Journal of the Korean Statistical Society (2002), **31**: *4, pp 415–431*

An Integrated Sequential Inference Approach for the Normal Mean

M. A. Almahmeed[1], H. I. Hamdy[1,2], Y. H. Alzalzalah[1] and M. S. Son[2]

ABSTRACT

A unified framework for statistical inference for the mean of the normal distribution to derive point estimates, confidence intervals and statistical tests is proposed. This optimal design is justified after investigating the basic information and requirements that are possible and impossible to control when specifying practical and statistical requirements. Point estimation is only credible when viewed in the larger context of interval estimation, since the information required for optimal point estimation is unspecifiable. Triple sampling is proposed and justified as a reasonable sampling vehicle to achieve the specifiable requirements within the unified framework.

Keywords. Confidence intervals, cost functions, decision theory, fixed precision hypothesis tests, Fisher information, loss functions, opportunity cost, point estimation, regret, risk functions, sampling cost, sequential sampling, squared error loss, triple sampling, type II error.

AMS 2000 subject classifications. Primary 62L10; Secondary 62L12.

1. Introduction

The literature in sequential analysis usually considers two main methodologies to tackle estimation problems: (1) point estimation, where a specific loss function is assumed to assess the encountered risk; (2) fixed width confidence interval estimation to attain a given nominal coverage value. To the best of our knowledge, these two methodologies have been treated as completely separate approaches to inference for the parameter(s) in the situations examined and it seems that few attempts have been made to combine these two methodologies

Received February 2000; accepted June 2002.

[1]Department of Quantitative Methods and Information Systems, Kuwait University, Kuwait

[2]Department of Mathematics and Statistics, University of Vermont, Burlington, VT 05401, U.S.A.

under a single unified framework to achieve maximal use of the available sample information to handle both problems simultaneously. It is certainly reasonable to inquire as to whether and when one might want to combine the criteria for point and fixed width confidence interval estimation. After all, if (for example) the primary focus in a given problem is on point estimation, should one even be concerned about confidence intervals (or *vice versa*)? In our view, the answers to both these questions should be a resounding "*yes*". We take this position in light of considering the most basic, inherent commonalities of point and interval estimators. In Section 2 we lay down the philosophical groundwork for determining what information can and can't be learned by sampling in the context of a "purely" point estimation problem. We conclude that inference in that context is essentially without credible content. In Section 3 we propose a first approach to credible estimation based on fixed width confidence interval estimators with specified coverage, which points the way towards credible point estimators as well. A second approach to credible testing based on confidence intervals controlled for type II error is also proposed along similar lines in Section 3. Section 4 emphasizes the worth of triple sampling as a credible vehicle for the approaches to inference outlined in Sections 3. Simulated results in typical situation to investigate the performance of the integrated triple sampling with discussion are also presented in Section 4. Some final remarks and discussion appears in Section 5.

2. The Unknowable Cost of Perfect Information for Point Estimation of the Mean

Let $X_1, X_2, \ldots$ be a sequence of independent and normally distributed random variables with mean μ and variance σ^2, both unknown. Suppose further that a random sample $X_1, X_2, \ldots, X_n$ of size ($n \geq 2$) is available, from which we propose the usual sample mean $\overline{X}_n$ to estimate μ. Moreover, we assume that the incurred loss (or cost) can be reasonably represented by

$$L_n = A\left(\overline{X}_n - \mu\right)^2 + cn, \tag{2.1}$$

where the constant $A(>0)$. We elaborate more on the interpretation of A shortly. The risk (or expected cost) associated with (2.1) is given by

$$R_n = E(L_n) = AE\left(\overline{X}_n - \mu\right)^2 + cn \tag{2.2}$$

$$= A\left(\sigma^2/n\right) + cn. \tag{2.3}$$

The optimal sample size that minimizes the expected cost in (2.3) is given by

$$n^* = \left(\sqrt{A/c}\right)\sigma \tag{2.4}$$

from which we determine that from which we determine that

$$A = cn^*\, I\left(n^*, \sigma^2\right). \tag{2.5}$$

The term cn^* in (2.5) is the *cost of optimal sampling* and the term $I(n^*, \sigma^2) = n^*/\sigma^2$ is the *perfect information* (in Fisher's sense), PI. By this designation we mean that PI represents the *optimal information, i.e., the amount of information required to explore a unit of variance in order to achieve the minimum expected cost (risk).* Hence, A is the *cost of perfect information*, C_{PI}, *i.e., the monetary amount that needs to be paid to achieve the minimum risk*, contrarily to what has been said in the literature regarding A as the known cost of estimation. See for example Robbins (1959), Starr (1966), Starr and Woodroofe (1968), Woodroofe (1977), Ghosh and Mukhopahyay (1979, 1980), Chow and Martinsek (1980), Chow and Yu (1981), Mukhopadhyay (1985), Woodroofe (1985, 1987), Mukhopadhyay *et al.* (1987), Hamdy *et al.* (1988), Martinsek (1988), Almahmeed and Hamdy (1990) and Hamdy *et al.* (1996), the interpretation of A in (2.5) does serve to emphasize that A reflects *both* estimation error and sampling cost.

