large n^* values in view of Theorem 3.2. This is indeed true as we find that nearly all $\bar{n} - n^*$ values lie inside the corresponding interval $[\psi, \psi + 1]$. Mukhopadhyay and Bapat (2016) briefly mentioned the case when σ_L may be slightly misspecified.

We provide Figure 2 showing empirical validation of asymptotic normality result described in (3.7). We considered two scenarios, namely, $\sigma_L = 3, b = 1, k = 1, n^* = 100$ (Figure 2(a)) and $\sigma_L = 3, b = 2, k = 3, n^* = 500$ (Figure 2(b)). Under each configuration, we recorded observed values:

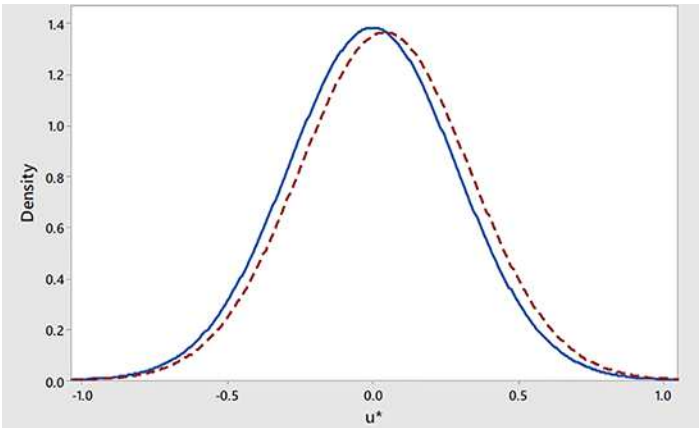
$$N = n_i, i = 1, \ldots, H(= 10{,}000)$$

Table 3. Simulation results from 10,000 replications for the modified two stage procedure (3.1)–(3.2) with $\mu_1 = 8, \mu_2 = 5, \sigma = 10, \sigma_L = 5, a = 1, c = 0.1$.

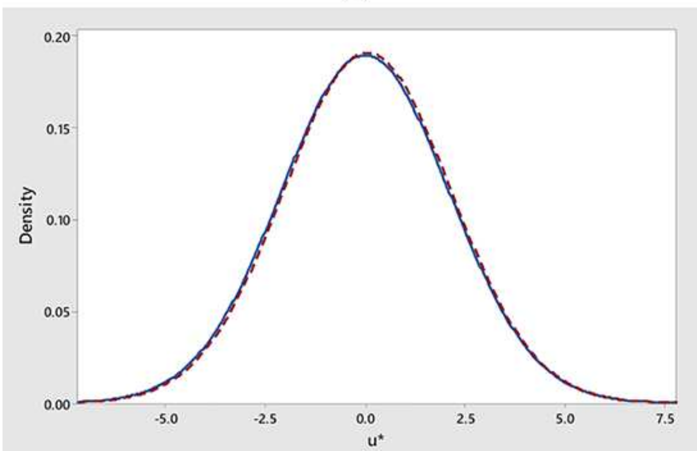
n^*	ω	$\bar{x}_{min}$ $s_{\bar{x}_{min}}$	$\bar{y}_{min}$ $s_{\bar{y}_{min}}$	$\overline{\Delta}$ $s_{\overline{\Delta}}$	$\bar{n}, (s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}, (s_{\bar{z}})$
			$k = 1, m_0 = 5, b = 2$				
20	3.75×10^{-2}	8.9922 0.0049	5.4933 0.0100	3.4989 0.0112	20.39,(0.0143)	1.0197	0.8992,(0.0019)
150	8.88×10^{-5}	8.1316 0.0006	5.0656 0.0013	3.0659 0.0014	150.46,(0.0388)	1.0031	0.9838,(0.0007)
500	2.4×10^{-6}	8.0402 0.0001	5.0199 0.0003	3.0202 0.0004	500.50,(0.0713)	1.0010	0.9948,(0.0004)
			$k = 2, m_0 = 6, b = 2$				
20	3.75×10^{-1}	9.0348 0.0052	5.5075 0.0105	3.5273 0.0015	20.31,(0.0332)	1.0158	1.0370,(0.0057)
150	8.88×10^{-4}	8.1332 0.0006	5.0666 0.0013	3.0665 0.0014	150.41,(0.0883)	1.0027	1.0014,(0.0017)
500	2.4×10^{-5}	8.0403 0.0001	5.0195 0.0002	3.0207 0.0004	500.44,(0.1601	1.0009	1.0001,(0.0009)
			$k = 1, m_0 = 5, b = 3$				
20	8.75×10^{-2}	9.1832 0.0050	5.5031 0.0149	3.6801 0.0156	20.39,(0.0168)	1.0195	0.9617,(0.0021)
150	2.07×10^{-4}	8.1988 0.0006	5.0673 0.0019	3.1315 0.0020	150.32,(0.0451)	1.0021	0.9801,(0.0008)
500	8×10^{-5}	8.0594 0.0002	5.0198 0.0005	3.0396 0.0005	500.31,(0.0816)	1.0006	0.9940,(0.0004)
			$k = 2, m_0 = 6, b = 3$				
20	8.75×10^{-1}	9.1187 0.0052	5.5093 0.0161	3.6094 0.0167	20.33,(0.0385)	1.0168	1.0309,(0.0067)
150	2.07×10^{-3}	8.2026 0.0006	5.0661 0.0020	3.1364 0.0021	150.60,(0.1004)	1.0040	0.9957,(0.0020)
500	5.6×10^{-5}	8.0597 0.0001	5.0198 0.0004	3.0399 0.0004	500.23,(0.1852)	1.0004	1.0011,(0.0011)

