Ann. Inst. Statist. Math.
Vol. 41, No. 1, 121-138 (1989)

STOPPING RULES, PERMUTATION INVARIANCE AND SUFFICIENCY PRINCIPLE*

NITIS MUKHOPADHYAY[1], PRANAB KUMAR SEN[2] AND BIKAS KUMAR SINHA[3]

[1]*Department of Statistics, Box U-120, 196 Auditorium Road, The University of Connecticut, Storrs, CT 06269, U.S.A.*
[2]*Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599, U.S.A.*
[3]*Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, 203 B. T. Road, Calcutta 700035, India*

(Received December 16, 1987; revised July 28, 1988)

Abstract. In the context of sequential (point as well as interval) estimation, a general formulation of permutation-invariant stopping rules is considered. These stopping rules lead to savings in the ASN at the cost of some elevation of the associated risk—a phenomenon which may be attributed to the violation of the sufficiency principle. For the (point and interval) sequential estimation of the mean of a normal distribution, it is shown that such permutation-invariant stopping rules may lead to a substantial saving in the ASN with only a small increase in the associated risk.

Key words and phrases: Permutation-invariant stopping rules, average sample numbers, percentage savings, sequential point estimation, fixed-width confidence interval, normal mean, unknown variance.

1. Introduction

Let $\{X_1, X_2, \dots\}$ be a sequence of independent and identically distributed random variables (i.i.d.r.v.) with a distribution function (d.f.) F. For every n (≥ 1), let $T_n = T(n; X_1, \dots, X_n)$ be a nonnegative statistic, and let $\{b_\nu: \nu = 1, 2, \dots\}$ be a nondecreasing sequence of real, positive numbers. Consider a general stopping variable

$$\tau_\nu = \inf\{n \geq m: n \geq b_\nu T_n\}, \tag{1.1}$$

where m is a preassigned positive integer (which may even depend on ν), $\nu = 1, 2, \dots$. It may be mentioned that the sequential estimation rules

*Work partially supported by (i) Office of Naval Research, Contract Number N00014-85-K-0548, and (ii) Office of Naval Research, Contract Number N00014-83-K-0387.

studied by Ray (1957), Robbins (1959), Chow and Robbins (1965), Ghosh and Mukhopadhyay (1975, 1976), Woodroofe (1982) and others may be unified in the form of (1.1).

We may notice that in (1.1), at the n-th stage, $b_\nu T(n; X_1,\dots, X_n)$ is compared with n, where $X_1,\dots, X_n$ have been observed in that order. In a general setup, T_n is a symmetric function of $X_1,\dots, X_n$, and, in a non-sequential setup, $X_1,\dots, X_n$ are *permutationally invariant* (PI) in the sense that their joint distribution remains invariant under any permutation of the n arguments. However, in a general sequential setup, for an arbitrary stopping rule τ_ν, $X_1,\dots, X_{\tau_\nu}$ need not be permutationally invariant. Thus, a natural question may arise: Could we have stopped earlier if the same set of $X_1,\dots, X_{\tau_\nu}$ had arrived in a possibly different order? This motivates us towards the formulation of *permutation-invariant stopping rules* (PISR), and we shall consider this concept in detail in Section 2.

There is an intricate relationship between *optimal stopping* rules and (*transitive*) *sufficiency* (see Bahadur (1954)). As we shall see in Section 2, by construction of the PISR, this (transitive) sufficiency principle is violated. Hence, a PISR may not share the optimality properties. Nevertheless, it will be shown that such PISR may lead to substantial savings in the ASN and there may not be any significant increase in the associated risk of the (sequential) estimators. To emphasize this vital aspect of PISR, in Sections 3 and 4, we will incorporate them in the case of sequential point and interval estimation problems for a normal mean (when the variance is also unknown), and show that the PISR compare very favorably with their classical (noninvariant) counterparts. In passing, we may remark that the above picture is largely asymptotic in nature, and there is a good scope for in-depth numerical studies in the non-asymptotic case which would be explored elsewhere.

2. PISR: General formulation

Note that we have taken (for $n \geq m$) $T_n = T(n; X_1,\dots, X_n)$. This is done merely to include the more general situation where

$$(2.1) \qquad T_n = T_n^* + h_n; \quad T_n^* = T^*(X_1,\dots, X_n), \quad n \geq m,$$

and $\{h_n\colon n \geq m\}$ is a (nonincreasing) sequence of real nonnegative numbers with $\lim_{n\to\infty} h_n = 0$. We shall see in later sections that (2.1) covers a more general setup than the simple case where $h_n = 0$ for all $n \geq m$.

For every n (≥ 1), let $\mathscr{P}_n$ be the set of $n!$ permutations $\{i_1,\dots, i_n\}$ of the first n natural integers $(1,\dots, n)$. For every k $(m \leq k \leq n)$ and $(i_1,\dots, i_k) \in \{1,\dots, n\}$, we define

(2.2) $$T_{n,(i_1,\ldots,i_k)} = T(k; X_{i_1},\ldots, X_{i_k}) \; .$$

Then, looking at (1.1), we may consider the following PISR:

(2.3) At the n-th stage ($n \geq m$), for each k ($m \leq k \leq n$) and every $(i_1,\ldots, i_k) \in \{1,\ldots, n\}$, we compute $T_{n,(i_1,\ldots,i_k)}$. If, for some k ($m \leq k \leq n$) and some $(i_1,\ldots, i_k)$, $k \geq b_\nu T_{n,(i_1,\ldots,i_k)}$, then we stop sampling at the n-th stage; otherwise, proceed to the next stage by taking one more observation. The associated stopping variable is denoted by τ_ν^*, $\nu = 1, 2,\ldots$.

If we write, for each k ($m \leq k \leq n$),

(2.4) $$T_{nk}^0 = \min \{T_{n,(i_1,\ldots,i_k)}\colon 1 \leq i_1 \neq \cdots \neq i_k \leq n\} \; ,$$

then, we may also write τ_ν^* equivalently as

(2.5) $$\tau_\nu^* = \inf \{n \geq m\colon k \geq b_\nu T_{nk}^0, \text{ for some } k\colon m \leq k \leq n\} \; ,$$

$\nu = 1, 2,\ldots$. Note that by (2.1), (2.2) and (2.4), for every k ($m \leq k \leq n$),

(2.6) $$T_{nk}^0 \leq T_k = T(k; X_1,\ldots, X_k) \quad \text{w.p. } 1 \; ,$$

and hence, by (1.1) and (2.5), we have

(2.7) $$\tau_\nu^* \leq \tau_\nu \quad \text{w.p. 1, for every} \quad \nu = 1, 2,\ldots \; .$$

It is also clear from (2.3)–(2.5) that τ_ν^* remains invariant under any permutation of the indices $i_1,\ldots, i_{\tau_\nu^*}$ (i.e., the order in which the X_i's enter into the (stopped) sample), while for τ_ν this invariance may not generally hold. Thus, τ_ν^* is a PISR, while τ_ν may not be PI.

