UNIVERSITY OF Southampton
Southampton Statistical Sciences Research Institute

Working Paper M09/06

Methodology

Sensitivity Of Normal-Based Triple Sampling Sequential Point Estimation To The Normality Assumption

A.S. Yousef, A.C. Kimber, H.I. Hamdy

Abstract

This article discusses the sensitivity of the sequential normal-based triple sampling procedure for estimating the population mean to departures from normality. We assume only that the underlying population has finite but unknown first four moments and find that asymptotically the behaviour of the estimator and the sample size depend on both the skewness and kurtosis of the underlying distribution, when using a squared error loss function with linear sampling cost. We supplement our findings with a simulation experiment to study the performance of the estimator and the sample size in a range of conditions.

Sensitivity of Normal-based Triple Sampling Sequential Point Estimation to the Normality Assumption

A. S. Yousef [1,2] , A. C. Kimber[1], H. I. Hamdy[2]

University of Southampton[1]and Kuwait University[2]

Abstract

This article discusses the sensitivity of the sequential normal-based triple sampling procedure for estimating the population mean to departures from normality. We assume only that the underlying population has finite but unknown first four moments and find that asymptotically the behaviour of the estimator and the sample size depend on both the skewness and kurtosis of the underlying distribution, when using a squared error loss function with linear sampling cost. We supplement our findings with a simulation experiment to study the performance of the estimator and the sample size in a range of conditions.

Keywords: Asymptotic relative efficiency, kurtosis, regret, sampling cost, simulation, skewness, squared error loss function, Taylor expansion.

Corresponding author:

A.C.Kimber**,** School of Mathematics and Southampton Statistical Sciences Research Institute, University of Southampton, SO17 1BJ, U.K.

Email A.C.Kimber@soton.ac.uk

Tel +44 (0)23 8059 5124

Fax +44 (0)23 8059 5147

1. Introduction

Hall (1981) introduced the triple sampling sequential technique to achieve the operational savings made possible by sampling in batches and at the same time to guarantee the asymptotic efficiency of the one by one sequential sampling of Anscombe (1953), Robbins (1959), and Chow and Robbins (1965). Although Hall's techniques were specifically devised to construct fixed width confidence for the normal mean with a prescribed coverage, his results may be modified to treat point estimation problems. Mukhophadhyay (1985) used Hall's (1981) results to derive a triple sampling sequential point estimation technique for the normal mean with bounded cost function. Since the publication of the Mukhophadhyay (1985) paper, attention has been given to the application of triple sampling sequential point estimation in two main directions. The first of these involves the assumption of normality of the underlying population and uses higher order cost functions. Here asymptotic higher order moments (positive or negative) are used to evaluate the performance of both the sampling scheme and inferences about the population. The second approach is to develop triple sampling schemes where the underlying population has a continuous distribution other than normal. Key references are Mukhophadhyay et. al. (1987), Hamdy and Pallotta (1987), Hamdy (1988) and Hamdy et. al. (1989); see also Johnson et. al. (1994, pp, 588, 623) for details. However, none of these papers discussed the sensitivity of the normal-based triple sampling sequential results to departures from the normality assumption. It is of interest to investigate this issue. If the procedure is insensitive to such departures, then the normal-based procedure may be used more generally, which will be useful in practical applications. On the other hand, if the procedure is sensitive to departures from normality, it is of interest to understand and quantify this sensitivity.

Early work on sensitivity to assumptions in sequential inference was conducted by Bhattacharjee and Nagendra (1964) in the context of the Wald sequential test for the mean, and by Bhattacharjee (1965), Blumenthal and Govindarajulu (1977) and Ramkaran (1983) for Stein's two stage sampling procedure. Jureckova and Sen (1996) devoted several chapters to the sensitivity of point and interval estimation methods in sequential statistical inference.

In the current study the sensitivity of normal-based triple sampling procedures to non-normality of the underlying distribution will be considered. The problem arises when the underlying population is misidentified (i.e., normality is assumed when in fact the population is not normal). Also, for some distributions (such as the gamma, beta, t and uniform distributions) explicit expression of the optimal sample size is intractable and therefore, if the normal-based triple sampling procedure were robust to departure from normality, it would be of practical help to use the triple sampling procedure with normal stopping rule.

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed random variables from a continuous distribution function $F(.)$ with mean μ, variance θ, skewness γ and kurtosis β, all unknown but finite. The main focus in this study is the estimation of μ in the presence of the nuisance parameter θ or some continuously differentiable and bounded function $g(\theta)$.

In the literature on sequential sampling for inference about the mean and for most distributions, it is assumed that the sample size required to satisfy the above conditions takes the general form

$$n \geq n^{*} = \lambda g(\theta), \qquad (1.1)$$

where λ is a function of some predetermined constants A and C, which may appear in a cost function incurred in point estimation of μ. Moreover, the function λ is also permitted to approach infinity such that the optimal sample size $n^{*} \to \infty$. Having observed a random sample $X_1, X_2, ..., X_n$ from the distribution $F(.)$ with $n \geq 2$, we propose to use the sample mean $\bar{X}_n = \sum_{i=1}^{n} X_i / n$ and the sample variance $S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 / (n-1)$ as estimators of μ and θ respectively.

The triple sampling scheme is introduced in section 2. The main asymptotic results are derived in section 3. The asymptotic distribution of the stopping sample size is obtained in section 4. The squared error loss function is discussed in the present context in section 5 and some asymptotic regret and efficiency results for the triple sampling scheme with squared error loss are presented in section 6. Finally, results of a simulation study are presented in section 7.

2. Triple sampling procedure for inference

Since n^{*} in (1.1) is unknown, no fixed sample size procedure provides the above point estimation for μ uniformly for all θ. Therefore, we use a sequential sampling procedure to estimate μ via estimation of the optimal sample size n^{*}. In the following section we give a rigorous account of the triple sampling procedure as described by Hall (1981). As the name suggests, triple sampling can be described by the following three phases:

Pilot Phase: An initial random sample $X_1, X_2, ..., X_m$ of size $m \geq 2$ is taken from the distribution $F(.)$. We use $\bar{X}_m$ and S_m^2 as our initial estimators for μ and θ respectively.

Main Study Phase: Let δ $(0 < \delta < 1)$ be a fixed design factor. The sample size required to complete the main study phase is defined by

$$N_1 = \max\{m, \ [\delta \lambda g(S_m^2)] + 1\}, \qquad (2.1)$$

where $[x]$ is the integer part of x. If $m \geq [\delta \lambda g(S_m^2)] + 1$, we stop sampling at this stage. Otherwise, we continue to observe an additional random sample $X_{m+1}, X_{m+2}, ..., X_{N_1}$ of size $N_1 - m$

from $F(.)$. We augment the $N_1 - m$ observations with the previous m observations and estimate the parameters using $\bar{X}_{N_1}$ and $S^2_{N_1}$.

Fine Tuning Phase: Let

$$N = \max\{N_1, \ [\lambda\, g(S^2_{N_1})]+1 \ \}. \qquad (2.2)$$

If $N_1 \geq \left[\lambda g(S^2_{N_1})\right]+1$, we stop at this stage. Otherwise we continue and take $N - N_1$ more observations from $F(.)$, $X_{N_1+1}, X_{N_1+2}, ..., X_N$, after which sampling is terminated and we propose $\bar{X}_N$ as a point estimator of μ.