Table 4. Values of $\bar{n} - n^*$, ψ from (3.5) and $\psi + 1$ for each k used in Table 2.

$n^*\backslash k$	1	2	3
20	0.38	0.40	0.48
50	0.40	0.32	0.65
100	0.45	0.42	0.24
300	0.45	0.65	0.46
500	0.43	0.20	0.46
ψ	−0.0829	−0.1239	0
$\psi + 1$	0.9170	0.8760	1



(a)



(b)

Figure 2. Plots of normality curves for the modified two-stage procedure (3.1)–(3.2) as validations for (3.7). 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 $\sigma_0^2 = \frac{1}{18}(\frac{\sigma}{\sigma_L})^{k/3}k^2$ coming from (3.5): (a) $\sigma_L = 3, b = 1, k = 1, n^* = 100$; (b) $\sigma_L = 3, b = 2, k = 3, n^* = 500$.

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 expected theoretical $N(0, \sigma_0^2)$ distributions (solid curve in blue) where $\sigma_0^2 = \frac{1}{18}(\frac{\sigma}{\sigma_L})^{k/3}k^2$ found from (3.5).

Both sets of plots in Figures 2(a) and 2(b) show very good fit. However, Figure 2(a) shows that the curves are slightly off from one another, but this corresponds to a moderate value of $n^*(= 100)$, Indeed, Figure 2(b) clearly shows that the empirical and theoretical distribution curves nearly lie on each other when $n^*(= 500)$ is large.

To supplement the graphical presentations in Figure 2, we also performed the customary Kolmogorov-Smirnov (K-S) test for normality in the case of each data set that generated Figures 2(a)–2(b). Table 5 shows associated K-S test statistic (D) values under the null hypothesis of normality with associated p-values.

Both p-values (Table 5) are much larger than 0.05. Thus, our earlier thoughts supported by simple visual examinations of Figures 2(a)–2(b) are clearly validated by 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.

6.2. *Purely sequential procedure (4.1)*

We now summarize performances for the purely sequential estimation methodology (4.1) in Table 6 for

$$n^* = 20, 50, 100, 300, 500 \text{ and}$$
$$(k, m) = (1, 4), (2, 6), (3, 9).$$

The estimation methodology (4.1) was implemented as described. Table 6 specifies $\mu_1, \mu_2, \sigma, b, a, c$ 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), $\overline{x}_{\min}$, $s_{\overline{x}_{\min}}$ (column 4), $\overline{\Delta}(= \overline{x}_{\min} - \overline{y}_{\min})$, $s_{\overline{\Delta}}$ (column 5) values $\overline{n}$, $s_{\overline{n}}$ (column 6), the ratio $\overline{n}/n^*$ (column 7), and values $\overline{z}$, $s_{\overline{z}}$ (column 8). An explanation of column 9 comes later.