For each n (≥ 1), let $\mathscr{X}^{(n)}$ be the sample space of $(X_1,\ldots, X_n)$, $\mathscr{B}^{(n)}$ the Borel sigma-field on $\mathscr{X}^{(n)}$, and let $\mathscr{P}^{(n)}$ be the family of probability measures on $(\mathscr{X}^{(n)}, \mathscr{B}^{(n)})$ which are assumed to be dominated by some sigma-finite measure. Let $\mathscr{B}_0^{(n)}$ be the sigma-subfield generated by T_n, $n \geq m$. Then, $\{T_n\colon n \geq m\}$ is a *transitive sequence* for the sequential model $\{(\mathscr{X}^{(n)}, \mathscr{B}^{(n)}, \mathscr{P}^{(n)}); n \geq m\}$ if for every $n \geq m$, any version of the conditional distribution of T_n given $(X_1,\ldots, X_{n-1})$ depends only on T_{n-1}. Actually, Wijsman's (1959) theorem on transitive sufficiency asserts that for $\mathscr{B}_0^{(n)} \subset \mathscr{B}^{(n)}$, for all $n \geq m$, the sequence $\{\mathscr{B}_0^{(n)}; n \geq m\}$ is transitive for $\{\mathscr{B}^{(n)}; n \geq m\}$ *if and only if* $\mathscr{B}^{(n)}$ and $\mathscr{B}_0^{(n+1)}$ are conditionally independent given $\mathscr{B}_0^{(n)}$. Bahadur (1954) has shown that in sequential decision problems, attention can be confined to procedures based on transitively sufficient sequences of statistics. Thus, whenever $\{T_n; n \geq m\}$ is such a transitive sequence, τ_ν may have some optimality properties, and it is counter-intuitive to consider the triangular scheme $\{T^0_{nk}, k \le n; n \ge m\}$ and the associated τ^*_ν. On the other hand, by (2.7), whenever the ASN (i.e., $E(\tau_\nu)$) exists, we have

$$E(\tau^*_\nu) \le E(\tau_\nu)\,, \tag{2.8}$$

so that τ^*_ν is more desirable than τ_ν. This apparent anomaly can be easily rectified by considering the optimality properties of the sequential estimation procedures based on the stopping rules in (1.1) and (2.5), respectively. In this context, we shall see that the violation of the sufficiency principle by τ^*_ν generally results in an elevated risk for the corresponding (sequential) estimator, and this is quite in line with Bahadur's (1954) basic result. Granted this explanation from the theoretical point of view, the natural question may arise: What would be the cost for the excess risk in compensation for the gain through reduction in the ASN? In other words, can some "near optimality" properties be achieved by such PISR? In the next two sections, we shall attempt to provide satisfactory answers to these questions with special reference to the sequential estimation of the normal mean problems. In this setup, we will mostly confine ourselves to the asymptotic case where "b_ν" is taken to be large (as has been done in Robbins (1959), Chow and Robbins (1965) and other places), and we believe that more favorable results can be obtained in the "non-asymptotic" case through extensive numerical work. We do report on some limited small sample studies; however, these are not very conclusive even though the overall picture appears to be extremely encouraging.

3. Minimum risk point estimation of the normal mean and PISR

Having recorded $X_1, \ldots, X_n$, suppose that the *loss function* in estimating the (unknown) mean μ by $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ is given by

$$L_n = (\bar{X}_n - \mu)^2 + cn\,, \tag{3.1}$$

where $c\ (>0)$ is the known cost per unit sample. If the variance σ^2 were known, the *risk* $E(L_n) = \sigma^2 n^{-1} + cn$ is minimized when $n = n^0_c \sim \sigma c^{-1/2}$, and ρ^0_c, the associated *minimum risk*, is $\sim 2\sigma c^{1/2}$ as $c \downarrow 0$; here $a \sim b$ means $a/b \to 1$. Since σ is, in fact, unknown, no fixed sample size procedure would minimize $E(L_n)$ uniformly in σ. Let $S_n^2 = (n-1)^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ for $n \ge 2$. Following Robbins (1959) and others, we may then consider the *stopping rule*:

$$N_c = \inf\{n \ge m\ (\ge 2) \colon n \ge c^{-1/2} S_n\}\,, \tag{3.2}$$

and estimate μ by $\bar{X}_{N_c}$. Note that the risk associated with $\bar{X}_{N_c}$ is given by

$$\sigma^2 E(N_c^{-1}) + cE(N_c) = \rho_c, \quad \text{say}. \tag{3.3}$$

Up to various orders of approximations (see Starr (1966*b*), Ghosh and Mukhopadhyay (1981) and Woodroofe (1982)), it has been shown that as $c \downarrow 0$,

$$E(N_c) \sim n_c^0 \quad \text{and} \quad \rho_c \sim 2\sigma c^{1/2}, \tag{3.4}$$

so that the sequential estimator $\bar{X}_{N_c}$ has asymptotically the minimum risk ρ_c^0 under a variety of conditions on m. Note that the stopping rule (3.2) is of the form (1.1) with $\nu = [c^{-1}]^* + 1$, $b_\nu = c^{-1/2}$ and $T_n = T_n^* = S_n$ for $n \geq 2$ where $[x]^*$ stands for the largest integer $< x$. Also, we may note that for every $n \geq 2$,

$$S_n^2 = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \phi(X_i, X_j); \qquad \phi(a, b) = \frac{1}{2}(a - b)^2. \tag{3.5}$$

Further, in dealing with possibly nonnormal d.f.'s (see Ghosh and Mukhopadhyay (1979)), one may modify (3.2) and consider $N_c = \inf\{n \geq m (\geq 2): n \geq c^{-1/2}(S_n + n^{-a})\}$ for some $0 < a < 1$, and this would correspond to $T_n = T_n^* + n^{-a}$ with $T_n^* = S_n$.