Direct substitution of A in (2.5) into the loss function in (2.1) yields

$$L_n = cn^* I\left(n^*, \sigma^2\right)\left(\overline{X}_n - \mu\right)^2 + cn \tag{2.6}$$

and the risk in (2.3) becomes

$$R_n = E(L_n) = cn^* I\left(n^*, \sigma^2\right) E\left(\overline{X}_n - \mu\right)^2 + cn. \tag{2.7}$$

Now suppose that the random sample used to obtain $\overline{X}_n$ in (2.7) is of *arbitrary* size n. It follows that $E(\overline{X}_n - \mu)^2 = I^{-1}(n, \sigma^2)$, where $I(n, \sigma^2)$ is the (Fisher) *sample information*, SI_n. The corresponding risk in (2.7) can be expressed as

$$R_n = cn^*\left(PI/SI_n\right) + cn. \tag{2.8}$$

The quantity (PI/SI_n) is the *relative information*, RI_n. In words, RI_n is a *measure of the estimation* (but not sampling) *efficiency of the sample information compared to the perfect information.* When $n = n^*$, see Figure 1 for example where $n^* = 20$, A=\$4.01, c=\$1 and $\sigma^2 = 100$, the sample information is *as efficient* as the perfect information (with the term efficient we mean: the ratio of the amount of sampling information compared to the amount of perfect information) since $RI_n = 1$ and $R_{n^*} = 2cn^*$, which is the optimal risk (minimum expected cost). On the other hand, when $n < n^*$, the sample information is *less efficient* than the perfect information since $RI_n > 1$. And when $n > n^*$, the sample information is *more efficient* than the perfect information since $RI_n < 1$. For any $n \neq n^*$, we have $R_n > R_{n^*}$. Apparently although (anomaly) that $R_n > R_{n^*}$ when $n > n^*$, the corresponding sample information is still considered to be more efficient than the perfect information. This is due to the fact that over sampling is good for estimation, *per se*, but bad in terms of sampling costs. In the sense that the sample information represents an estimate of the perfect information, SI_n is also termed the *imperfect information*, II_n.

Another way to compare the values of the sample II_n and the PI is through the *expected cost of missed opportunity*, which we express as

$$R_{MISS} = (R_n - R_{n^*}) = c\{n^*(RI_n - 1) + (n - n^*)\}. \qquad (2.9)$$

[Other names for R_{MISS} (up to sign) appearing in the literature include the "regret" and the "opportunity cost". However, since it is true that you can't regret what you never have, and false that missing the mark is an "opportunity", our terminology seems preferable.] This measure is conceptually useful since it clearly indicates the separate effects of potential estimation and sampling deficiencies for both under sampling and over sampling.