All $\overline{x}_{\min}$ and $\overline{y}_{\min}$ values appear closer to $8(= \mu_1)$ and $5(= \mu_2)$ respectively with very small estimated standard error values $s_{\overline{x}_{\min}}$, $s_{\overline{y}_{\min}}$ for sample size 100 or over. $\overline{\Delta}$ also accurately estimates the actual difference in locations which is 3. For all (k, m) values, it appears that $\overline{n}$ estimates n^* very accurately throughout. These features are consistent with Theorem 4.1, parts (i)–(ii). Column 8 shows that the purely sequential estimation methodology (4.1) is very successful for all values of k and n^* under consideration in delivering a risk-bound approximately ω. However, when $k = 1, 2$, we see an underestimation in the risk, which improves when $k = 3$. We feel that the purely sequential procedure provides estimates for the risk per unit cost, which are generally closer to 1 than the two-stage methodology (3.1)–(3.2).

Theorem 4.3 showed that $\omega^{-1}E[\text{RPUC}_N]$ should be close to $1 + \varepsilon$ where we have

$$\varepsilon = (6p - 3\eta_k - 3 + a\left(\frac{b^2 - b^3 - b + 1}{1 + b^2 - b}\right))n^{*-1} \tag{6.2}$$

Table 5. Kolmogorov-Smirnov test results corresponding to Figures 2(a)–2(b).

Case	Parameter configuration	K-S stat D	p-value
2a	$\sigma_L = 3, b = 1, k = 1, n^* = 100$	0.46252	0.8212
2b	$\sigma_L = 3, b = 2, k = 4, n^* = 500$	0.49874	0.9025

Table 6. Simulation results from 10,000 replications of the purely sequential procedure (4.1) with $\mu_1 = 8$, $\mu_2 = 5, \sigma = 10, b = 1, a = 1, c = 0.1$

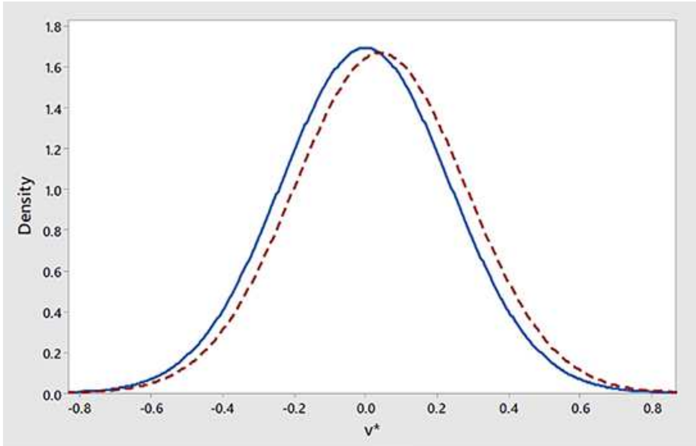
n^*	ω	$\bar{x}_{\min}$ $s_{\bar{x}_{\min}}$	$\bar{y}_{\min}$ $s_{\bar{y}_{\min}}$	$\overline{\Delta}$ $s_{\overline{\Delta}}$	$\bar{n}, (s_{\bar{n}})$	$\bar{n}/n^*$	$\bar{z}, (s_{\bar{z}})$	ε in (7.2)
				k = 1, m = 4				
20	1.25×10^{-2}	8.4902 0.0049	5.4975 0.0049	2.9926 0.0069	20.40,(0.0110)	1.0201	0.8612,(0.0016)	−0.1203
50	8×10^{-4}	8.1989 0.0020	5.1960 0.0019	3.0028 0.0027	50.44,(0.0168)	1.0088	0.9810,(0.0009)	−0.0482
100	1×10^{-4}	8.0990 0.0010	5.1005 0.0009	2.9986 0.0014	100.40,(0.0238)	1.0040	0.9914,(0.0007)	−0.0241
300	3.7×10^{-6}	8.0335 0.0003	5.0336 0.0003	2.9999 0.0004	300.31,(0.0414)	1.0015	0.9979,(0.0004)	−0.0080
500	8×10^{-7}	8.0197 0.0001	5.0202 0.0002	2.9994 0.0002	500.45,(0.0527)	1.0009	0.9979,(0.0003)	−0.0048
				k = 2, m = 6				
20	1.25×10^{-1}	8.4937 0.0050	5.5018 0.0051	2.9918 0.0071	20.30,(0.0218)	1.0150	0.9610,(0.0040)	−0.0555
50	8×10^{-3}	8.2003 0.0020	5.2037 0.0020	2.9966 0.0028	50.34,(0.0336)	1.0069	0.9847,(0.0021)	−0.0222
100	1×10^{-3}	8.0990 0.0009	5.0994 0.0010	2.9995 0.0014	100.37,(0.0472)	1.0037	0.9965,(0.0014)	−0.0111
300	3.7×10^{-5}	8.0332 0.0003	5.0333 0.0003	2.9998 0.0004	300.38,(0.0822)	1.0012	0.9980,(0.0008)	−0.0037
500	8×10^{-6}	8.0197 0.0001	5.0203 0.0001	2.9994 0.0002	500.44,(0.1046)	1.0008	0.9986,(0.0005)	−0.0022
				k = 3, m = 9				
20	1.25	8.5199 0.0054	5.5157 0.0053	3.0041 0.0075	20.14,(0.0329)	1.0074	1.2011,(0.0091)	0.0492
50	8×10^{-2}	8.2035 0.0020	5.1993 0.0019	3.0042 0.0028	50.21,(0.0506)	1.0042	1.0542,(0.0035)	0.0197
100	1×10^{-2}	8.1014 0.0010	5.1012 0.0010	3.0002 0.0014	100.23,(0.0706)	1.0023	1.0242,(0.0022)	0.0098
300	3.7×10^{-4}	8.0335 0.0003	5.0333 0.0003	3.0002 0.0004	300.19,(0.1236)	1.0006	1.0083,(0.0012)	0.0032
500	8×10^{-5}	8.0200 0.0001	5.0199 0.0001	3.0000 0.0002	500.04,(0.1594)	1.0000	1.0058,(0.0009)	0.0019