To introduce the PISR, we define $S^2_{n,(i_1,\ldots,i_k)}$ as in (2.2) and let

$$Z_{n,k} = \min\{S^2_{n,(i_1,\ldots,i_k)}: 1 \leq i_1 < \cdots < i_k \leq n\}, \tag{3.6}$$

for $k = 2, \ldots, n$. Then, the stopping variable N_c^* of the PISR is of the form

$$N_c^* = \inf\{n \geq m(\geq 2): k \geq c^{-1/2}(Z_{n,k}^{1/2} + k^{-a}), \quad \text{for some } k: m \leq k \leq n\}. \tag{3.7}$$

We estimate μ by $\bar{X}_{N_c^*}$ and denote the risk of this sequential estimator by ρ_c^*. Our main contention is to compare $E(N_c)$ and ρ_c with $E(N_c^*)$ and ρ_c^*, respectively.

Let $\mathscr{C}_n$ be the sigma-field generated by the tail sequence $\{S_l^2; l \geq n\}$, for $n \geq 2$, so that $\mathscr{C}_n$ is nonincreasing in n. Then, $\{S_n^2, \mathscr{C}_n; n \geq 2\}$ is a *reverse martingale*, so that

$$E(S_k^2 | \mathscr{C}_n) = S_n^2 \quad \text{w.p. } 1, \quad \forall\, 2 \leq k \leq n, \tag{3.8}$$

$$S_n^2 \to \sigma^2 \quad \text{w.p. } 1/\text{1st mean}, \quad \text{as } n \to \infty, \tag{3.9}$$

where for both these results, $E(X^2) < \infty$ suffices and the normality of the d.f. of X is not all that crucial. Also, by definition,

$$(3.10)\qquad E(S_k^2 | \mathscr{C}_n) = \binom{n}{k}^{-1} \sum_{1 \le i_1 < \cdots < i_k \le n} S_{n,(i_1,\ldots,i_k)}^2 \quad (= S_n^2)\,,$$

so that by (3.6) and (3.10), we obtain

$$(3.11)\qquad Z_{n,k} \le S_n^2 \quad \text{for every} \quad k\colon 2 \le k \le n;\ n \ge 2\,.$$

Suppose now that $J_{k+1,n}^0 = \{j_1^0, \ldots, j_{k+1}^0\} \in \{1, \ldots, n\}$ be such that

$$(3.12)\qquad Z_{n,k+1} = S_{n,(j_1^0,\ldots,j_{k+1}^0)}^2, \quad k = m-1, \ldots, n-1\,.$$

Then, for every k-element subset $J_k = (j_1, \ldots, j_k) \subset J_{k+1,n}^0$, letting $j_{k+1} = J_{k+1,n}^0 \setminus J_k$, we have

$$(3.13)\qquad S_{n,(j_1,\ldots,j_k)}^2$$
$$= \binom{k}{2}^{-1} \left\{ \binom{k+1}{2} S_{n,(j_1^0,\ldots,j_{k+1}^0)}^2 - \sum_{l=1}^{k} \phi(X_{j_l}, X_{j_{k+1}}) \right\}$$
$$= S_{n,(j_1^0,\ldots,j_{k+1}^0)}^2 - \binom{k}{2}^{-1} \left\{ \sum_{l=1}^{k} \phi(X_{j_l}, X_{j_{k+1}}) - k S_{n,(j_1^0,\ldots,j_{k+1}^0)}^2 \right\},$$

so that for $X_{j_{k+1}}$ being one of the two extreme values (within the set $\{X_j\colon j \in J_{k+1,n}^0\}$), the term within the parenthesis $\{\cdot\}$ is nonnegative, and, hence,

$$(3.14)\qquad \min_{J_k \subset J_{k+1,n}^0} S_{n,(j_1,\ldots,j_k)}^2 \le S_{n,(j_1^0,\ldots,j_{k+1}^0)}^2 = Z_{n,k+1}\,.$$

On the other hand, by construction,

$$(3.15)\qquad Z_{n,k} = \min_{1 \le i_1 < \cdots < i_k \le n} S_{n,(i_1,\ldots,i_k)}^2 \le \min_{J_k \subset J_{k+1,n}^0} S_{n,(j_1,\ldots,j_k)}^2\,,$$

so that by (3.14) and (3.15), we obtain for every $n \ge 2$,

$$(3.16)\qquad Z_{n,k} \le Z_{n,k+1} \quad \text{w. p. 1,} \quad \text{for every} \quad k \ge m\,.$$

We may again recall that $\phi(a,b) = (a-b)^2/2$, so that using the order statistics $X_{n:1} \le \cdots \le X_{n:n}$ corresponding to $X_1, \ldots, X_n$ and following some routine steps based on the usual inclusion-exclusion principle, we obtain for every $n \ge k \ge m$ (≥ 2),

$$(3.17)\quad Z_{n,k} = \min\left\{\binom{k}{2}^{-1} \sum_{q\le i<j\le q+k-1} \phi(X_{n:i}, X_{n:j}) \colon 1 \le q \le n-k+1\right\}$$
$$= \min\{Z_{n,k}^{(q)} \colon 1 \le q \le n-k+1\}, \quad \text{say},$$

where $Z_{n,k}^{(q)}$ is the sample variance for the (ordered) sub-sample $\{X_{n:q}, \dots, X_{n:q+k-1}\}$ of size k, for $1 \le q \le n-k+1$, $k \ge 2$. Next, we note that

$$(3.18)\qquad Z_{n,k}^{(q)} = \begin{cases} Z_{n+1,k}^{(q)} & \text{if} \quad X_{n+1} > X_{n:q+k-1} \\ Z_{n+1,k}^{(q+1)} & \text{if} \quad X_{n+1} < X_{n:q}\,, \end{cases}$$

while for $X_{n:q} \le X_{n+1} \le X_{n:q+k-1}$, we may use (3.13) to compute $Z_{n+1,k}^{(q)}$ from $Z_{n,k}^{(q)}$. In fact, in this case, it follows that

$$(3.19)\qquad Z_{n+1,k}^{(q)} \text{ is smaller than at least one of } Z_{n,k}^{(q)} \text{ and } Z_{n,k}^{(q+1)}.$$

Thus, from the viewpoint of computation, given the picture at the n-th stage, we do not have to exhaust the full computation of $Z_{n+1,k}^{(q)}$ at the $(n+1)$-th stage, and this observation would be of considerable help, particularly if n is large.