Throughout the following sections, the asymptotic characteristics of the triple sampling scheme are developed under the assumption made by Hall (1981) that

$$\ell im \ \ Sup(m/n^*) < \delta \ \ as \ \ m \to \infty, \ \ and \ \ n^* = O(m^r), \ \ for \ \ r > 1 \qquad (2.3)$$

Before developing the theory of the triple sampling procedure specified above, we state the following preliminary results in the fixed sample size case which will facilitate proofs of the theorems to follow.

Lemma 1

Let X be a random variable with distribution function $F(.)$ as above and let $Z = X - \mu$. For a random sample of $Z_1, Z_2, ..., Z_m$ we have:

i) $E(Z) = 0, \ \ E(Z^2) = \theta, \ \ E(Z^3) = \gamma\,\theta^{3/2}, \ \ E(Z^4) = \beta\,\theta^2$.

ii) $E(\bar{Z}_m) = 0, \ \ E(\bar{Z}_m{}^2) = \dfrac{\theta}{m}, \ \ E(\bar{Z}_m{}^3) = \dfrac{\gamma\,\theta^{3/2}}{m^2}, \ \ E(\bar{Z}_m{}^4) = \dfrac{\theta^2(\beta+3m-3)}{m^3}$

iii) $E(\overline{Z_m}\, S_m{}^2) = \dfrac{\gamma\,\theta^{3/2}}{m}, E(\bar{Z}_m{}^2\, S_m{}^2) = \dfrac{\theta^2(\beta+m-3)}{m^2}, \ \ E(\sum_{i\neq j}^{m}\sum Z_i Z_j\, S_m{}^2) = -2\theta^2$.

Proof of (i) is immediate and for (ii) the two terms $E\left(\bar{Z}_m{}^3\right)$ and $E\left(\bar{Z}_m{}^4\right)$ follow immediately from Rohatgi (1976, p303). The proof of (iii) follows by taking the expectation over the identities:

$$\bar{Z}_m\, S_m{}^2 = \left(m(m-1)\right)^{-1}\left\{\sum_{i=1}^{m} Z_i^3 + \sum_{i\neq j}^{m}\sum Z_i Z_j{}^2 - m^2\bar{Z}_m{}^3\right\},$$

$$\bar{Z}_m{}^2\, S_m{}^2 = (m-1)^{-1}\left\{m^{-2}\left[\sum Z_i^4 + \sum_{i\neq j}\sum Z^2{}_i Z_j{}^2 + \sum Z_i^2 \sum_{i\neq j}\sum Z_i Z_j\right] - m\bar{Z}_m{}^4\right\}$$

and

$$\sum_{i\neq j}\sum Z_i Z_j S_m^{\ 2} = m^{-1}\left[2\sum_{i\neq j}\sum Z_i^3 Z_j + \sum_{i\neq j\neq k}\sum\sum Z_i^{\ 2} Z_j Z_k\right]$$
$$-(m(m-1))^{-1}\left[2\sum_{i\neq j}\sum Z_i^{\ 2} Z_j^{\ 2} + 4\sum_{i\neq j\neq k}\sum\sum Z_i^{\ 2} Z_j Z_k + \sum_{i\neq j\neq k\neq l}\sum\sum Z_i Z_j Z_k Z_l\right].$$

3. Asymptotic characteristics of the triple sampling procedure

The asymptotic characteristics of the triple sampling procedure are thoroughly discussed through the following theorems. Theorem 1 provides results regarding the asymptotic characteristics of the main study phase. Specifically, second order approximations of the expectation and the variance of the second stage sample mean are given as the initial sample size gets large.

Theorem 1

For the triple sampling rule (2.1)-(2.2), if condition (2.3) holds, then as $m \to \infty$, we have

$$i) E(\bar{X}_{N_1}) = \mu - \gamma\,\theta^{3/2}\frac{d}{d\theta}\ell n(g(\theta))\ (\delta n^*)^{-1} + o(\lambda^{-1})$$
$$ii) Var(\bar{X}_{N_1}) = \theta(\delta n^*)^{-1} - 2\theta^2(\beta-3)\frac{d}{d\theta}\ell n(g(\theta))\ (\delta n^*)^{-2} + o(\lambda^{-2})$$

Proof: To prove (i), consider the transformation $Z = X - \mu$, and we may write

$$E(\bar{X}_{N_1}) = E\left(N_1^{-1}\sum_{i=1}^{N_1} Z_i\right) + \mu = E\left\{E\left(N_1^{-1}\sum_{i=1}^{N_1} Z_i \,\middle|\, N_1\right)\right\} + \mu. \qquad (3.1)$$

Then, conditioning on the $\sigma-$field generated by the random variables $Z_1, Z_2, ..., Z_m$, we have

$E(\bar{X}_{N_1}) = E\left\{N_1^{-1}E\left(\sum_{i=1}^{m} Z_i + \sum_{i=m+1}^{N_1} Z_i \,\middle|\, Z_1, Z_2, ..., Z_m\right)\right\} + \mu.$ Given $Z_1, Z_2, ..., Z_m$, the first sum, $\sum_{i=1}^{m} Z_i$, is non-random, and the second sum has an expectation zero. Hence:

$$E(\bar{X}_{N_1}) = E\left\{N_1^{-1}\sum_{i=1}^{m} Z_i\right\} + \mu \qquad (3.2)$$

A Taylor series expansion of N_1^{-1} around δn^* gives:

$$N_1^{-1} = (\delta n^*)^{-1} - (\delta n^*)^{-2}(N_1 - \delta n^*) + \eta^{-3}(N_1 - \delta n^*)^2, \qquad (3.3)$$

where η is a random variable lying between N_1 and δn^*. For large λ we approximate $N_1 \approx \delta\lambda g(S_m^{\ 2})$ *as* $\lambda \to \infty$, where g is a differentiable bounded function around θ. By expanding $g(S_m^2)$ and $g^2(S_m^2)$ around θ and substituting in (3.3), we obtain:

$$E\left(N_1^{-1}\sum_{i=1}^{m} Z_i \right) = -(\delta n^*)^{-1}\left[\frac{d}{d\theta}\ln g(\theta) \right]\gamma\,\theta^{3/2} + o(\lambda^{-1}), \qquad (3.4)$$

where the remainder term is of order $o(\lambda^{-1})$. By substituting this in (3.2), (i) follows.