from (4.7) when n^* is large. η_k was defined by (4.5) and it was tabulated in Table 1. Column 9 in Table 6 shows these ε values under each configuration. Upon comparing $\bar{z}$ values from column 8 with ε values from column 9, we see that for all practical purposes, the $\bar{z}$ values are explained fairly well by $1 + \varepsilon$ for practical purposes.

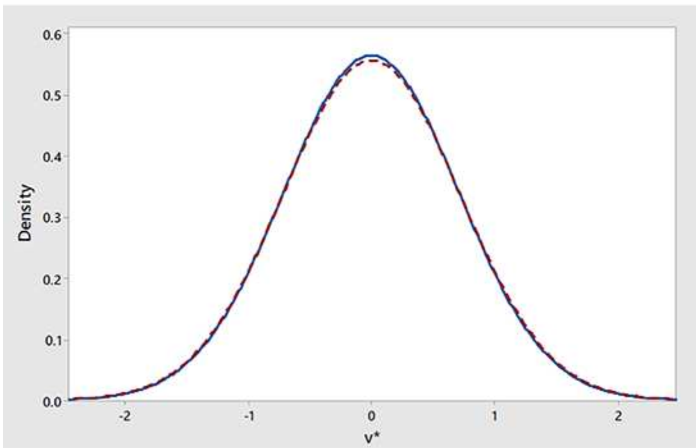
We now address Figure 3 consisting of two side-by-side plots in our attempt to validate empirically the normality result described in Theorem 4.1, part (v). We considered two scenarios, $b = 3, k = 1, m = 5,\ n^* = 100$ in Figure 3(a) and $b = 1, k = 3, m = 13,\ n^* = 500$ in Figure 3(b). Under each configuration, we recorded observed values:

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

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 $\sigma_1^2 = \frac{1}{18}k^2$ coming from Theorem 4.1, part (v).



(a)



(b)

Figure 3. Plots of normality curves for the purely sequential procedure (4.1) as validation of Theorem 4.1, 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 $\sigma_1^2 = \frac{1}{18}k^2$ coming from Theorem 4.1, part (v): (a) $b = 3, k = 1, m = 5, n^* = 100$; (b) $b = 1, k = 3, m = 13, n^* = 500$.

Both sets of plots in Figures 2(a) and 2(b) show very good fit. However, Figure 2(a) shows that the curves are slightly off from one another, but this corresponds to a moderate value of $n^*(= 100)$, Indeed, Figure 2(b) clearly shows that the empirical and theoretical distribution curves nearly lie on each other when $n^*(= 500)$ is large.