THEOREM 3.1. *If the X_i's have a normal d.f. with a finite variance σ^2, then*

$$(3.20)\qquad \lim_{c\downarrow 0}\{E(N_c^*)/E(N_c)\} = (\pi/6)^{1/2}\,,$$

$$(3.21)\qquad \lim_{c\downarrow 0}\{\rho_c^*/\rho_c\} = (6+\pi)(24\pi)^{-1/2}\,,$$

so that as $c \downarrow 0$ for the PISR, there is about 27.6% reduction in the ASN at the expense of only about 5% increase in the risk.

PROOF. We start with some identities on the variance of truncated normal distributions. Let g and G be, respectively, the standard normal density function and d.f., and for every α, β such that $0 < \alpha$, $\beta < 1$ and $\alpha + \beta \le 1$, let $a = G^{-1}(\beta)$, $b = G^{-1}(\alpha+\beta)$. Then, we have

$$(3.22)\qquad \int_a^b x\,dG(x) = g(a) - g(b)\,,$$

$$(3.23)\qquad \int_a^b x^2\,dG(x) = \alpha - \{bg(b) - ag(a)\}\,,$$

so that

$$(3.24)\qquad \alpha^{-1}\int_a^b x^2 dG(x) - \left\{ \alpha^{-1}\int_a^b x dG(x) \right\}^2$$
$$= 1 - \alpha^{-1}\{bg(b) - ag(a)\} - \alpha^{-2}\{g(b) - g(a)\}^2 .$$

Note that $g'(x) = -xg(x)$ and $g''(x) = (x^2 - 1)g(x)$, so that for any $\alpha \in (0, 1)$, the right-hand side (rhs) of (3.24) is minimized for $\beta = (1 - \alpha)/2$. Thus, if we define d_α by $G(d_\alpha) = (1 + \alpha)/2$, then for $\beta = (1 - \alpha)/2$, the rhs of (3.24) reduces to

$$(3.25)\qquad 1 + 2\alpha^{-1}g'(d_\alpha) = 1 - 2\alpha^{-1}d_\alpha g(d_\alpha)$$
$$= \frac{\pi}{6}\alpha^2 + O(\alpha^4), \quad \text{as} \quad \alpha \to 0 ,$$

by Taylor series expansion.

Returning now to the proof of the theorem, we note that in (3.7), an admissible k must satisfy the condition $k^{1+a} \geq c^{-1/2}$, so that as $c \downarrow 0$, $k \to \infty$. We rewrite (3.7) as

$$(3.26)\qquad N_c^* = \inf\{n \geq m\ (\geq 2): n^2 c \geq (k/n)^{-2}(Z_{n,k}^{1/2} + k^{-a})^2,$$
$$\text{for some } k\text{: } m \leq k \leq n\} .$$

Keeping this in mind, we first study the asymptotic behavior of $(n/k)^2 Z_{n,k}$ when k and n are both large. Let $F_n(x) = n^{-1}\sum_{i=1}^{n} I(X_i \leq x)$ be the empirical d.f. of the sample of size n. Then, we have

$$(3.27)\qquad (1 - k^{-1})Z_{n,k}^{(q)} = (n/k)\int_{I_{nkq}} x^2 dF_n(x) - \left\{ (n/k)\int_{I_{nkq}} x dF_n(x) \right\}^2 ,$$

where

$$(3.28)\qquad I_{nkq} = \{x\text{: } X_{n:q} \leq x \leq X_{n:q+k-1}\} ,$$

$1 \leq q \leq n - k + 1$, $m \leq k \leq n$. Since $Z_{n,k}^{(q)}$'s are translation invariant, without any loss of generality, we may set $\mu = 0$, so that the d.f. F of X_1 is taken as $F(x) = G(x/\sigma)$, $x \in R$. Then,

$$(3.29)\qquad \int_{I_{nkq}} x dF(x) = \sigma\{g(\sigma^{-1}X_{n:q}) - g(\sigma^{-1}X_{n:q+k-1})\} ,$$

$$(3.30)\qquad \int_{I_{nkq}} x^2 dF(x) = \sigma^2(k/n) - \sigma\{X_{n:k+q-1}g(\sigma^{-1}X_{n:k+q-1})$$
$$- X_{n:q}g(\sigma^{-1}X_{n:q})\} ,$$

which are both *smooth* (i.e., bounded and differentiable) functions of $X_{n:q}$ and $X_{n:q+k-1}$. Let us now write

$$Z_{n,k}^{(q)*} = (n/k)\int_{I_{nkq}} x^2 dF(x) - \left\{ (n/k)\int_{I_{nkq}} x dF(x) \right\}^2, \tag{3.31}$$

$$I_{nkq}^0 = \left\{ x\colon \sigma G^{-1}\left(\frac{q}{n}\right) \leq x \leq \sigma G^{-1}\left(\frac{q+k-1}{n}\right)\right\}, \tag{3.32}$$

$$\xi_{n,k}^{(q)} = (n/k)\int_{I_{nkq}^0} x^2 dF(x) - \left\{ (n/k)\int_{I_{nkq}^0} x dF(x) \right\}^2. \tag{3.33}$$

Note that for $r = 1, 2, \ldots$, we have

$$\begin{aligned} \int_{I_{nkq}} x^r dF_n(x) - \int_{I_{nkq}} x^r dF(x) &= \int_{I_{nkq}} x^r d(F_n(x) - F(x)) \\ &= n^{-1/2}\left\{ n^{1/2} x^r [F_n(x) - F(x)]_{X_{n:q}}^{X_{n:q+k-1}} \right. \\ &\qquad \left. - r\int_{I_{nkq}} n^{1/2}[F_n(x) - F(x)] x^{r-1} dx \right\}. \end{aligned} \tag{3.34}$$

Also, for a normal d.f. F, for every finite r (>0),

$$\int_{-\infty}^{\infty} |x|^r \{F(x)[1 - F(x)]\}^{\gamma} dx < \infty, \tag{3.35}$$

for every $\gamma > 0$, while (see Ghosh (1972)) as $n \to \infty$

$$\sup_{-\infty<x<\infty} \{n^{1/2}|F_n(x) - F(x)|/[F(x)(1 - F(x))]^{1/4}\} = O(1), \tag{3.36}$$

w.p. 1 as well as in the r-th mean. Thus, the rhs of (3.34) is $O(n^{-1/2})$ w.p. 1, as well as in the s-th mean for all s (>0). On the other hand, proceeding as in (3.24)–(3.25), we can verify that for any given k/n, (3.33) is a minimum when $q = (n+k)/2$ and this minimum value is given by the rhs of (3.25) with $\alpha = k/n$. Thus, using (3.29), (3.30), (3.33) and the lower bound in (3.25), it follows that whenever $k \to \infty$, $n \to \infty$ with $n^{1/2}(k/n)^3 \to \infty$, at a rate faster than $\log n$, then for every $\varepsilon > 0$, there exists $c(\varepsilon) \in (0, \infty)$ such that

$$P\{|Z_{n,k}^{(q)*}/\xi_{n,k}^{(q)} - 1| > \varepsilon\} \leq c(\varepsilon) n^{-5}, \qquad \forall\, n \geq n_0\,; \tag{3.37}$$

$$\xi_{n,k}^{(q)} \gg \frac{\pi}{6}(k/n)^2 \gg O((n^{-1/2}\log n)^{2/3}). \tag{3.38}$$