To prove (ii), we also write

$$Var(\bar{X}_{N_1}) = E(\bar{X}_{N_1} - \mu)^2$$

$$= E\left(E\left[N_1^{-2}\left(\sum_{i=1}^{N_1} Z_i^{\,2} + \sum_{i\neq j}^{N_1}\sum Z_i Z_j \right) \middle| Z_1, Z_2, ..., Z_m \right] \right). \qquad (3.5)$$

The first term in (3.5), conditioned on the $\sigma-$ field generated by the random variables $Z_1, Z_2, ..., Z_m$, can be written as

$$E\left(N_1^{-2}\sum_{i=1}^{N_1} Z_i^{\,2} \right) = E\left\{ N_1^{-2}\, E\left(\sum_{i=1}^{m} Z_i^{\,2} + \sum_{i=m+1}^{N_1} Z_i^{\,2} \middle| Z_1, Z_2, ..., Z_m \right) \right\} \qquad (3.6)$$

and therefore,

$$E\left(N_1^{-2}\sum_{i=1}^{N_1} Z_i^{\,2} \right) = E\left(N_1^{-2}\sum_{i=1}^{m} Z_i^{\,2} \right) + \theta E\left(N_1^{-1} - m\, N_1^{-2} \right). \qquad (3.7)$$

The first term on the right hand side of (3.7) can be treated along the same lines that led to (3.4). The second term of (3.7) can be written as

$$E\left\{ N_1^{-2}\left(N_1 - m \right) \right\} \le E\left(N_1^{-1} \right)$$

$$= o(\lambda^{-1}) \quad as \;\; m \to \infty$$

Finally,

$$E\left(N_1^{-2}\sum_{i=1}^{N_1} Z_i^2 \right) = \theta\left(\delta n^* \right)^{-1} - 2\theta^2\left(\beta - 1 \right)\left(\delta n^* \right)^{-2}\left[\frac{d}{d\theta}\ln g(\theta) \right] + o(\lambda^{-1}). \qquad (3.8)$$

Similarly, the second term of (3.5) yields

$$E\left(N_1^{-2}\sum_{i\neq j}^{N_1}\sum Z_i Z_j \right) = 4\theta^2\left(\delta n^* \right)^{-2}\left[\frac{d}{d\theta}\ln g(\theta) \right] + o(\lambda^{-2}), \qquad (3.9)$$

where we have used Lemma 1 and the assumption that g and its derivative are bounded. Hence, the proof of part (ii) is complete.

From part (i) of Theorem 1 the bias of $\bar{X}_{N_1}$ depends on the variance and skewness of the underlying distribution and the form of the function *g*, as well as on the optimal sample size n^* and the design factor δ. In the normal case, where $\gamma = 0$, we observe that the bias is zero.

The variance of the second stage sample mean $Var(\overline{X}_{N_1})$ depends on $g'(\theta)$, the kurtosis, the variance θ, n^* and δ. Suppose $g'(\theta) > 0$. Then platykurtic, leptokurtic and mesokurtic distributions yield estimators of the mean with variances larger than, less than and equal to $\theta(\delta n^*)^{-1}$ respectively. Note that, for large n^*, the variance of $\overline{X}_{N_1}$ tends to zero, as expected.

It is of interest to see whether the fine tuning stage (third stage) reduces the amount of bias evident from Theorem 1. The following results in Theorem 2 provide asymptotic characteristics of the second stage sample size N_1 which are essential for proving Theorems 3 and 4 later.

Lemma 2

For the triple sampling rule (2.1)-(2.2), if condition (2.3) holds, then conditioning on the σ-field generated by $Z_1, Z_2, \ldots, Z_m$, we have

$$(i)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1}\sum_{i\neq j}^{N_1} Z_i Z_j \,\middle|\, Z_1, Z_2, \ldots, Z_m\right] = -2N_1^{-2}\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i^2 Z_j^2 + \theta^2\left(N_1 - m\right)\left(N_1 - m - 1\right)\right)$$

$$(ii)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1} Z_i^2 \,\middle|\, Z_1, Z_2, \ldots, Z_m\right] = N_1^{-1}\Big(\sum_{i=1}^{m} Z_i^4 + \beta\theta^2\left(N_1 - m\right) + \sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i^2 Z_j^2 + \theta^2\left(N_1^2 - 2mN_1 - N_1 + m^2 + m\right)\Big) - \theta^2 N_1$$

$$(iii)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1} Z_i \,\middle|\, Z_1, Z_2, \ldots, Z_m\right] = N_1^{-1}\left(\sum_{i=1}^{m} Z_i^3 + \sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i^2 Z_j - \mu\sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i Z_j\right) + N_1^{-1}\left(\gamma\theta^{3/2}\left(N_1 - m\right)\right) + \mu\left(\sum_{i=1}^{m} Z_i^2 + \theta\left(N_1 - m\right)\right)$$

$$(iv)\; E\left[N_1\left(S_{N_1}^2 - \theta\right)\middle| Z_1, Z_2, \ldots, Z_m\right] = -N_1^{-1}\sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i Z_j$$

$$(v)\; E\left[N_1\left(S_{N_1}^2 - \theta\right)^2 \,\middle|\, Z_1, Z_2, \ldots, Z_m\right] = \theta^2 N_1 + N_1^{-1}\left(\sum_{i=1}^{m} Z_i^4 + \beta\theta^2\left(N_1 - m\right)\right) + N_1^{-1}\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i^2 Z_j^2 + \theta^2\left(N_1^2 - 2mN_1 - N_1 + m^2 + m\right)\right) + N_1^{-3}\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m} Z_i^2 Z_j^2 + \theta^2\left(N_1 - m\right)\left(N_1 - m - 1\right)\right)$$

Theorem 2

For the triple sampling rule (2.1)-(2.2), if condition (2.3) holds, then as $m \to \infty$, we have

$$(i)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1}\sum_{i\neq j}^{N_1} Z_i Z_j\right] = -2\theta^2 + 2\theta^2\left(\delta n^*\right)^{-1} + o(\lambda^{-1})$$

$$(ii)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1} Z_i^2\right] = \theta^2\left(\beta - 1\right) + o(\lambda^{-1})$$

$$(iii)\; E\left[\left(S_{N_1}^2 - \theta\right)\sum_{i=1}^{N_1} Z_i\right] = \gamma\,\theta^{3/2} + o(\lambda^{-1})$$

$(iv)\ E\left[N_1\left(S_{N_1}^2-\theta\right)\right]=-2\theta^2\left(\delta n^*\right)^{-1}\left[\frac{d}{d\theta}\ln g(\theta)\right]+o(\lambda^{-1})$

$(v)\ E\left[N_1\left(S_{N_1}^2-\theta\right)^2\right]=\theta^2(\beta-1)+2\theta^2\left(\delta n^*\right)^{-1}+o\left(\lambda^{-1}\right)$

Proof:
To prove (i), we write

$$E\left[\left(S_{N_1}^2-\theta\right)\sum_{i=1}^{N_1}\sum_{i\neq j}^{N_1}Z_iZ_j\right]=E\left\{E\left[\left(S_{N_1}^2-\theta\right)\sum_{i=1}^{N_1}\sum_{i\neq j}^{N_1}Z_iZ_j\,|Z_1,Z_2,\ldots Z_m\right]\right\}$$

Hence, from (i) of lemma 2, we have

$$E\left[\left(S_{N_1}^2-\theta\right)\sum_{i=1}^{N_1}\sum_{i\neq j}^{N_1}Z_iZ_j\right]=E\left\{-2N_1^{-2}\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m}Z_i^{\,2}Z_j^{\,2}+\theta^2\left(N_1-m\right)\left(N_1-m-1\right)\right)\right\}.$$

Consider the expansion of N_1^{-1} and N_1^{-2} around δn^*. The first term leads to

$$E\left\{-2N_1^{-2}\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m}Z_i^{\,2}Z_j^{\,2}\right)\right\}=-2\theta^2+2\theta^2\left(\delta n^*\right)^{-1}+o(\lambda^{-1})$$

and the second term leads to

$$E\left\{-2N_1^{-2}\left(\theta^2\left(N_1-m\right)\left(N_1-m-1\right)\right)\right\}=o(\lambda^{-1})\ ,$$

where we have used the assumptions in (2.3) and the fact that g and its derivatives are bounded. Therefore, (i) follows.