To supplement the graphical presentations in Figure 3, we also performed the customary K-S test for normality in the case of each dataset that generated Figures 3(a)–3(b). Table 7 shows associated K-S test statistic (D) values under the null hypothesis of normality with associated p-values.

Both p-values (Table 7) are much larger than 0.05 and our earlier thoughts supported by visual examinations of Figures 2(a)–2(b) are clearly validated by K-S test of normality under

Table 7. Kolmogorov-Smirnov test results corresponding to Figures 3(a)–3(b).

Case	Parameter configuration	K-S stat D	p-value
3a	$b = 3, m = 5, k = 1, n^* = 100$	0.55487	0.8978
3b	$b = 1, m = 13, k = 3, n^* = 500$	0.52314	0.9487

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.

7. Data analysis: Illustrations using real data

In this section, we illustrate applications of the modified two-stage estimation methodology (3.1)–(3.2) and the purely sequential estimation methodology (4.1) using two real data sets, one from cancer research and the other from reliability. The first example (Section 7.1) uses parallel data sets on survival times of cancer patients from Shanker et al. (2016). The data were utilized in Efron's (1988) paper.

A second example (Section 7.2) uses lifetimes of steel components data available from the textbook by Lawless (2003; *Statistical Models and Methods for Lifetime Data*). These data were presented earlier by Crowder (2000).

7.1. Cancer studies data

These data consist of two independent data sets representing the survival times of a group of patients suffering from head and neck cancer. The first group was treated using a combination of radiotherapy and chemotherapy whereas the second group was treated with radiotherapy, alone. These data on survival times, X from the first group and Y from the second group, have been presented in Shanker et al. (2016) and were reported earlier in Efron (1988).

The full data consisted of survival times of 44 and 51 patients, respectively. As a check for a negative exponential model's fit, in Figure 4 we provide the exponential Q-Q plots for both data sets. Indeed, the data points seem to lie fairly well within or near the bands, indicating a good fit.

Treating these two data sets as the universe, we first found $\hat{\mu}_1 = 12.2$, $\hat{\sigma}_1 = 216.19$ and $\hat{\mu}_2 = 6.53$, $\hat{\sigma}_2 = 233.88$ from full data. We note that the scale parameters appear nearly the same. Thus, we assumed $b = 1$. Utilizing (2.1), we found the pooled estimator of the scale, $\hat{\sigma} = 225.04$. We then implemented both modified two-stage and purely sequential estimation procedures drawing observations (X, Y) from the full set of data as needed. It is emphasized, however, that implementation of sampling strategies did not exploit the observed numbers $\hat{\mu}_1$, $\hat{\mu}_2$, or $\hat{\sigma}$.

For estimating $\Delta \equiv \mu_1 - \mu_2$, we carried out a single run under both procedures. Tables 8–9 provide the results from implementing the stopping rules from (3.1)–(3.2) and (4.1) respectively corresponding to certain fixed values of c, k, a and the preset risk-bound ω, chosen arbitrarily. We fixed m or m_0 as needed.

The outcome of these methodologies are summarized in Tables 8–9. Table 8 summarizes results from the modified two-stage procedure with $\sigma_L = 80$ but assuming otherwise that the scale parameter σ remains unknown. Table 9 summarizes the results from the purely sequential procedure.

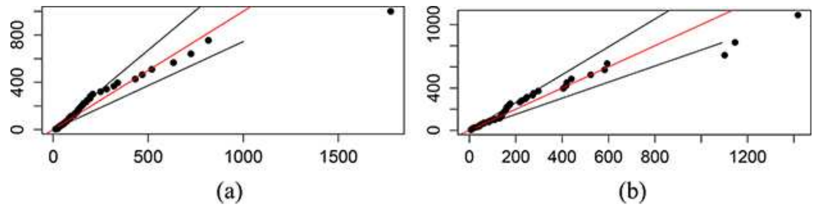


Figure 4. Exponential Q-Q plots for cancer data: (a) radiotherapy and chemo (X) and (b) radiotherapy alone (Y).