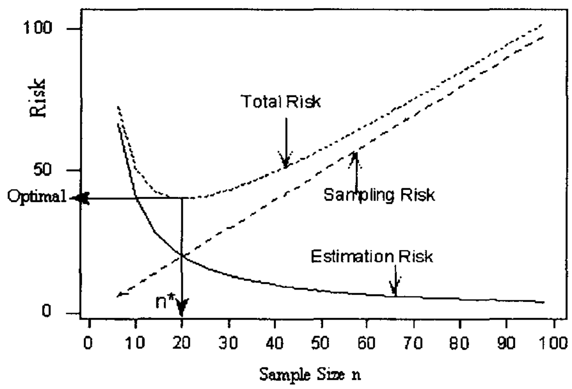

FIGURE 1

Most of the work done in the area of multistage point estimation of the normal mean have involved the assumption of *complete knowledge* of the constant A. As we have shown in (2.5) for the case of squared error loss in (2.1), such claims tacitly imply that the decision maker *knows* the value of the ratio of the square of the optimal sample size to the population variance or, in our present terminology, that the decision maker *knows* the cost of perfect information. Is this a viable assumption in the context of obtaining a point estimator for μ via multistage sampling, the basic premise of which is that σ^2 is completely unknown? For that matter, is it a viable assumption even in fixed sample size point estimation applications where the decision maker's experience with previous data leads him/her to believe that σ^2 is (at least approximately) "known"? In either case, and regardless of whether multistage or fixed size sampling is actually used, it is logically apparent that the assumption of complete (or even approximate) knowledge of the value of A in a "purely" point estimation problem begs the question *with or without* knowledge of σ^2. Thus, the practice of pre-specifying the value of A in point estimation applications inevitably leads to "spuriously optimal" sample sizes. This is because their determination via (2.4) is based on a self-fulfilling criterion, which is unrelated to, and thereby artificially, and mysteriously restricts the population of inference. In effect, the sampled population is not the target population that needs to be explored freely through the sampling procedure in the first place. In particular, if the resulting sample information is less than the perfect information, under sampling will likely occur and the corresponding estimator will have larger variance compared to an estimator based on the true [but underestimated (or under explored)] optimal sample size. While if the resulting sample information is greater than the perfect information, over sampling will likely occur and the sampling costs will be larger compared to what they would have been had the true [but overestimated (or over explored)] optimal sample size been employed.

In the next section, we elaborate further on both the specific consequences of arbitrarily specifying A in the contexts of fixed width confidence interval estimators and tests of hypotheses. Further a *knowable portion* of A *can* be specified to satisfy practically and statistically meaningful target requirements within those inferential frameworks.

3. The C_{PI} for Both Confidence Interval and Tests for the Normal Mean

In constructing a fixed $2d$ width confidence interval for the normal mean μ with unknown variance σ^2 such that the coverage probability is at least $100(1-\alpha)\%$, it is natural to employ the interval $I_n = (\overline{X}_n - d,\, \overline{X}_n + d)$. To satisfy the requirements that

$$P(\mu \in I_n) = P\left(-d \leq \overline{X}_n - \mu \leq d\right) \geq 1 - \alpha$$

which implies that $2\Phi(\sqrt{n}d/\sigma) - 1 \geq 2\Phi(a) - 1$, where $\Phi(\cdot)$ is the cumulative distribution function and a is the $100(1-\alpha/2)^{th}$ percentile of the standard normal distribution. Hence, the corresponding optimal sample size is

$$n^* = (a/d)^2 \sigma^2. \tag{3.1}$$

Comparing (3.1) with (2.4) and (2.5) provides the cost of perfect information as

$$C_{PI} = c\,(a/d)^4 \sigma^2 \tag{3.2}$$

and the corresponding perfect information is

$$PI = I\left(n^*, \sigma^2\right) = (a/d)^2. \tag{3.3}$$

We see immediately that the perfect information in (3.3) for the point estimation of the normal mean in the context of the fixed width confidence interval is completely knowable and it reflects both the influence of the target population and the required precision through the constants a and d.

Furthermore, assume that we need to test the null hypothesis $H_{0,n} : \mu \in I_n$ versus the alternative hypothesis $H_{a,n} : \mu \notin I_n$. To develop theory for testing the hypothesis using the already constructed confidence interval we assume that $H_{0,n} : \mu = \mu_0 \in I_n$ versus the alternative hypothesis $H_{a,n} : \mu = \mu_0 \pm d(1+k) \notin I_n$ for all $k > 0$. Under the above decision rule, $H_{0,n}$ is rejected only if μ_0 lies outside the interval. To develop theory, we assume that the specified value of μ_0 is located in the center of the interval to provide equal probability of type II error for equidistant shifts in the mean occurring outside the interval. See for example Son *et al.* (1997). We also emphasize that under the null hypothesis all values of μ inside the interval equally represent the true value of μ. Consequently, the alternative hypothesis only makes sense for shifts occurring outside the interval I_n. Here our main objective is to control type II error, see for example Son *et* *al.* (1997), Costanza *et al.* (1995) and Hamdy (1997), therefore we impose the restriction that the probability of committing such an error when the true mean is shifted away by an amount $\pm d(1+k)$ from μ_0 for a given k is at most $\beta(k)$. It follows that the optimal sample size required to ensure the desired nominal probability and at the same time provide protection against type II error (see Kupper and Hafner, 1989) is given by

$$n^* = \{(a+b)/d(1+k)\}^2 \sigma^2, \tag{3.4}$$

where b is the $100(1-\beta(k))^{th}$ percentile of the standard normal distribution. It is obvious in (3.4) that for k close to 0, we would expect a coverage probability greater than the required nominal value. By setting $k=0$ in (3.4) to obtain an upper bound for the optimal sample size given by

$$n^* = \{(a+b)/d\}^2 \sigma^2. \tag{3.5}$$

It follows that the corresponding cost of perfect information is

$$C_{PI} = c\{(a+b)/d\}^4 \sigma^2 \tag{3.6}$$

and the corresponding perfect information is

$$PI = \{(a+b)/d\}^2. \tag{3.7}$$

Consequently, incorporating the cost of perfect information in the loss function in (2.1) would serve several objectives. It still provides $2d$ fixed width confidence interval for the mean μ with coverage probability at least the nominal value. Moreover, it controls the probability of type II error as required. Finally, it also provides a point estimator for the mean μ with minimum expected cost.