Actually, in (3.37), we could have used an exponential rate in n, but $O(n^{-5})$ suffices. Similarly, using (3.34)–(3.36), we have for the same set

$$P\{(\xi_{n,k}^{(q)})^{-1}|Z_{n,k}^{(q)} - Z_{n,k}^{(q)*}| > \varepsilon\} \leq c(\varepsilon)n^{-5}, \qquad \forall\, n \geq n_0 . \tag{3.39}$$

Thus, for $k, n \to \infty$ with $n^{1/2}(k/n)^3(\log n)^{-1} \to \infty$, we have

$$P\left\{ Z_{n,k} < \sigma^2 \frac{\pi}{6}(k/n)^2(1-\varepsilon') \right\} \leq 2c(\varepsilon)n^{-4}, \qquad \forall\, n \geq n_0 , \tag{3.40}$$

where ε' ($\geq \varepsilon$) can be made to converge to 0 when $\varepsilon \downarrow 0$. Also, refer to Kiefer (1961, 1967). Now,

$$P\left\{ Z_{n,k} > \sigma^2 \frac{\pi}{6}(k/n)^2(1+\varepsilon') \right\} \leq 2c(\varepsilon)n^{-4}, \qquad \forall\, n \geq n_0 . \tag{3.41}$$

On the other hand, whenever $k \to \infty$, $n \to \infty$, but $k/n \to 0$, we write

$$(n/k)\int_{I_{nkq}} x^r dF_n(x) = k^{-1}\sum_{i=q}^{q+k-1}(X_{n:i})^r, \qquad r = 1, 2 , \tag{3.42}$$

as a linear combination of functions of order statistics (i.e., $h_1(x) = x$, $h_2(x) = x^2$). For normal distributions, it is known that the sample order statistics have finite moments of any finite order, and hence using the classical moment convergence results on order statistics (see Sen (1959), van Zwet (1964)), it follows that the $2p$-th central moment of (3.42) is

$$O(k^{-p}(\log n)^r), \qquad \forall\, r \text{ and } p = 1, 2, \dots . \tag{3.43}$$

Now, as has been pointed out immediately before (3.26) that $k^{1+a} \geq c^{-1/2}$, so that whenever n^2c is bounded, $k \gg O(n^{1/(1+a)})$, and hence, choosing p such that $p/(1+a) > 5$, we again arrive at (3.40)–(3.41). A similar technique may be used when n^2c is large, where we need to adjust p (in (3.43)) accordingly. Thus, we conclude that whenever $k \to \infty$, $n \to \infty$ with $k/n > n^{-\eta}$, for some $\eta > 0$, we have for every $n \geq n_0$, $\varepsilon > 0$,

$$p\left\{ \left| \left(\frac{k}{n}\right)^{-2} Z_{n,k} - \sigma^2 \frac{\pi}{6} \right| > \varepsilon \right\} \leq 2c(\varepsilon)n^{-4} . \tag{3.44}$$

Next, note that

$$(n_c^0)^2 c \to \sigma^2 \quad \text{as} \quad c \downarrow 0 , \tag{3.45}$$

so that using (3.26), (3.44) and (3.45) we readily obtain

$$N_c^* / \{(\pi/6)^{1/2} n_c^0\} \to 1 \quad \text{w.p. 1} \quad \text{as} \quad c \downarrow 0 . \tag{3.46}$$

Also, it follows from Robbins (1959) and others that

$$E(N_c)/n_c^0 \to 1 \quad \text{as} \quad c \downarrow 0\,, \tag{3.47}$$

while by construction, $N_c^*/N_c \le 1$ w.p. 1 for all $c > 0$. Hence, by the Dominated Convergence Theorem, we get

$$E(N_c^*)/\{(\pi/6)^{1/2} n_c^0\} \to 1 \quad \text{as} \quad c \downarrow 0\,, \tag{3.48}$$

and this completes the proof of (3.20).

Next, we note that by (3.7),

$$\begin{aligned} P(N_c^* > n) &\le P\{Z_{n,k} > (c^{1/2}k - k^{-a})^2,\ \forall\, k \le n\} \\ &\le P\{Z_{n,k_n} > (c^{1/2}k_n - k_n^{-a})^2\}, \qquad \forall\, k_n \in [m, n]\,. \end{aligned} \tag{3.49}$$

Thus, if we define $n_c^* \sim \sigma c^{-1/2}(\pi/6)^{1/2}$ as $c \downarrow 0$, then for every $n \ge n_c^*(1+\varepsilon)$, $\varepsilon > 0$, letting $k_n \sim \alpha n$ ($\alpha > 0$, small), by using (3.41), we obtain

$$P(N_c^* > n) = O(n^{-4})\,. \tag{3.50}$$

Similarly, for every $n \le n_c^*(1-\varepsilon)$, $\varepsilon > 0$, by (3.40), we get

$$\begin{aligned} P(N_c^* = n) &= P\{Z_{n',k} > (c^{1/2}k - k^{-a})^2,\ \forall\, n' \le n-1, k \le n', \\ &\qquad Z_{n,k} \le (c^{1/2}k - k^{-a})^2, \text{ for some } k \le n\} \\ &\le P\{Z_{n,k} \le (c^{1/2}k - k^{-a})^2, \text{ for some } k\colon n_0^* \le k \le n\} \\ &= O(n^{-4}(n - n_0^*))\,, \end{aligned} \tag{3.51}$$

where $n_0^* \sim c^{-1/(2+2a)}$. Thus, using the Hölder inequality, for $n' \le n_c^*(1-\varepsilon)$ and $r^{-1} + s^{-1} = 1$, we get