Similar arguments can be used to verify (ii) and (v) using $E\left[N_1^{-1}\sum_{i=1}^{N_1}Z_i^{\,4}\right]=\beta\,\theta^2+o(\lambda^{-1})$ and $E\left[N_1^{-3}\sum_{i\neq j}^{N_1}\sum Z_i^{\,2}Z_j^{\,2}\right]=o(\lambda^{-1})$. Part (iii) follows along similar lines and the fact that $E\left[N_1^{-1}\sum_{i=1}^{m}Z_i^{\,3}\right]=\gamma\,\theta^{3/2}+o(\lambda^{-1})$. To prove (iv), recall (iv) of Lemma 2 and part (iii) of Lemma 1.

The proof of Theorem 2 is completed. We delete details for brevity. The following Theorem 3 presents the asymptotic characteristics of the third stage sample.

Theorem 3

For the triple sampling rule (2.1)-(2.2), if condition (2.3) holds, then as $m\to\infty$, we have

$$i)\, E(\bar{X}_N) = \mu - \gamma\, \theta^{3/2} \left(n^*\right)^{-1} \left[\frac{d}{d\theta}\left(\ln(g(\theta))\right) \right] + o(\lambda^{-1})$$

$$ii) Var(\bar{X}_N) = \theta\left(n^*\right)^{-1} - 2\theta^2 \left(\beta - 3\right)\left(n^*\right)^{-2} \left[\frac{d}{d\theta}\left(\ln\left(g(\theta)\right)\right) \right] .$$

$$+\theta^3(\beta-1)\delta^{-1}(n^*)^{-2}\left[2\left(\frac{d}{d\theta}\ln g(\theta)\right)^2 - \frac{1}{2n^*}\frac{d^2 n^*}{d\theta^2}\right] + o(\lambda^{-2})$$

Proof: To prove (i), conditioning on the σ-field generated by the random variables $Z_1, Z_2, ..., Z_{N_1}$, we write

$$E(\bar{X}_N) = E\left\{ N^{-1}\, E\left(\sum_{i=1}^{N_1} Z_i + \sum_{i=N_1+1}^{N} Z_i \,\middle|\, Z_1, Z_2, ..., Z_{N_1} \right) \right\} + \mu \; . \qquad (3.10)$$

Again the first sum $\sum_{i=1}^{N_1} Z_i$ in (3.10) is non-random. Thus, (3.10) reduces to

$$E(\bar{X}_N) = E\left(N^{-1} \sum_{i=1}^{N_1} Z_i \right) + \mu \qquad (3.11)$$

Also, $N \approx \lambda g({S_{N_1}}^2)$, as $\lambda \to \infty$, where $g(.)$ is a differentiable and bounded function around θ. By expanding N^{-1} around n^*, and $g(S^2_{N_1})$ around θ, using Taylor series we obtain

$$E\left(N^{-1} \sum_{i=1}^{N_1} Z_i \right) = -\left[\frac{d}{d\theta} \ln g(\theta) \right] (n^*)^{-1}\, E\left\{ \left({S_{N_1}}^2 - \theta \right) \sum_{i=1}^{N_1} Z_i \right\} + o(\lambda^{-1}). \qquad (3.12)$$

Application of part (iii) of Theorem 2 completes the proof of (i) of Theorem 3.

Proof of (ii) can be obtained directly from the following

$$Var(\bar{X}_N) = E\left(N^{-2} E\left(\sum_{i=1}^{N_1} {Z_i}^2 + \sum_{i \neq j}^{N_1} \sum Z_i Z_j \,\middle|\, Z_1, Z_2, ..., Z_{N_1} \right) \right)$$

$$+\theta E\left(E\left(N^{-1} - N_1 N^{-2} \,\middle|\, Z_1, Z_2, ..., Z_{N_1} \right) \right) \qquad (3.13)$$

By conditioning on the σ-field generated by $Z_1, Z_2, ..., Z_{N_1}$ and expanding N^{-2} and $g({S_{N_1}}^2)$ around n^* and θ respectively, it can be shown that

$$E\left(N^{-2} \left(\sum_{i=1}^{N_1} Z_i^2 \,\middle|\, Z_1, Z_2, ..., Z_{N_1} \right) \right) = \theta\left(n^*\right)^{-2} E(N_1) - 2\theta^2 \left(\beta - 1\right)\left(n^*\right)^{-2} \left[\frac{d}{d\theta} \ln g(\theta) \right] + o(\lambda^{-2}) \qquad (3.14)$$

$$E\left(N^{-2} \left(\sum_{i \neq j}^{N_1} \sum Z_i Z_j \,\middle|\, Z_1, Z_2, ..., Z_{N_1} \right) \right) = 4\theta^2 \left[\frac{d}{d\theta} \ln g(\theta) \right] \left(n^*\right)^{-2} \left(1 - \left(\delta n^*\right)^{-1}\right) + o(\lambda^{-3}). \qquad (3.15)$$

Also

$$E\left(N^{-1}-N_1N^{-2}\right)=-\left(n^*\right)^{-2}E\left(N_1\right)+\left(n^*\right)^{-1}$$

$$+\theta^2\left(\beta-1\right)\left(n^*\right)^{-2}\left[\frac{g^{''}(\theta)}{g(\theta)}\left(4-\frac{1}{2}\delta^{-1}\right)+2\delta^{-1}\left(\frac{g^{'}(\theta)}{g(\theta)}\right)^2-\frac{3g^{2''}(\theta)}{2g^2(\theta)}\right]+o(\lambda^{-2}) \qquad (3.16)$$

By adding (3.14), (3.15) and (3.16) the proof of Theorem 3 is completed.
In view of (i) and (ii) of Theorem 3, it is worth mentioning that the third stage has indeed reduced the bias noticed in (i) and (ii) of Theorem 1.

Theorem 4 gives asymptotic results regarding the second stage estimator of the unknown variance θ.

Theorem 4

For the triple sampling rule (2.1)-(2.2), if condition (2.3) holds, then as $m\to\infty$, we have:

$$(i)E(S^2_{N_1})=\theta-\theta^2(\beta-1)\left(\frac{d}{d\theta}\ln(g(\theta))\right)(\delta n^*)^{-1}+o(\lambda^{-1})$$

$$(ii)\,E\left(S^4_{N_1}\right)=\theta^2+\theta^2\left(\beta-1\right)\left(\delta n^*\right)^{-1}+o(\lambda^{-1})$$

$$(iii)\ Var(S^2_{N_1})=\theta^2(\beta-1)(\delta n^*)^{-1}+o(\lambda^{-1})$$

$$(iv)E(g(S^2_{N_1}))=g(\theta)-\theta^2(\beta-1)(\lambda\delta n^*)^{-1}\left[\left(n^*\right)^{-1}\left(\frac{dn^*}{d\theta}\right)^2-(1/2)\frac{d^2n^*}{d\theta^2}\right]+o(\lambda^{-1})$$

Proof: To prove (i), we write

$$E\left(S^2_{N_1}\right)=E\left(N_1^{-1}E\left(\sum_{i=1}^{N_1}Z_i^2\,|Z_1,Z_2,\ldots,Z_m\right)\right)$$

$$-E\left(\left(N_1\left(N_1-1\right)\right)^{-1}E\left(\sum_{i\neq j}^{N_1}\sum Z_iZ_j\,|Z_1,Z_2,\ldots,Z_m\right)\right)$$