However, the optimal sample size in (3.5) required to achieve the above objectives (of having a confidence interval of at least the nominal value and at the same time controlling for type II error) is inherently unknown, for σ^2 is essentially unknown. Therefore we resort to multistage methods of sampling techniques to estimate σ^2 via estimation of n^* and in sequel propose the required point and interval estimation for μ. The triple sampling technique provides a unified sampling framework within which all objectives can be achieved. Unlike Stein (1945) and Cox (1952) two-stage sampling procedures, triple sampling technique was developed to overcome the problem of oversampling, especially when the initial sample size is chosen much less than the optimal sample size. On the other hand, triple sampling enjoys the asymptotic efficiency of Anscombe (1953), Robbins

(1959) and Chow and Robbins (1965) one-by-one sequential sampling. Therefore, sampling in three stages was engineered to serve two main purposes, to achieve operational savings made possible by sampling in batches by utilizing Stein (1945) and Cox (1952) two-stage sampling, and at the same time maintain the asymptotic efficiency of one-by-one sequential sampling. In the following section we mimic n^* in (3.5) and describe the triple sampling technique along the lines of Hall (1981, 1983), Hamdy (1988) and Woodroofe (1987).

4. An Integrated Triple Sampling Procedure for Statistical Inference for μ

The triple sampling procedure begins with a fixed sample size $X_1, X_2, \dots, X_m$, $m \geq 2$, from which we propose $\overline{X}_m$ and S_m^2 as initial estimates for both μ and σ^2 respectively. The second stage sample size is then decided for, according to the rule

$$N_1 = \max\left\{m, \left[\gamma(a+b)^2 S_m^2/d^2\right]\right\}, \tag{4.1}$$

where $0 < \gamma < 1$ is the design factor which represents the fraction of $n^*(\gamma n^*)$ to be estimated in this stage and $[\cdot]$ is the integer function. The design factor γ was introduced by Hall (1981) to reduce the possibility of over-sampling during the second stage. If the decision is to continue sampling, a second sample of size $N_1 - m$ is to be taken and augmented with the first sample to bring the final stage sample size to

$$N = \max\left\{N_1, [(a+b)^2 S_{N_1}^2/d^2] + 1\right\}. \tag{4.2}$$

If necessary, we continue to observe $N - N_1$ further observations and terminate the sampling process. Upon realization of N, we compute $\overline{X}_N$ as the point estimate for μ. Accordingly, we construct the type II controlled triple sampling interval $I_N = (\overline{X}_N - d, \overline{X}_N + d)$ for μ. The following Theorem 4.1 provides the large sample feature of the triple sample size N.

Theorem 4.1. *Let $g(>0)$ be continuously twice differentiable function in a neighborhood of n^* such that $\sup_{n\geq m}|g'''(n)| = O(|g'''(n^*)|)$, then*

$$E(g(N)) = g(n^*) + (2\gamma)^{-1}\left\{(\gamma-4)g'(n^*) + 2n^*g''(n^*)\right\} + o\left(d^{-4}\left|g'''(n^*)\right|\right).$$

The expectation of Taylor series expansion of $g(N)$ around n^* and the asymptotic results of Hall (1981) and Hamdy (1988) that $E(N) = n^* - 2\gamma^{-1} + 0.5 + o(1)$,

$\text{Var}(N) = 2\gamma^{-1}n^* + o(d^{-2})$ and $E|N - n^*|^3 = o(d^{-4})$ as well as the assumption that g''' is bounded, provide the statement of Theorem 4.1.