Table 8. Analysis of cancer survival data using modified two-stage procedure (3.1)–(3.2) with $b = 1, a = 1, c = 0.1, \sigma_L = 80$.

n^*	m_0	k	ω	$\hat{\mu}_1$: $x_{n:1}$	$\hat{\mu}_2$: $y_{n:1}$	$\widehat{\Delta}$	n	n/n^*	z
15	3	1	0.6667	11.64	6.84	4.8	16	1.067	0.9274
25	3		0.1440	12.52	6.21	6.31	24	0.96	0.9526
15	4	2	150.04	12.03	6.47	5.56	15	1.00	1.4201
25	4		32.410	12.38	6.64	5.74	26	1.04	1.2317

Table 9. Analysis of cancer survival data using purely sequential procedure (4.1) with $b = 1, a = 1, c = 0.1$.

n^*	m	k	ω	$\hat{\mu}_1$: $x_{n:1}$	$\hat{\mu}_2$: $y_{n:1}$	$\widehat{\Delta}$	n	n/n^*	z	ε in (7.2)
15	5	1	0.66677	12.40	6.52	5.88	17	1.13	0.8629	−0.1609
25	5		0.14402	12.27	6.19	6.08	26	1.04	0.8897	−0.0965
15	7	2	150.046	13.02	6.49	6.53	15	1.00	0.9248	−0.0740
25	7		32.4101	13.11	6.62	6.49	28	1.12	0.9589	−0.0444

Under both methodologies, we notice that the terminal estimated values of μ_1, μ_2 are not too far away from corresponding $\hat{\mu}_1 = 12.2$ and $\hat{\mu}_2 = 6.53$ obtained from full data. In addition, the estimate of the difference in locations given by $\widehat{\Delta}$ is also close to the actual value $\Delta = 5.67$. One other comment is in order: The n^* values shown in the first column are computed using (2.7) after replacing σ with $\hat{\sigma} = 225.04$ obtained from full data. Again, in running the estimation methodologies, we did not exploit the number $\hat{\sigma} = 225.04$.

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 is nice to see. However, these correspond to a single run each and hence observed n may not always be very close to n^*. The values of z—that is, the ratio of achieved risk per unit cost and preset goal ω—also appear reasonably under (or close to) 1.

7.2. Lifetimes of steel specimen data

We considered Data G4 presented in Appendix G of the textbook *Statistical Models and Methods for Lifetime Data*, by Lawless (2003). These data, presented earlier in Crowder (2000), consisted of lifetimes of steel specimens tested at 14 stress levels. We divided the whole data into two parts, one corresponding to stress levels 35 or lower (low stress) and the other corresponding to stress levels 35.5 or higher (high stress).

We then considered these two groups to be our two independent groups, low (high) stress group giving rise to failure time X (Y). The two data sets consisted of 140 observations each.

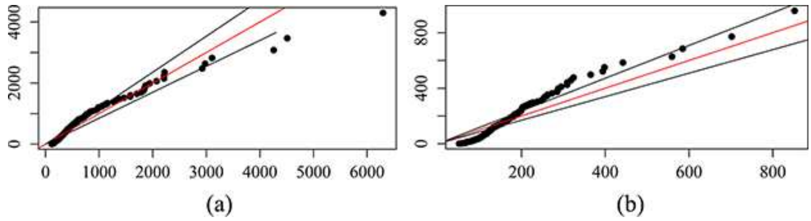


Figure 5. Exponential Q-Q plots for stress data (a) low stress (X) and (b) high stress (Y).

As a check for a negative exponential (1.1) fit, Figure 5 provides the exponential Q-Q plots for both groups. Indeed, the data seem to lie largely within or near the bands, indicating a reasonable fit.

The scale parameters are taken to be σ and $b\sigma$ for X, Y, respectively, whereas the formulation demands that one should specify a value of b before obtaining the data, perhaps from previous analogous studies or prior experiences. We have thus resorted to the following practical approach.