Consequently

$$E\left(N_1^{-1}\ E\left(\sum_{i=1}^{N_1}Z_i^2\,|Z_1,Z_2,\ldots,Z_m\right)\right)=\theta-\theta^2\left(\beta-1\right)\left(\delta n^*\right)^{-1}\left[\frac{d}{d\theta}\ln g\left(\theta\right)\right]+o(\lambda^{-1})$$

while

$$E\left(\left(N_1\left(N_1-1\right)\right)^{-1}E\left(\sum_{i\neq j}^{N_1}\sum Z_iZ_j\,|Z_1,Z_2,\ldots,Z_m\right)\right)=E\left(\left(N_1\left(N_1-1\right)\right)^{-1}\sum_{i=1}^{m}\sum_{i\neq j}^{m}Z_iZ_j\right)$$

$$\leq E\left(\sum_{i=1}^{m}\sum_{i\neq j}^{m}Z_iZ_j\right)=0 \qquad (3.17)$$

To prove (ii), we write

$$E\left(S_{N_1}^4\right) = E\left(N_1^{-2}\left(E\left(\sum_{i=1}^{N_1} Z_i^4\right)\left|Z_1, Z_2, \ldots, Z_m\right.\right)\right) + E\left(N_1^{-2}E\left(\sum_{i\neq j}^{N_1}\sum Z_i^2 Z_j^2\left|Z_1, Z_2, \ldots, Z_m\right.\right)\right)$$

$$+ E\left(\left(N_1\left(N_1-1\right)\right)^{-2} E\left(\sum_{i\neq j}^{N_1}\sum Z_i^2 Z_j\left|Z_1, Z_2, \ldots, Z_m{}^2\right.\right)\right)$$

Arguments similar to those used to verify (3.17) can be used to prove

$$E\left(\left(N_1\left(N_1-1\right)\right)^{-2} E\left(\sum_{i\neq j}^{N_1}\sum Z_i^2 Z_j\left|Z_1, Z_2, \ldots, Z_m{}^2\right.\right)\right) = 0,$$

while

$$E\left(N_1^{-2}\left(E\left(\sum_{i=1}^{N_1} Z_i^4\right)\left|Z_1, Z_2, \ldots, Z_m\right.\right)\right) = \beta\theta^2\left(\delta n^*\right)^{-1} + o(\lambda^{-1})$$

and

$$E\left(N_1^{-2}E\left(\sum_{i\neq j}^{N_1}\sum Z_i^2 Z_j^2\left|Z_1, Z_2, \ldots, Z_m\right.\right)\right) = \theta^2 - \theta^2\left(\delta n^*\right)^{-1} + o(\lambda^{-1}).$$

The proof of (iii) follows immediately from parts (i) and (ii).
A more general second order approximation of the expectation of a continuously differentiable function g of $S_{N_1}^2$ as $m \to \infty$ is presented in (iv). The proof of (iv) follows by expanding the function $g(S_{N_1}^2)$ around $g(\theta)$ and using (i), (ii) and (iii) of Theorem 4.

Clearly, (i) of Theorem 4 illustrates that $S_{N_1}^2$ is biased for θ. The bias depends on the kurtosis of the underlying distribution, the form of g, δ and n^*. Similar arguments can be applied to discuss (ii) and (iii) above.

Next, the expectation of the final stage sample size N and other asymptotic characteristics may be easily obtained from (iv) of Theorem 4 above as given in the following Theorem 5.

Theorem 5

Let N be defined as in (2.2) and assume that condition (2.3) holds, then as $m \to \infty$, we have:

$$i) E(N) = n^* - \theta^2(\beta-1)(\delta n^*)^{-1}\left(n^*\left(\frac{d}{d\theta}\ell n(n^*)\right)^2 - (1/2)\left(\frac{d^2 n^*}{d\theta^2}\right)\right) + E(\varepsilon_{N_1}) + o(1)$$

$$ii) Var(N) = \theta^2(\beta-1)(\delta n^*)^{-1}\left(\frac{dn^*}{d\theta}\right)^2 + o(\lambda)$$

$$iii) E\,|\,N - E(N)\,|^3 = o(\lambda^2),$$

where the continuous random variable $\varepsilon_{N_1} = 1 - \left\{\lambda g(S_{N_1}^2) - [\lambda g(S_{N_1}^2)]\right\}$ is defined over the interval (0, 1).

Proof: To prove part (i), note that $N = \left[\lambda g(S_{N_1}{}^2)\right] + 1$, *a.s.* except possibly on a set $\xi = \left\{\left(N_1 > \left[\lambda g\left(S_{N_1}^2\right)\right] + 1\right) \cup \left(m > \left[\delta\lambda g\left(S_m^2\right)\right] + 1\right)\right\}$ of measure zero, such that $\int_\xi N \, dP = o(1)$; see, for example, Hall (1981) for details. Further discussion of this will be given later in section 7. Hence,

$$
\begin{aligned}
N &= \left[\lambda g(S_{N_1}{}^2)\right] + 1 \\
&= \lambda g(S_{N_1}{}^2) - \left\{\lambda g(S_{N_1}{}^2) - \left[\lambda g(S_{N_1}{}^2)\right]\right\} + 1 \\
&= \lambda g(S_{N_1}{}^2) + \varepsilon_{N_1} \qquad (3.18)
\end{aligned}
$$

Thus, $E(N) = \lambda E\left(g(S_{N_1}{}^2)\right) + E(\varepsilon_{N_1}), \quad as \ m \to \infty$.

Using Theorem 4 part (iv), we obtain the result.

Proof of (ii)*:* From (3.16), $Var(N) \approx \lambda^2 Var\left(g(S_{N_1}^2)\right) \quad as \quad m \to \infty$,

$$
\begin{aligned}
Var\left(g(S_{N_1}^2)\right) &= Var\left(g(\theta) + g'(\theta)\left(S_{N_1}^2 - \theta\right) + \dots + \frac{1}{2} g''(\tau)\left(S_{N_1}^2 - \theta\right)\right) \\
&= \left(g'(\theta)\right)^2 Var\left(S_{N_1}^2\right) + o(\lambda^{-1}),
\end{aligned}
$$

where τ is a random variable between $S_{N_1}{}^2$ and θ. By using Theorem 4 part (iii), and the assumption that g and its derivatives are bounded, part (iii) of Theorem 5 follows similarly.

Remarks

If the underlying distribution is normal, Hall (1981, pg. 1237) proved that $\varepsilon_{N_1} \xrightarrow{L} U(0,1)$ as $\lambda \to \infty$. More generally, ε_{N_1} is continuous over the interval $(0,1)$ and is independent of λ. This can be shown easily from (2.2) and the inequality $\lambda g\left(S_{N_1}^2\right) < \left[\lambda g\left(S_{N_1}^2\right)\right] + 1 \le \lambda g\left(S_{N_1}^2\right) + 1$. A simulation study of the behaviour of ε_{N_1} for non-normal underlying distributions is given in section 7 and the results tend to support the conjecture that Hall's result applies for any continuous distribution.
It is also evident from Theorem 5 above that Hall's (1981) Theorem 1 is obtained for the normal distribution case, when the optimal sample size $n^* = \lambda\theta$.