Also, conditioning on N and write $E(\overline{X}_N) = E\{E(\overline{X}_N|N = n)\}$ and the fact that the event $\{N = n\}$ and $\overline{X}_n$ are independent for all $n = m, m+1, \ldots,$ it is easily shown that $E(\overline{X}_N) = \mu$, which proves that the triple sampling mean $\overline{X}_N$ is unbiased for the population mean μ. Again, we condition on N and write $\text{Var}(\overline{X}_N)$ as

$$\begin{aligned}\text{Var}(\overline{X}_N) &= E\left\{\text{Var}(\overline{X}_N \mid N = n)\right\} + \text{Var}\left\{E\left(\overline{X}_N \mid N = n\right)\right\} \\ &= \mu^2\text{Var}(N^{-1}) + \sigma^2 E(N^{-1}).\end{aligned}$$

Now, by applying Theorem 4.1 and ignoring terms of order higher than $o(d^4)$, it can be proved that $Var(\overline{X}_N) = (\sigma^2/n^*) - (\gamma - 8)(\sigma^2/2\gamma n^{*2}) + o(d^4)$. Moreover, it is not hard to show that the distribution of $Z_N = (N - n^*)/\sqrt{2\gamma^{-1}n^*}$ is asymptotically N(0,1) via the moment generating function $E(e^{tZ_N})$ and Theorem 4.1.

Theorem 4.2. *The coverage probability of I_N is given by*

$$P(\mu \in I_N) = \{2\Phi(a+b) - 1\} - Q_0(n^*, \gamma) + o(d^2),$$

where $Q_0(n^*, \gamma) = (a+b)\phi(a+b)\{(a+b)^2 - \gamma + 5\}(2\gamma n^*)^{-1}$, and $\phi(\cdot), \Phi(\cdot)$ are the standardized probability density and cumulative distribution function of the normal distribution, respectively.

The proof of Theorem 4.2 is direct application of Theorem 4.1 since $P(\mu \in I_N) = 2E\{\Phi(d\sqrt{N}/\sigma)\} - 1$. Moreover, it is also evident that

$$P(\mu \in I_N) = \{2\Phi(a+b) - 1\}\, Q_0(n^*, \gamma) + o(d^2) > (1-\alpha) - Q_0(n^*, \gamma) + o(d^2)$$

and for large n^* (as $d \to 0$) the quantity $Q_0(n^*, \gamma) \to 0$ and hence, $P(\mu \in I_N) > (1-\alpha) + o(d^2)$. Simulation results presented in Section 5 support this asymptotic behavior of the confidence interval even for small to moderate values of n^*. The following Theorem 4.3 provides the large sample type II error probability function.

Theorem 4.3. *For the controlled triple sampling procedure (4.1) and (4.2), and for fixed n^* in (3.5) and any $k \geq 0$, the operating characteristic function is given by*

$$\beta(k) = \Phi\left\{-k(a+b)\right\} - \Phi\left\{-(2+k)(a+b)\right\} + Q_1(n^*, \gamma, k) + Q_2(n^*, \gamma, k) + o(d^2),$$

as $d \to 0$, *where*

$$Q_1(n^*, \gamma, k) = (4\gamma n^*)^{-1} k(a+b)\phi(-k(a+b)) \left\{k^2(a+b)^2 - \gamma + 5\right\}$$

and $Q_2(n^*, \gamma, k)$ *is given by*

$$(4\gamma n^*)^{-1}(2+k)(a+b)\phi(-(2+k)(a+b)) \left\{(2+k^2)(a+b)^2 - \gamma + 5\right\}.$$

To prove Theorem 4.3, we write

$$\begin{aligned} \beta(k) &= P\left(\mu \in I_N | H_a\right) = P\left\{\overline{X}_N - d \leq \mu \leq \overline{X}_N + d \,|\, \mu = \mu_0 + d(1+k)\right\} \\ &= E\left\{\Phi\left(-kd\sqrt{N}\sigma^{-1}\right)\right\} - E\left\{\Phi\left(-(2+k)d\sqrt{N}\sigma^{-1}\right)\right\} \end{aligned}$$

and make use of Theorem 4.1, then the statement of Theorem 4.3 is straight forward.

It is clear from Theorem 4.3 that for $k = 0$, we have $\beta(0) = 0.5$. It is also evident that $\beta(k)$ approaches the specific target value of β quickly in the interval $0.4 \leq k \leq 0.5$. We discuss the moderate sample performance of the above integrated triple sampling technique in the following section.