We randomly picked 10 data points X (Y) from each group under low (high) stress levels and treat them as our data, say, from previous analogous study. In the population from where this prior data supposedly became available through a similar previous stress test, suppose that the low (high) stress level group had the scale parameter value σ_{Low} (σ_{High}). Following were those observed data values:

$$\begin{aligned} &\textbf{Low}\,(X) : 206, 273, 312, 385, 415, 568, 669, 714, 767, 1056 \\ &\Rightarrow \widehat{\sigma}_{\text{Low}} = \frac{1}{9}\sum_{i=1}^{n}(X_i - X_{10:1}) = \frac{1}{9}(5365 - 2060) = 367.22; \\ &\textbf{High}(Y) : 66, 90, 105, 108, 121, 122, 127, 164, 255, 318 \\ &\Rightarrow \widehat{\sigma}_{\text{High}} = \frac{1}{9}\sum_{i=1}^{n}(Y_i - Y_{10:1}) = \frac{1}{9}(1476 - 660) = 90.67. \end{aligned} \tag{7.1}$$

We have

$$\begin{aligned} &F_{18,18;0.025} = 0.38527 \text{ and } F_{18,18;0.975} = 2.5956 \\ &\Rightarrow 95\% \textit{ confidence interval for } \sigma_{\text{High}}/\sigma_{\text{Low}} \text{ is given by} \\ &\left(\frac{90.67}{367.22} \times \frac{1}{2.5959}, \frac{90.67}{367.22} \times \frac{1}{0.38527}\right);\ \text{that is, } (0.0951, 0.6409). \end{aligned}$$

Thus, a null hypothesis postulating that $\sigma_{\text{High}}/\sigma_{\text{Low}} = 1$ will be rejected in favor of a two-sided test at 5% level. Indeed, any number lying between 0.0951 and 0.6409 would seem reasonable for the ratio $\sigma_{\text{High}}/\sigma_{\text{Low}}$.

Thus, we decided to implement our proposed methodologies on the remainder of the stress test data after removing the observations shown in (7.1) from the full data set with two possible choices for the number b, namely, Case 1: $b = 0.33$ and Case 2: $b = 0.25$.

We treat the remaining 130 values from each stress groups as our universe and carry out our proposed methodologies. From full data sets, we found $\hat{\mu}_1 = 115$, $\hat{\mu}_2 = 51$ and $\widehat{\sigma} = 507.22$ when $b = 0.33$. whereas $\widehat{\sigma} = 570.78$ when $b = 0.25$. Table 10 shows findings from implementing the stopping rule (3.1)–(3.2) with $\sigma_L = 150$ whether $b = 0.33$ or $b = 0.25$.

Table 11 shows findings from implementing the stopping rule (4.1) when $b = 0.33, 0.25$. Both tables correspond to certain fixed choices of c, k, a and the preset (4.1) risk-bound ω and m or m_0 as needed. The implementation of sampling strategies did not exploit the numbers $\hat{\mu}_1$, $\hat{\mu}_2$ or $\hat{\sigma}$ obtained from full data sets. For estimating $\Delta = \mu_1 - \mu_2$, we carried out a single run under both methodologies.

Here again, we show n^* values just so that one is able to gauge whether the observed n-values look reasonable. The estimate of difference Δ in locations given by $\overline{\Delta}$ appears fairly close to the actual value of Δ under either choices of b. The ratio n/n^* appears reasonably close to 1 which should be desirable. The value z—that is, the ratio of achieved risk per unit cost and preset goal ω—appears reasonably under (or close to) 1.

7.3. Concluding thoughts

In Tables 8–11, we have shown n^* values just so that one is able to gauge whether the observed n-values look reasonable. The estimate of difference Δ in locations given by $\overline{\Delta}$ appears fairly close to the actual value of Δ under either choice(s) of b that we had made in either problem. The ratio n/n^* appears reasonably close to 1 which should be desirable. The value z—that

Table 10. Analysis of lifetimes of steel components in stress data using modified two-stage procedure (3.1)–(3.2) with $a = 1, c = 0.1$.

n^*	m_0	k	ω	$\hat{\mu}_1$: $x_{n:1}$	$\hat{\mu}_2$: $y_{n:1}$	$\widehat{\Delta}$	n	n/n^*	z
				$b = 0.33$, $\sigma_L = 150$					
50	4	1	0.0316	129	51	78	52	1.04	0.9056
80	4		0.0077	115	57	58	79	0.98	1.0511
50	5	2	16.0311	129	65	64	47	0.94	1.2289
80	5		3.9138	146	51	95	73	0.91	1.3335
				$b = 0.25$, $\sigma_L = 150$					
50	4	1	0.0371	115	57	58	48	0.96	0.9608
80	4		0.0090	115	57	58	73	0.91	1.1816
50	5	2	21.1763	129	51	78	48	0.96	1.1538
80	5		5.1700	129	57	50	72	0.90	1.3340