We also stress that both the expectation of N and its variance depend on the kurtosis of the underlying distribution and accordingly will reflect the amount of departure from normality while estimating the optimal sample size n^*. To reiterate, for distributions with kurtosis close to 3, $E(N)$ and $Var(N)$ will be little affected by such departures from normality. However, a more substantial effect will accrue for underlying distributions that are either very flat or very peaked.

It is also of interest to consider the general form of the expectation of a real valued continuously differential function on the final stage sample size N to be able to derive asymptotic results for all moments of N, provided they exist. We also stress that we have not assumed independence between $\overline{X}_{N_1}$ and S_m^2 or $\overline{X}_N$ and $S_{N_1}^2$, and therefore the above results are more general in that sense.

Theorem 6

Let $h(>0)$ be a continuously differentiable real valued function in a neighborhood around n^* such that $\sup_{n\geq m} h(n) = O(h'''(n^*))$. Then, as $\lambda \to \infty$

$$E(h(N)) = h(n^*) - \theta^2(\beta-1)(\delta n^*)^{-1}\left\{\left(\frac{dn^*}{d\theta}\right)^2\left[\left(n^*\right)^{-1} h'(n^*) - (1/2)h''(n^*)\right] - (1/2)h'(n^*)\frac{d^2n^*}{d\theta^2}\right\}$$

$$+h'(n^*)E(\varepsilon_{N_1}) + o(\lambda^{-1}(|h'''(n^*)|)).$$

Proof: The proof follows by expanding the differentiable function $h(N)$ around $h(n^*)$ using Taylor series and applying the results of Theorem 5. The general form of the second order asymptotic expansion of the expectation of a real valued continuously differentiable function $h(>0)$ enables one to obtain the expectations of positive and negative moments of N in subsequent analysis.

4. Asymptotic normality of the stopping sample size N

Theorem 7

Let N be defined as in (2.2) and assume that condition (2.3) holds. Then as $m \to \infty$, N is asymptotically normal with mean n^* and variance

$$Var(N) = \theta^2(\beta-1)(\delta n^*)^{-1}\left(\frac{dn^*}{d\theta}\right)^2 + o(\lambda).$$

Proof: The proof of Theorem 7 is a straightforward application of Theorem 6 above with $h(\nu) = e^{t\nu}$ and $\nu = \left(N - n^*\right)/\sqrt{Var(N)}$.

5. Squared error loss function to estimate the mean

In this section our main objective is to develop a triple sampling point estimation procedure to estimate the mean μ of the population. In particular, if a point estimate of the unknown μ is required, we assume that the incurred cost of estimating mean μ by the corresponding sample mean $\overline{X}_n$ can be approximated by the following squared error loss function in (5.1) with a linear sampling cost. The literature in sequential sampling has considered several forms of higher order cost functions to model estimation cost. However, squared error loss functions are recommended and commonly used in sequential point estimation problems (see, for example, Degroot, 1962). Therefore, we write the cost (loss) function as

$$L_n(A) = A(\overline{X} - \mu)^2 + Cn, \qquad (5.1)$$

where C is the known cost per unit sample and A is a constant permitted to approach ∞. We elaborate further on determination of A in subsequent developments. The risk associated with (1.1) can be written as

$$R_n(A) = E(L_n(A)) = AE(\overline{X}_n - \mu)^2 + Cn = A\theta / n + Cn \qquad (5.2)$$

Treating n as a continuous variable in (5.2), we differentiate (5.2) with respect to n and equate the results to zero to obtain the optimal sample size as

$$n \geq n^* = \sqrt{\frac{A\ \theta}{C}} \quad . \qquad (5.3)$$

The value of n^* in (5.3) is unknown because the population variance θ is unknown. It has been shown by Dantzig (1940), Stein (1945, 1949) and Seelbinder (1953) that no fixed sample size procedure exists to achieve the above optimal requirement uniformly over θ. In light of (5.3), $n^* \to \infty$ in two ways, first, either A is permitted to approach infinity (extremely high cost of estimation error) or the cost of sampling is cheap.

Since the optimal sample size in (5.3) depends on the unknown variance θ, no fixed sample size procedure can be used to estimate μ uniformly over all θ. Therefore, the triple sampling procedure in (2.1)-(2.2) can be used to provide a point estimator of μ with $\lambda = \sqrt{A/C}$ and $g(\theta) = \sqrt{\theta}$. The question arises: how efficient is this estimator?

6. The asymptotic regret and efficiency of triple sampling point estimation under squared error loss function

In the literature on sequential point estimation several measures have been developed of the efficiency of the sequential procedures (triple sampling, or accelerated sequential schemes) relative to the fixed sample size counterpart had the form of $g(\theta)$ in (1.1) been completely specified. The regret reflects the expected cost of missed opportunity which measures the risk in using the triple sampling procedures to perform point estimation of the population mean instead of the fixed sample size procedure, had the nuisance parameter(s) been known. Other weaker measures like the asymptotic relative efficiency, $\eta(A) = E(L_N(A))/E(L_{n^*}(A))$, which is the ratio of the triple sampling risk to the optimal risk are also used. For an efficient sampling procedure we expect $\eta(A) \to 1$ and that $\omega(A)$ is bounded as $n^* \to \infty$.

Theorem 8

The risk associated with the squared error loss equation (5.1) as $m \to \infty$ is given by

$$\begin{aligned} R_N(A) &= E\left[L_N(A)\right] \\ &= 2Cn^* - C(\beta - 3) + (1/4)(\beta - 1)C\delta^{-1} + C\,E\left(\varepsilon_{N_1}\right) + o(1) \end{aligned}$$

Moreover, the asymptotic relative efficiency of the triple sampling scheme and the asymptotic regret are given by

$i)\ \eta(A) = 1 + o(\lambda^{-1})$,

$ii)\ \omega(A) = -C(\beta-3) + (1/4)(\beta-1)C\delta^{-1} + C\,E(\varepsilon_{N_1}) + o(1)$, as $m \to \infty$.

Proof of Theorem 8:

This is immediate if we recall (ii) of Theorem 3 and (i) of Theorem 5. Moreover, the regret of triple sampling associated with equation (5.1) is given by

$$\omega(A) = R_N(A) - R_{n^*}(A)$$
$$= -C(\beta-3) + (1/4)(\beta-1)C\delta^{-1} + C\,E(\varepsilon_{N_1}) + o(1)$$

If the cost of unit sampling tends to zero, then we expect zero regret. However, for C non-zero, the regret is bounded, as illustrated in Theorem 8, and depends on the kurtosis β.

Theorem 8 has several consequences. First, the case of normal distributions treated by Mukhophadhyay et al. (1987), Hamdy (1988) and Hamdy et al. (1988) are special cases. Secondly, for distributions with β<3 a non-vanishing positive regret is expected. In addition, for distributions with β>3 we expect either positive or negative non-vanishing regret, depending on the values of the kurtosis and the design factor δ. Specifically, for distributions with $\beta > 6$, negative regret is expected with $\delta = 1/2$. Martinsek (1988) argued that for one-by-one sequential procedures negative regret is expected when β>3. It is also worth mentioning that the regret of purely sequential procedures depends on both the kurtosis and skewness of the underlying distribution, as indicated by Martinsek (1988), while our findings in Theorem 8 emphasize that the triple sampling procedure depends only on the kurtosis. This could be due to the nature of one-by-one purely sequential procedure which filters data. This filtration may cause either acceleration or delay. On the other hand, triple sampling is based on bulks (batches). Therefore, if an extreme observation presents, its effect on such the decision to stop or continue sampling will be diluted by the rest of the bulk at that stage, which may cause the triple sampling procedure to be less sensitive to extreme observations than one-by-one sequential procedures.