5. Simulation Study

Since the results of Theorems 4.1 to 4.3 are asymptotic in nature, Monte Carlo investigation in typical situations was performed to provide a feel regarding the moderate sample size performance of the integrated triple sampling technique proposed in Section 4. We considered sampling from standard normal distribution where $\mu = 0$ and $\sigma = 1$. We also let the optimal sample size n^* range from small to large (5, 10, 15, 20, 30, 50, 100, 150, 200, 500, 1000) while the design factor γ was fixed to 0.3, 0.5 and 0.8 in all cases. The targeted coverage probability was set to 0.90, 0.95 and 0.99, and the type II error $\beta = 0.05$. We took the starting sample size m to be 5, 10, 15 and 20. However, we only report the cases for $\alpha = 0.05$, $\gamma = 0.5$ and $m = 5$, 10 and 15. Tables (I-1) to (III-2) present the simulation results of the integrated triple sampling. As indicated in all tables, the triple sampling sample size $\overline{N}$ is very close to the optimal sample size n^*. The triple sampling point estimate Avg($\overline{X}$) as shown in all tables provides credible estimate to the mean of the normal distribution $\mu = 0$. It is also obvious that the estimated coverage probability $\widehat{P}$ is strictly greater than the nominal targeted value 0.95. Tables (I-2), (II-2) and (III-2) present the simulated type II errors of the integrated triple sampling procedure. As noted in all tables the estimated type II error $\beta(k)$ is controlled to zero outside the interval $(k \geq 0)$ as expected. However, when the shift (dk) is downward, we expect large values of type II errors inside the interval which decreases as we approach the boundaries, *i.e.*, $k \to 0$. It is also evident that as the pilot sample size m increases all estimated values approach the targeted values. The bias in $\overline{N}$ decreases and the Avg$(\overline{x})$ attains the true value 0 as m increases.

6. Conclusions

We have proposed an integrated triple sampling procedure, which provides a credible point estimate to the mean of the normal distribution. Moreover, it guarantees a fixed width confidence interval for the unknown mean with a prescribed confidence coefficient at least the nominal value. In addition, the constructed integrated triple sampling procedure insures the protection of the interval against type II error. We recommend the use of the proposed procedure in practical implementation. For example, suppose that a quality characteristics in a continuous manufacturing process is normally distributed with mean μ and variance σ^2, both unknown. Suppose also that it is required to estimate μ by a confidence interval such that the precision, $\pm d$, and the coverage probability is at least $100(1-\alpha)\%$, of the interval are determined according to manufacturing specifications. The specific objective might be to reliably establish the center line in $\overline{X}$-chart to be employed in a planned quality monitoring scheme.

Acknowledgements

The authors are grateful to the referees' comments which improved the contents in various places.

TABLE (I-1) *Three stage sampling for the mean of* $N(0,1)$
$\alpha = 0.05$, $m = 5$, $\gamma = 0.5$, *number of simulations* $= 5000$

d	n^*	$\overline{N}$	$S.E.(\overline{N})$	$Avg(\overline{x})$	$S.E.(\overline{x})$	$\widehat{p}$
1.7531	5	6.3070	0.0290	−0.0002	0.4182	1.0000
1.2396	10	9.5030	0.0630	−0.0044	0.3562	0.9986
1.0121	15	13.0190	0.0970	0.0010	0.3177	0.9932
0.8765	20	17.2180	0.1300	−0.0009	0.2919	0.9906
0.7157	30	26.3930	0.1890	0.0007	0.2388	0.9892
0.5544	50	45.8440	0.2770	−0.0002	0.1782	0.9896
0.3920	100	95.2530	0.4330	−0.0009	0.1207	0.9908
0.3201	150	146.8570	0.5670	−0.0012	0.0971	0.9948
0.2772	200	198.1650	0.7020	0.0004	0.0788	0.9958
0.1753	500	507.8800	1.3240	−0.0002	0.0481	0.9992
0.1240	1000	1021.2310	2.1790	−0.0001	0.0318	0.9998

TABLE (I-2) *Simulated values of* β *(type II error) for different values of* k

	Downward Shift k				
n^*	0.1	0.05	0.01	0.001	0
5	0.3284	0.1632	0.0338	0.0040	0.0000
10	0.2954	0.1490	0.0294	0.0036	0.0000
15	0.2600	0.1302	0.0252	0.0018	0.0000
20	0.2684	0.1378	0.0288	0.0028	0.0000
30	0.2740	0.1346	0.0260	0.0030	0.0000
50	0.2738	0.1414	0.0246	0.0024	0.0000
100	0.2904	0.1486	0.0278	0.0030	0.0000
150	0.3000	0.1522	0.0258	0.0034	0.0000
200	0.3000	0.1490	0.0306	0.0036	0.0000
500	0.2970	0.0498	0.0272	0.0024	0.0000
1000	0.3084	0.1602	0.0308	0.0022	0.0000

NOTE : *The estimated type II errors* $\beta(k)$*'s are all zero when* $k > 0$.