$$\begin{aligned} &E\{(\bar{X}_{N_c^*} - \mu)^2 I(N_c^* \le n')\} \\ &\quad = \sum_{n_0^* \le n \le n'} E\{(\bar{X}_n - \mu)^2 I(N_c^* = n)\} \\ &\quad \le \sum_{n_0^* \le n \le n'} \{P(N_c^* = n)\}^{1/r}\{E[|\bar{X}_n - \mu|^{2s}]\}^{1/s} \\ &\quad \le \sum_{n_0^* \le n \le n'} O(n^{-4/r})O(n^{-1})O((n - n_0^*)^{1/r}) \\ &\quad \le O((n' - n_0^*)^{1/r}) \sum_{n_0^* \le n \le n'} O(n^{-1-4/r}) \\ &\quad = O((n' - n_0^*)^{1/r})O((n_0^*)^{-4/r} - (n')^{-4/r}) \\ &\quad = O(c^{-1/(2r)})O(c^{2/(r(1+a))}) \end{aligned} \tag{3.52}$$

$$= O(c^{(3-a)/2(r+ra)}) .$$

Since $0 < a < 1$, $3 - a > 2 > 1 + a$. Also, since $E(|\bar{X}_n - \mu|^{2s}) = O(n^{-s})$ for every $s = 2, 3, \ldots$, choosing s so large that $(3 - a)/(r + ra) = (3 - a)(s - 1)/(s + sa) > 1$, we obtain from (3.52), as $c \downarrow 0$,

$$E\{(\bar{X}_{N_c^*} - \mu)^2 I(N_c^* \leq n')\} = O(c^{(1+\eta)/2}) , \tag{3.53}$$

for $\eta > 0$, while using (3.50) and a similar inequality, we have as $c \downarrow 0$,

$$E\{(\bar{X}_{N_c^*} - \mu)^2 I(N_c^* \geq n)\} = O(c^{(1+\eta)/2}) , \tag{3.54}$$

for all $n \geq n_c^*(1 + \varepsilon)$. On the central domain, that is, $n_c^*(1 - \varepsilon) \leq N_c^* \leq n_c^*(1 + \varepsilon)$, $\varepsilon > 0$, we may virtually repeat the steps in Sen and Ghosh (1981) and conclude that as $c \downarrow 0$, we have

$$\begin{aligned} E\{(\bar{X}_{N_c^*} - \mu)^2 I(|N_c^* - n_c^*| \leq \varepsilon n_c^*)\} \\ \sim \sigma^2 / n_c^* + o(c^{1/2}) = \sigma(6/\pi)^{1/2} c^{1/2} + o(c^{1/2}) . \end{aligned} \tag{3.55}$$

Therefore, as $c \downarrow 0$, by (3.48), (3.53)–(3.55), we get

$$\begin{aligned} \rho_c^* &\sim \sigma c^{1/2}\{(6/\pi)^{1/2} + (\pi/6)^{1/2}\} \\ &\sim \rho_c(6 + \pi)/(24\pi)^{1/2} , \end{aligned} \tag{3.56}$$

and this completes the proof of (3.21). □

Suppose now that in (3.7) we replace $c^{1/2}$ by $dc^{1/2}$, for some $d > 1$, and denote the corresponding stopping variable by $N_c^*(d)$, that is,

$$\begin{aligned} N_c^*(d) = \inf\{n \geq m\ (\geq 2)\colon k \geq c^{-1/2} d(Z_{n,k}^{1/2} + k^{-a}) \\ \text{for some } k\colon m \leq k \leq n\} . \end{aligned} \tag{3.57}$$

Let us also denote by $\rho_c^*(d)$, the risk of the PISR based on the estimator $\bar{X}_{N_c^*(d)}$. Then, virtually repeating the proof of Theorem 3.1, we obtain

$$\lim_{c \downarrow 0}\{E(N_c^*(d))/E(N_c)\} = d(\pi/6)^{1/2} , \tag{3.58}$$

$$\lim_{c \downarrow 0}\{\rho_c^*(d)/\rho_c\} = \frac{1}{2}\{d(\pi/6)^{1/2} + d^{-1}(6/\pi)^{1/2}\} . \tag{3.59}$$

Then, if we let

(3.60) $$d_{0,\eta} = (6/\pi)^{1/2}(1 - \eta)\,,$$

for some arbitrary small η (> 0), from (3.58) and (3.59), we obtain

(3.61) $$\lim_{c \downarrow 0}\{E(N_c^*(d_{0,\eta}))/E(N_c)\} = 1 - \eta\,,$$

(3.62) $$\lim_{c \downarrow 0}\{\rho_c^*(d_{0,\eta})/\rho_c\} = 1 + \frac{1}{2}\eta^2/(1 - \eta)\,.$$

Thus, for $\eta = .1$ (or .05), we have 10% (or 5%) reduction in the ASN at the expense of only .5% (or .13%) increase in the relative risk. Thus, allowing $\eta \to 0$ and noting that $\eta^2/(2 - 2\eta) \sim \eta^2/2$, we arrive at the following.

THEOREM 3.2. *Let $\{\varepsilon_v\colon v = 1, 2, \ldots\}$ be a sequence of positive numbers such that $\lim_{v \uparrow \infty} \varepsilon_v = 0$. Then, there exists a sequence $\{\eta_v\colon v = 1, 2, \ldots\}$ of positive numbers such that $\eta_v \le (2\varepsilon_v)^{1/2}$ for all $v = 1, 2, \ldots$, and defining $d_v = (1 - \eta_v)(6/\pi)^{1/2}$ and $N_{c,v}^* = N_c^*(d_v)$, we have for every $\varepsilon > 0$ the existence of v_0 such that $\varepsilon_{v_0} \le \varepsilon$ and*

(3.63) $$\lim_{c \downarrow 0}\{E(N_{c,v_0}^*)/E(N_c)\} \le 1 - (2\varepsilon)^{1/2}\,,$$

(3.64) $$\lim_{c \downarrow 0}\{\rho_c^*(d_{v_0})/\rho_c\} = 1 + \varepsilon\,.$$

Thus, the modified PISR $N_c^*(d_v)$ is ε-risk efficient and has a smaller ASN than the usual stopping rule N_c.

Although it appears that for the PISR the sufficiency principle is violated, nevertheless in the asymptotic setup, the PISR or its modified version compares very favorably with the original stopping rule.