However, a general formula for the regret incurred in estimating μ with squared error loss function (6.1) can be written as

$$\omega(A) = -2C\theta(\beta-3)\left(\frac{g'(\theta)}{g(\theta)}\right) + C\theta^2(\beta-1)\delta^{-1}\left(\frac{g'(\theta)}{g(\theta)}\right)^2 + CE(\varepsilon_{N_1}) + o(1) \qquad (6.1)$$

Obviously, the non-vanishing regret in (6.1) depends on β, δ, C and the form of $g(\theta)$. To elaborate further, consider the loss function in Martinsek (1988) of the form

$$L_n(A) = A\theta^{b-1}(\bar{X}_n - \mu)^2 + n\,.$$

This, under the triple sampling scheme, provides the following regret:

$$\omega(A) = -(\beta-3)b + b^2(\beta-1)/4\delta + E(\varepsilon_{N_1}) + o(1) \qquad (6.2)$$

The regret of the triple sampling procedure in (6.2) is the same as Martinsek's (1988) equation (7) for symmetric underlying distributions.

The above results are asymptotic and therefore to study the small to moderate sample size performance of triple sampling a series of Monte Carlo simulation experiments were performed and are presented in section 7.

7. Simulation results

Since the results obtained above are asymptotic in nature, a series of simulation studies, each based on 50,000 replications were undertaken to investigate the small, moderate, and large sample size performance of the normal-based triple sampling procedure (2.1) and (2.2) under squared error loss with (without loss of generality) C=1 and consequently from (5.3) $A = (n^*)^2 / \theta$. The optimal sample size n^* was allowed to vary from small to large (24, 43, 61, 76, 96, 125, 171, 246, 500); δ = 0.3, 0.5, 0.8; m =10, 15 and the underlying distributions were $U(0, 1)$, $N(0, 1)$ and exponential with mean 1, thereby giving θ = 1/12, 1 and 1 respectively. For each replicate the stopping sample size, the stage at which sampling was stopped, the estimate of the mean, the loss and the value of ε_{N_1} were recorded. Relevant summary statistics for each combination of conditions were recorded across replicates and these results form the basis of the discussion given below in sections 7.1 and 7.2.

7.1 The estimator for the mean and the stopping sample size

Table 1 shows some results for the $N(0, 1)$ underlying distribution. All the selected values of n^* are included for the case m = 15 and δ = 0.5. Here the estimated mean final sample size, $\bar{N}$, and its standard error, the mean estimate for μ , $\hat{\mu}$, its standard error and estimated sampling variance, and the estimated regret, $\hat{\omega}$, are shown. Corresponding results are shown in Table 2 for the $U(0, 1)$ distribution but with m = 10, and in Table 3 for the exponential distribution with mean 1, again with m = 15. Only the δ = 0.5 results are shown because these always gave the most satisfactory results.

In Tables 1 and 2 there is good agreement with the asymptotic results (Theorem 3) for the estimator of the mean: it is clearly unbiased and the estimated variance is close to the asymptotic value except for n^*=24, where m/n^* is not less than the design factor δ (as required by (2.3)). Again, apart from n^*=24, the estimated mean stopping sample size agrees well with the results of Theorem 5.

In Table 3 the slight negative bias for the estimator of the mean predicted by Theorem 3 is clearly present but the asymptotic variance is much lower than the actual estimated variance unless n^* is quite large. The observed tendency for early stopping seen in Table 3 is in line with the result of Theorem 5 where the large kurtosis of the exponential distribution seriously affects E(N). The estimated regret is rather volatile in the simulation results but it is interesting that in Table 3 the estimated regret approaches the limiting value, -1.5, from above.

(Tables 1, 2 and 3 near here)

7.2 Termination stage and the distribution of ε_{N_1}

Tables 4, 5 and 6 show the estimated probabilities of stopping after the first and second stages, together with summary statistics for the mean and standard deviation of ε_{N_1}, together with the p-value of a Kolmogorov-Smirnov test of uniformity for the distribution of ε_{N_1}.

Table 6 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the exponential case with mean 1 with $m = 15$ and $\delta = 0.5$.

For the normal and uniform cases (Tables 4 and 5) there is little early stopping unless n^* is small, whereas for the exponential case the proportion of early stopping declines more slowly.

For all three distributions the hypothesis that the ε_{N_1} follow a $U(0, 1)$ distribution mostly cannot be rejected even on the basis of 50,000 replicates when n^*, which suggests that the asymptotic result of Hall(1981) for the distribution of ε_{N_1} may also hold for other underlying distributions.

(Tables 4, 5 and 6 near here)

References

Anscombe, F.J., 1953. Sequential estimation. *J. Roy. Statist. Soc.* Ser. B **15** 1-21.
Bhattacharjee, G.P., 1965. Effect of non-normality on Stein’s two sample test. *Ann. Math. Statist*. **36**, No. 2, 651-663.
Bhattacharjee, G.P., Nagendra, Y., 1964. Effect of non-normality on a sequential test for mean. *Biometrika,* **51** 281-287.
Blumenthal, S., and Govindarajulu, Z., 1977. Robustness of Stein’s two-stage procedure for mixtures of normal populations. *J. Amer. Statist. Assoc*. **72** 192-196.
Chow, Y.S., and Robbins, H., 1965. On the asymptotic theory of fixed width confidence intervals for the mean. *Ann. Math. Statist*. **36** 457-462.
Dantzig, G.B., 1940. On the non existence of tests of “students” hypothesis having power functions independent of σ . *Ann. Math. Statist*. **11** 186- 192.
Hall, P., 1981. Asymptotic theory of triple sampling for sequential estimation of a mean. *Ann. Statist.*, **9** 1229-1238.
Hamdy, H.I., 1988. Remarks on the asymptotic theory of triple stage estimation of the normal mean. *Scand. J. Statist*. **15** 303-310.
Hamdy, H.I., Mukhophadhyay, N., Costanza, M.C., and Son, M.C., 1988. Triple stage point estimation for the exponential location parameter. *Ann. Instit. Statist. Math*., **40** 785-797.
Hamdy, H.I., AL-Mahameed, M., Nigm, A., and Son, M.S., 1989. Three stage estimation procedure for the exponential location parameters. *Metron*, XLVII, 279-294.
Hamdy, H.I., Al-Mahameed, M.A., Son, M.S., and AL-Hussainan, A., 1998. Three stage estimation for the mean of a one parameter exponential family. *The Korean Commun. Statist*. **5** 539-557.
Hamdy, H.I., and Plotter, W. J., 1987. Triple sampling procedure for estimating the scale parameter of Pareto distribution. *Commun. Statist., Theory and Methods*, **16** 2155-2164.

Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994. Continuous Univariate Distributions, Vol. I. Wiley, New York.
Jureckova, J., and Sen, P.K., 1940. Robust Statistical Procedures, Asymptotics and Interrelations. Wiley, New York.
Liu, W., 1997. A k-stage sequential sampling procedure for estimation of normal mean. *J. Statist. Planning and Inference*. **65** 109-127.
Martinsek, A.T., 1988. Negative regret, optimal stopping, and the elimination of outliers. *J. Amer. Statist. Assoc*. **83** 160-163.
Mukhophadhyay, N., 1985. A note on three stage and sequential point estimation procedures for a normal mean. *Sequential Anal*., **4** 311-319.
Mukhophadhyay, N., Hamdy H.I., AL-Mahmeed M., and Costanza M.C., 1987. Three stage point estimation procedures for a normal mean. *Sequential Anal*., **6** 21-36.
Mukhophadhyay, N., and Mavromoustakos, A., 1987. Three stage estimation procedures of the negative exponential distribution. *Metrika*, **34** 83-93.
Ramkaran, 1983. The robustness of Stein's two-stage procedure. *Ann. Statistic.* ***11***, No. 4, 1251-1256.
Robbins, H., 1959. Sequential estimation of the mean of a normal population. Probability and Statistics (Harald Cramer Volume) 235-245. Almquist and Wiksell, Uppsala, Sweden.
Rohatgi, V.K., 1976. An Introduction to Probability Theory and Mathematical Statistics. Wiley, New York, 303.
Seelbinder, D.M., 1953. On stein's two-stage procedure. *Ann. Math. Statist.***24** 640-649. [458, 459]
Stein, C., 1945. A two-stage sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist*. **16** 243-258.
Stein, C. 1949. Some problems in sequential estimation. *Econometrika*. **17** 77-78.

Table headings:

Table 1 Simulation results for the $N(0, 1)$ case with $m = 15$ and $\delta = 0.5$.
Table 2 Simulation results for the $U(0, 1)$ case with $m = 10$ and $\delta = 0.5$.
Table 3 Simulation results for the exponential case (mean 1) with $m = 15$ and $\delta = 0.5$.
Table 4 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the $N(0, 1)$ case with $m = 15$ and $\delta = 0.5$.
Table 5 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the $U(0, 1)$ case with $m = 10$ and $\delta = 0.5$.
Table 6 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the exponential case with mean 1 with $m = 15$ and $\delta = 0.5$.

n^*	$\overline{N}$	s.e.($\overline{N}$)	$\widehat{\mu}$	s.e.($\widehat{\mu}$)	Est. var. of $\widehat{\mu}$	$\widehat{\omega}$
24	16.35	0.021	-0.0009	0.0011	0.06403155	5.2324
43	41.07	0.039	0.0001	0.0007	0.02642890	3.9383
61	59.73	0.038	0.0001	0.0006	0.01717961	1.6524
76	74.80	0.042	-0.0002	0.0005	0.01368532	1.8419
96	94.83	0.046	0.0002	0.0005	0.01061260	0.6333
125	123.86	0.053	-0.0003	0.0004	0.00815736	1.3200
171	169.89	0.061	-0.0002	0.0003	0.00593427	1.4134
246	244.90	0.073	0.0003	0.0003	0.00412565	2.5638
500	498.94	0.103	0.0003	0.0002	0.00200386	-0.1007

Table 1 Simulation results for the $N(0, 1)$ case with $m = 15$ and $\delta = 0.5$.

n^*	$\overline{N}$	s.e.($\overline{N}$)	$\widehat{\mu}$	s.e.($\widehat{\mu}$)	Est. var. of $\widehat{\mu}$	$\widehat{\omega}$
24	22.36	0.027	0.5001	0.0003	0.00512696	9.8001
43	42.64	0.022	0.5001	0.0002	0.00210675	3.3893
61	60.73	0.025	0.5002	0.0002	0.00141767	2.0271
76	75.77	0.027	0.5002	0.0001	0.00111595	1.1225
96	95.73	0.030	0.5000	0.0001	0.00088364	1.4543
125	124.78	0.034	0.4999	0.0001	0.00067339	1.0434
171	170.80	0.039	0.4999	0.0001	0.00049422	2.2209
246	245.77	0.047	0.5000	0.0001	0.00034111	1.4751
500	499.86	0.065	0.5000	0.0001	0.00016605	-1.9790

Table 2 Simulation results for the $U(0, 1)$ case with $m = 10$ and $\delta = 0.5$.

n^*	$\overline{N}$	s.e.($\overline{N}$)	$\widehat{\mu}$	s.e.($\widehat{\mu}$)	Est. var. of $\widehat{\mu}$	$\widehat{\omega}$
24	18.07	0.032	0.9712	0.0010	0.0493740	-1.4936
43	36.76	0.064	0.9519	0.0008	0.0362480	17.7787
61	55.44	0.072	0.9723	0.0007	0.0239964	22.7263
76	70.58	0.080	0.9806	0.0006	0.0173673	18.8950
96	90.82	0.091	0.9866	0.0005	0.0126232	15.1528
125	119.65	0.104	0.9906	0.0004	0.0088950	8.6339
171	165.92	0.123	0.9936	0.0004	0.0062784	7.5051
246	240.73	0.150	0.9951	0.0003	0.0043172	9.9935
500	494.89	0.218	0.9976	0.0002	0.0020597	9.8092

Table 3 Simulation results for the exponential case (mean 1) with m = 15 and δ = 0.5.

n^*	$P(N=m)$	$P(N=N_1)$	$E(\varepsilon_{N_1})$	σ_{N_1}	p-value
24	0.768	0.000	0.4965	0.2880	0.007
43	0.221	0.001	0.4953	0.2885	0.001
61	0.061	0.003	0.4981	0.2883	0.032
76	0.024	0.005	0.4967	0.2884	0.003
96	0.007	0.007	0.4998	0.2890	0.821
125	0.002	0.009	0.5011	0.2885	0.507
171	0.000	0.010	0.4991	0.2879	0.507
246	0.000	0.012	0.4974	0.2889	0.015
500	0.000	0.013	0.4982	0.2884	0.154

Table 4 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the N(0, 1) case with m = 15 and δ = 0.5.

n^*	$P(N=m)$	$P(N=N_1)$	$E(\varepsilon_{N_1})$	σ_{N_1}	p-value
24	0.839	0.000	0.4961	0.2884	0.006
43	0.113	0.000	0.4971	0.2894	0.006
61	0.016	0.000	0.5006	0.2884	0.250
76	0.004	0.000	0.4984	0.2884	0.759
96	0.001	0.000	0.4999	0.2897	0.113
125	0.000	0.000	0.4984	0.2882	0.400
171	0.000	0.000	0.5011	0.2895	0.151
246	0.000	0.000	0.4974	0.2887	0.211
500	0.000	0.000	0.4999	0.2875	0.303

Table 5 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the $U(0, 1)$ case with $m = 10$ and $\delta = 0.5$.

n^*	$P(N=m)$	$P(N=N_1)$	$E(\varepsilon_{N_1})$	σ_{N_1}	p-value
24	0.748	0.010	0.4951	0.2891	<0.001
43	0.413	0.035	0.4990	0.2894	0.131
61	0.232	0.056	0.4981	0.2881	0.076
76	0.150	0.065	0.4964	0.2889	0.013
96	0.084	0.072	0.4985	0.2887	0.164
125	0.040	0.078	0.4986	0.2884	0.145
171	0.014	0.085	0.4988	0.2893	0.324
246	0.005	0.083	0.4977	0.2885	0.019
500	0.000	0.084	0.5012	0.2883	0.439

Table 6 Estimated probabilities of early stopping and the behaviour of ε_{N_1} for the exponential case with mean 1 with $m = 15$ and $\delta = 0.5$.