TABLE (II-1) *Three stage sampling for the mean of* $N(0,1)$
$\alpha = 0.05,\ m = 10,\ \gamma = 0.5,$ *number of simulations* $= 5000$

d	n^*	$\overline{N}$	$S.E.(\overline{N})$	$Avg(\overline{x})$	$S.E.(\overline{x})$	$\widehat{p}$
1.7631	5	10.0810	0.0070	0.0078	0.3171	1.0000
1.2396	10	11.9340	0.0430	−0.0018	0.2996	1.0000
1.0121	15	15.4190	0.0720	−0.0057	0.2671	0.9996
0.8765	20	19.3390	0.0970	−0.0009	0.2467	0.9990
0.7157	30	27.3170	0.1450	0.0007	0.2092	0.9974
0.5544	50	46.0080	0.2240	−0.0001	0.1616	0.9964
0.3920	100	95.4930	0.3430	−0.0011	0.1085	0.9980
0.3201	150	146.4460	0.4140	0.0007	0.0846	0.9994
0.2772	200	196.5800	0.4810	0.0010	0.0727	0.9992
0.1753	500	498.3810	0.7830	0.0007	0.0455	0.9994
0.1240	1000	1002.2590	1.2240	−0.0003	0.0314	0.9998

TABLE (II-2) *Simulated values of* β *(type II error) for different values of* k

	Downward Shift k				
n^*	0.1	0.05	0.01	0.001	0
5	0.4250	0.2168	0.0406	0.0042	0.0000
10	0.3270	0.1680	0.0342	0.0024	0.0000
15	0.3048	0.1592	0.0324	0.0026	0.0000
20	0.2922	0.1512	0.0276	0.0028	0.0000
30	0.2834	0.1442	0.0286	0.0030	0.0000
50	0.2850	0.1452	0.0284	0.0042	0.0000
100	0.3082	0.1522	0.0332	0.0032	0.0000
150	0.3006	0.1486	0.0280	0.0032	0.0000
200	0.2982	0.1534	0.0272	0.0028	0.0000
500	0.3108	0.1556	0.0328	0.0042	0.0000
1000	0.3084	0.1568	0.0314	0.0028	0.0000

NOTE : *The estimated type II errors* $\beta(k)$*'s are all zero when* $k > 0$.

TABLE (III-1) *Three stage sampling for the mean of* $N(0,1)$
$\alpha = 0.05$, $m = 15$, $\gamma = 0.5$, *number of simulations* $= 5000$

d	n^*	$\overline{N}$	$S.E.(\overline{N})$	$Avg(\overline{x})$	$S.E.(\overline{x})$	$\widehat{p}$
1.7531	5	15.0000	0.0000	−0.0020	0.2607	1.0000
1.2396	10	15.3020	0.0160	0.0020	0.2587	1.0000
1.0121	15	17.4530	0.0530	0.0048	0.2424	1.0000
0.8765	20	20.8410	0.0830	−0.0022	0.2244	0.9996
0.7157	30	29.1100	0.1260	−0.0049	0.1900	0.9996
0.5544	50	46.3890	0.2010	−0.0003	0.1558	0.9988
0.3920	100	95.5150	0.3130	−0.0021	0.1068	0.9992
0.3201	150	145.5490	0.3840	0.0004	0.0838	0.9998
0.2772	200	196.0420	0.4410	0.0009	0.0737	0.9998
0.1753	500	495.9750	0.6910	0.0000	0.0451	0.9996
0.1240	1000	996.7490	0.9950	0.0003	0.0315	1.0000

TABLE (III-2) *Simulated values of* β *(type II error) for different values of* k

	Downward Shift k				
n^*	0.1	0.05	0.01	0.001	0
5	0.4902	0.2572	0.0508	0.0046	0.0000
10	0.3604	0.1848	0.0372	0.0030	0.0000
15	0.3200	0.1598	0.0328	0.0032	0.0000
20	0.3090	0.1642	0.0332	0.0024	0.0000
30	0.3030	0.1528	0.0288	0.0030	0.0000
50	0.2770	0.1412	0.0306	0.0026	0.0000
100	0.3044	0.1512	0.0274	0.0030	0.0000
150	0.3048	0.1508	0.0304	0.0044	0.0000
200	0.2848	0.1474	0.0300	0.0030	0.0000
500	0.3034	0.1568	0.0308	0.0026	0.0000
1000	0.3068	0.1612	0.0360	0.0032	0.0000

NOTE : *The estimated type II errors* $\beta(k)$*'s are all zero when* $k > 0$.

REFERENCES

Almahmeed, M. A. and Hamdy, H. I. (1990). "Sequential estimation of linear models in three stages", *Metrika*, **37**, 19–36.