In order to get some ideas regarding the comparative behaviors of N_c and N_c^* for small values of n_c^0, we ran small-scale simulations with 200 replications. We fixed $\sigma = 1$, $\mu = 20$, $m = 5$, 10, $n_c^0 = 25$, 50, 75, $a = .5$, .9, 1.3 and $d = 1.05$, 1.2, 1.35. Note that d should lie in the interval $(1, \sqrt{6/\pi})$ so that we may expect some saving in the ASN. We observed negative savings, that is, $E(N_c^*)$ exceeded $E(N_c)$ all the time when $a = .5$. When $a = .9$, we noted that $E(N_c^*)$ and $E(N_c)$ came out to be nearly the same. Table 1 summarizes our findings when $a = 1.3$ and $m = 10$. When d gets closer to $\sqrt{6/\pi} \sim 1.382$, naturally percentage savings are expected to go down. For $a = 1.3$ and small values of n_c^* like 50 and 75, we do observe significant savings in the ASN while using the stopping rule (3.26); however, the associated "risk-efficiency" seems to be nearly the same as it would be for the stopping rule (3.2). This is definitely encouraging and we foresee the need for future in-depth numerical studies along these lines.

Table 1. Savings and associated risk.

n_c^0	d	Estimated Savings (%)	Estimated ρ_c^*/ρ_c
25	1.05	21.6	1.051
	1.20	12.2	1.025
	1.35	2.6	1.018
50	1.05	28.8	1.087
	1.20	20.4	1.047
	1.35	10.9	1.025
75	1.05	31.6	1.096
	1.20	21.4	1.050
	1.35	12.6	1.027

4. Fixed-width confidence interval for the normal mean and PISR

Given two numbers d (>0) and $\alpha \in (0, 1)$, we wish to construct a confidence interval I_n for μ such that

$$\text{the width of } I_n = 2d \text{ and } P(\mu \in I_n) \sim 1 - \alpha \,. \tag{4.1}$$

The sequential procedure of Ray (1957), later studied in Starr (1966*a*), can be defined by introducing a stopping variable

$$N_d = \inf\{n \geq m\ (\geq 2): n \geq \rho^2 S_n^2 d^{-2}\}\,, \tag{4.2}$$

where S_n^2 is defined as in (3.2) and ρ is the upper $50\alpha\%$ point of the standard normal d.f., that is, $G(\rho) = 1 - \alpha/2$. Recall from Section 3 that g and G stand for the standard normal density function and d.f. The confidence interval I_{N_d} for μ is then taken as

$$I_{N_d} = [\bar{X}_{N_d} - d,\ \bar{X}_{N_d} + d]\,. \tag{4.3}$$

One may again note that (4.2) is of the form (1.1) with $\nu = [d^{-2}]^* + 1$ and $b_\nu = \rho^2/d^2$. From Starr (1966*b*), it follows that as $d \downarrow 0$,

$$E(N_d) \sim \rho^2\sigma^2 d^{-2} = n_d^* \quad \text{(say)} \quad \text{and} \quad P(\mu \in I_{N_d}) \sim 1 - \alpha\,. \tag{4.4}$$

Following Chow and Robbins (1965), we may as well introduce the fudge factor n^{-a^*} in (4.2), for some $0 < a^* < 1$.

To introduce the PISR, we conceive of a positive number ρ^* and then parallel to (3.7) and (3.57), we define

$$N_d^*(\rho^*) = \inf\{n \geq m\ (\geq 2): k \geq (\rho^*/d)^2(Z_{n,k} + k^{-a^*}), \tag{4.5}$$

for some k: $m \le k \le n\}$.

Then, we propose the confidence interval $I_{N_d^*(\rho^*)} = [\bar{X}_{N_d^*(\rho^*)} \pm d]$ for μ.

From the detailed analysis given in Section 3, we can show that as $d \downarrow 0$, we get

$$N_d^*(\rho^*) \bigg/ \left\{ \frac{\pi}{6} \left(\frac{\rho^*}{\rho} \right)^2 n_d^* \right\} \to 1 \quad \text{w.p. } 1/\text{1st mean ,} \tag{4.6}$$

and

$$P\{\mu \in I_{N_d^*(\rho^*)}\} \sim 2G\left(\rho^* \left(\frac{\pi}{6} \right)^{1/2} \right) - 1 \, . \tag{4.7}$$

Thus, if we let

$$(\pi/6)^{1/2} \rho^* = \rho \, , \tag{4.8}$$

then from (4.6) and (4.7), we note that for such particular choice of ρ^*, (4.1) holds for the PISR introduced in (4.5). Thus, for the PISR in (4.5), we achieve both the "asymptotic consistency" and "asymptotic efficiency" results by simply adjusting ρ in (4.2) and replacing it by ρ^* given by (4.8). More generally, if we conceive of a sequence $\{\varepsilon_\nu : \nu = 1, 2, \ldots\}$ of positive numbers such that $\varepsilon_\nu \to 0$ as $\nu \to \infty$ and define

$$\rho_\nu^* = (6/\pi)^{1/2} (1 + \varepsilon_\nu) \rho \, , \tag{4.9}$$

then for the corresponding $N_d^*(\rho_\nu^*)$, we have as $d \downarrow 0$,

$$E\{N_d^*(\rho_\nu^*)\} / E(N_d) \to (1 + \varepsilon_\nu)^2 \, . \tag{4.10}$$

Next, we obtain

$$\begin{aligned} P\{\mu \in I_{N_d^*(\rho_\nu^*)}\} &\sim 2G(\rho(1 + \varepsilon_\nu)) - 1 \\ &\sim 2G(\rho) + 2\varepsilon_\nu \rho g(\rho) + \varepsilon_\nu^2 \rho^2 g'(\rho) - 1 + o(\varepsilon_\nu^2) \\ &= 1 - \alpha + 2\varepsilon_\nu \rho g(\rho) - \varepsilon_\nu^2 \rho^3 g(\rho) + o(\varepsilon_\nu^2) \\ &= 1 - \alpha + \varepsilon_\nu \rho g(\rho) \{2 - \rho^2 \varepsilon_\nu\} + o(\varepsilon_\nu^2) \\ &\sim 1 - \alpha + \varepsilon_\nu \rho^2 \{1 - G(\rho)\} \{2 - \rho^2 \varepsilon_\nu\} + o(\varepsilon_\nu^2) \\ &= 1 - \alpha + \varepsilon_\nu \rho^2 \left\{ \alpha - \frac{1}{2} \alpha \rho^2 \varepsilon_\nu \right\} + o(\varepsilon_\nu^2) \end{aligned} \tag{4.11}$$

Table 2. Savings and coverage probability: α=.05, m=10.