Table 11. Analysis of lifetimes of steel components in stress data using purely sequential procedure (4.1) with $a = 1, c = 0.1$.

n^*	m	k	ω	$\hat{\mu}_1$: $x_{n:1}$	$\hat{\mu}_2$: $y_{n:1}$	$\widehat{\Delta}$	n	n/n^*	z	ε in (6.2)
					$b = 0.33$					
50	5	1	0.0316	115	57	58	51	1.02	0.9603	−0.0101
80	5		0.0077	140	51	89	78	0.97	0.9796	−0.0063
50	7	2	16.0311	129	51	78	49	0.98	1.0568	0.0159
80	7		3.9138	143	57	86	79	0.98	1.0233	0.0099
					$b = 0.25$					
50	5	1	0.0371	129	51	78	51	1.02	0.9824	−0.0090
80	5		0.0090	115	59	56	79	0.98	0.9912	−0.0056
50	7	2	21.1763	146	51	95	49	0.98	1.0478	0.0170
80	7		5.1700	129	51	78	81	1.01	1.0321	0.0106

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

One may notice that the purely sequential strategy appears to perform better overall than the modified two-stage strategy. However, it is also true 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. A practitioner, however, may consider employing one of the two procedures (3.1)–(3.2) or (4.1) that will provide an acceptable level of logistical comfort in running an experiment as one balances benefit under the presence of additional factors, namely, feasibility, efficiency, accuracy, operational convenience, and cost.

Acknowledgments

We are grateful to Professor Jerry Lawless for enthusiastically giving us the permission to use the reliability and stress test Data G4 (Appendix G) from his 2003 textbook, *Statistical Models and Methods for Lifetime Data*. We heartily thank Professor Lawless for this. We also remain indebted to the Associate Editor and the referees for their constructive comments.

References

Aoshima, M., Mukhopadhyay, N., and Kobayashi, Y. (2011). Two-Stage Procedures for Estimating the Difference of Means When the Sampling Cost is Different, *Sequential Analysis* 30: 160–171.

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

Basu, A. P. (1971). On a Sequential Rule for Estimating the Location Parameter of an Exponential Distribution, *Naval Research Logistics Quarterly* 18: 329–337.

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

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.

Crowder, M. J. (2000). Tests for a Family of Survival Models Based on Extremes, in *Recent Advances in Reliability Theory,* N. Limnios and N. Nikulin, eds., pp. 307–321, Boston: Birkhauser.

Efron, B. (1988). Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve, *Journal of American Statistical Association* 83: 414–425.

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.

Ghurye, S. G. (1958). Note on Sufficient Statistics and Two Stage Procedures, *Annals of Mathematical Statistics* 29: 155–166.

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

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.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*, New York: Wiley.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data,* second edition, Hoboken, NJ: Wiley.

Lawless, J. F. and Singhal, K. (1980). Analysis of Data from Life Test Experiments Under an Exponential Model, *Naval Research Logistics Quarterly* 27: 323–334.

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. (1980). A Consistent and Asymptotically Efficient Two-Stage Procedure to Construct Fixed-Width Confidence Intervals for the Mean, *Metrika* 27: 281–284.

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. and Bapat, S. R. (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: 175–206. http://dx.doi.org/10.1080/07474946.2016.1165532.

Mukhopadhyay, N. and Darmanto, S. (1988). Sequential Estimation of the Difference of Means of Two Negative Exponential Populations, *Sequential Analysis* 7: 165–190

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

Mukhopadhyay, N., de Silva, B. M. and Waikar, V. B. (2010). On Two Stage Confidence Interval Procedures and Their Comparisons for Estimating the Difference of Normal Means, *Sequential Analysis* 29: 44–68.

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 Hamdy, H. I. (1984). On Estimating the Difference of Location Parameters of Two Negative Exponential Distributions, *Canadian Journal of Statistics* 12: 67–76

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

Shanker, R., Fesshaye, H., and Selvaraj, S. (2016). On Modeling of Lifetime Data Using One Parameter Akash, Lindley and Exponential Distributions, *Biometrics and Biostatistics International Journal* 2: 1–10.

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.

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 Arnold Zellner, eds., pp. 195–208, North Holland, Amsterdam: 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.

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.