Anscombe, F. J. (1953). "Sequential estimation", *Journal of the Royal Statistical Society*, **B15**, 1–29.

Chow, Y. S. and Martinsek, A. T. (1980). "Bounded regret of a sequential procedure for estimation of the mean", *The Annals of Statistics*, **10**, 909–914.

Chow, Y. S. and Robbins, H. (1965). "On the asymptotic theory of fixed width sequential confidence intervals for the mean", *The Annals of Mathematical Statistics*, **36**, 457–462.

Chow, Y. S. and Yu, K. F. (1981). "The performance of a sequential procedure for the estimation of the mean", *The Annals of Statistics*, **9**, 184–189.

Costanza, M. C., Hamdy, H. I., Haugh, L. D. and Son, M. S. (1995). "Type II error performance of triple sampling fixed precision confidence intervals for the normal mean", *Metron*, **31**, 69–82.

Cox, D. R. (1952). "Estimation by double sampling", *Biometrika*, **39**, 217–227.

Ghosh, M. and Mukhopadhyay, N. (1979). "Sequential point estimation of the mean when the distribution is unspecified", *Communications in Statistics*, **A8**, 637–652.

Ghosh, M. and Mukhopadhyay, N. (1980). "Sequential point estimation of the difference of two normal means", *The Annals of Statistics*, **8**, 221–225.

Hall, P. (1981). "Asymptotic theory of triple sampling for sequential estimation of a mean", *The Annals of Statistics*, **9**, 1229–1238.

Hall, P. (1983). "Sequential estimation saving sampling operations", *Journal of the Royal Statistical Society*, **B45**, 219–223.

Hamdy, H. I. (1988). "Remarks on the asymptotic theory of triple sampling estimation of the normal mean", *Scandinavian Journal of Statistics*, **15**, 303–310.

Hamdy, H. I. (1997). "Performance of fixed width confidence intervals under type II errors: The exponential case", *South African Statistical Journal*, **31**, 259–269.

Hamdy, H. I., Al-Mahmeed, M. and Al-Zalzalah, Y. (1996). "A certain accelerated sequential procedure to construct simultaneous confidence region : The exponential case", *Statistics*, **28**, 57–71.

Hamdy, H. I., Mukhopadhyay, N., Costanza, M. C. and Son, M. S. (1988). "Triple stage point estimation for the exponential location parameter", *Annals of the Institute of Statistical Mathematics*, **40**, 785–797.

Kupper, L. L. and Hafner, K. B. (1989). "How appropriate are popular sample size formulas?", *The American Statistician*, **43**, 101–105.

Martinsek, A. T. (1988). "Negative regret, optional stopping and the elimination of outliers", *Journal of the American Statistical Association*, **83**, 160–163.

Mukhopadhyay, N. (1985). "A note on three-stage and sequential point estimation procedures for a normal mean", *Sequential Analysis*, **4**, 311–319.

Mukhopadhyay, N., Hamdy, H. I., Al-Mahmeed, M. and Costanza, M. C. (1987). "Three-stage point estimation procedures for a normal mean", *Sequential Analysis*, **6**, 21–36.

Robbins, H. (1959). "Sequential estimation of the mean of a normal population", *Probability & Statistics-The Harald Cramer Volume*, Almqvist and Wiksell, Uppsala, 235–245.

Son, M. S., Haugh, L.D., Hamdy, H. I. and Costanza, M. C. (1997). "Controlling type II error while constructing triple sampling fixed precision confidence intervals for the normal mean", *Annals of the Institute of Statistical Mathematics*, **49**, 681–692.

Starr, N. (1966). "On the asymptotic efficiency of a sequential procedure for estimating the mean", *The Annals of Mathematical Statistics*, **37**, 1173–1185.

Starr, N. and Woodroofe, M. B. (1968). "Remarks on a sequential point estimation", *Proceedings of the National Academy of Sciences of the United States of America*, **63**, 285–288.

Stein, C. (1945). "A two stage test for a linear hypothesis whose power is independent of the variance", *The Annals of Mathematical Statistics*, **17**, 243–258.

Woodroofe, M. (1977). "Second order approximation for sequential point and interval estimation", *The Annals of Statistics*, **5**, 984–995.

Woodroofe, M. (1985). "Asymptotic local minimaxity in sequential point estimation", *The Annals of Statistics*, **7**, 676–688.

Woodroofe, M. (1987). "Asymptotically optimal sequential point estimation in three stages", In *New Perspectives in Theoretical and Applied Statistics* (M. L. Puri, J. P. Vilaplana and W. Wertz, eds.), Wiley, New York, 397–411.