n_d^*	ρ^*	a	Estimated Savings (%)	Estimated C.P.
25	1.05ρ	.6	18.5	.914
		.7	25.7	.899
50	1.05ρ	.6	34.9	.889
		.7	42.0	.867
25	1.3ρ	.6	20.2	.971
		.7	28.3	.960
50	1.3ρ	.6	34.2	.955
		.7	41.3	.941
70	1.1ρ	.5	24.7	.910
		.7	41.4	.885
90	1.1ρ	.5	27.1	.930
		.7	45.0	.840
70	1.37ρ	.5	−7.8	.960
		.7	18.4	.915
90	1.37ρ	.5	−3.5	.940
		.7	22.1	.930

$$= 1 - \alpha + \alpha\varepsilon_v\rho^2\left\{1 - \frac{1}{2}\rho^2\varepsilon_v + o(\varepsilon_v)\right\}$$

$$= 1 - \alpha + \eta_v, \qquad \text{say}.$$

Thus, in order to be able to claim that the asymptotic coverage probability in this case is at least $1 - \alpha$, we may select ε_v so small that η_v is small, while in (4.10), $2\varepsilon_v + \varepsilon_v^2$ is also small. Note that $\rho^2\alpha = 2\rho^2\{1 - G(\rho)\}$ is bounded by $(2/\pi)^{1/2}\rho \exp(-\rho^2/2)$ for all ρ and it converges to zero as $\rho^2 \to \infty$. Thus, compared to the increase in the relative ASN (i.e., $\varepsilon_v(2 + \varepsilon_v)$), the gain in the coverage probability (i.e., $2\rho^2\{1 - G(\rho)\}\varepsilon_v\{1 - \rho^2\varepsilon_v/2\}$) is relatively small. Thus, there may not be any practical gain in choosing ρ_v^* as in (4.9). On the other hand, if we replace ε_v by $-\varepsilon_v$ in (4.9), then instead of (4.10) we would have $(1 - \varepsilon_v)^2 < 1$ and (4.11) would be changed to $1 - \alpha - \alpha\varepsilon_v\rho^2\{1 + \rho^2\varepsilon_v/2 + o(\varepsilon_v)\}$. Thus, sacrificing only a small fraction of the coverage probability, we may achieve a somewhat larger fraction of the reduction of ASN. For example, if we let $\varepsilon_v = .05$ (or .01), we have 9.75% (or 2%) reduction in the ASN along with a reduction of $.05\alpha\rho^2(1 + .025\rho^2)$ (or $.01\alpha\rho^2(1 + .005\rho^2)$) of the target coverage probability. For $\alpha = .05$, we thus observe the possibility of 5% reduction of the ASN for our PISR by lowering the coverage probability from .95 to .945.

In order to get some feeling regarding the comparative behaviors of N_d and N_d^* for small values of n_d^*, we ran small-scale simulations with 200 replications. We fixed $\sigma = 1$, $\mu = 20$, $\rho = 1.96$, $m = 5$, 10, $n_d^* = 25$, 50, 70, 90, $\rho^*/\rho = 1.05$, 1.1, 1.3, 1.37. Notice that $E(N_d^*)$ is expected to be smaller than $E(N_d)$ when $1 < \rho^*/\rho < \sqrt{6/\pi} \sim 1.382$. We estimate the coverage probability (C.P.) merely by the relative frequency of the constructed intervals covering μ-value. Table 2 partially summarizes our findings.

Again, we think that Table 2 speaks for itself. It seems that by properly choosing a and ρ^*, the rule (4.5) would provide substantial saving compared to $E(N_d)$ and at the same time, the achieved coverage probability may not be unattractive at all in comparison with the target $1 - \alpha$. Let us point out another aspect. Suppose we implement the original stopping rule (4.2) where we plug in $1 - \alpha =$ estimated C.P. from Table 2. The ASN for that adaptive rule (which is not permutationally invariant) always came out within a small fraction of $E(N_d^*)$ in our simulations, and, of course, asymptotically they are the same. The picture is definitely encouraging and worth pursuing in the future.

Acknowledgements

Eight students enrolled in the sequential analysis course at The University of Connecticut during the fall semester, 1986, took part in the related simulation exercises. We are grateful to them for sharing with us their insights and understanding. The authors also thank the referees for helpful comments.

References

Bahadur, R. R. (1954). Sufficiency and statistical decision functions, *Ann. Math. Statist.*, **25**, 423–462.

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean, *Ann. Math. Statist.*, **36**, 457–462.

Ghosh, M. (1972). On the representation of linear functions of order statistics, *Sankhyā Ser. A*, **34**, 349–356.

Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic normality of stopping times in sequential analysis (unpublished manuscript).

Ghosh, M. and Mukhopadhyay, N. (1976). On two fundamental problems of sequential estimation, *Sankhyā Ser. B*, **38**, 203–218.

Ghosh, M. and Mukhopadhyay, N. (1979). Sequential point estimation of the mean when the distribution is unspecified, *Comm. Statist. A—Theory Methods*, **8**, 637–652.

Ghosh, M. and Mukhopadhyay, N. (1981). Consistency and asymptotic efficiency of two-stage and sequential procedures, *Sankhyā Ser. A*, **43**, 220–227.

Kiefer, J. (1961). On large deviations of the empiric D. F. of vector chance variables and a law of iterated logarithm, *Pacific J. Math.*, **11**, 649–660.

Kiefer, J. (1967). On Bahadur's representation of sample quantiles, *Ann. Math. Statist.*, **38**, 1323–1342.

Ray, W. D. (1957). Sequential confidence intervals for the mean of a normal population with unknown variance, *J. Roy. Statist. Soc. Ser. B*, **19**, 133–143.

Robbins, H. (1959). Sequential estimation of the mean of a normal population, *Probability and Statistics* (H. Cramér Vol., ed. U. Grenander), 235–245, Almquist & Wiksell, Uppsala.

Sen P. K. (1959). On the moments of the sample quantiles, *Calcutta Statist. Assoc. Bull.*, **9**, 1–20.

Sen, P. K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on *U*-statistics, *Sankhyā Ser. A*, **43**, 331–344.

Starr, N. (1966*a*). The performance of a sequential procedure for fixed-width interval estimation of the mean, *Ann. Math. Statist.*, **37**, 36–50.

Starr, N. (1966*b*). On the asymptotic efficiency of a sequential procedure for estimating the mean, *Ann. Math. Statist.*, **37**, 1173–1185.

van Zwet, W. R. (1964). Convex transformations of random variables, Math. Center Tract #7, Amsterdam.

Wijsman, R. (1959). On the theory of BAN estimator, *Ann. Math. Statist.*, **30**, 185–191.

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