The Annals of Statistics
1974, Vol. 2, No. 3, 540-552

SEQUENTIAL RANK TESTS FOR LOCATION[1]

By Pranab Kumar Sen and Malay Ghosh

University of North Carolina, Chapel Hill
and Indian Statistical Institute, Calcutta

In this paper, for the one and two sample location problems, a class of sequential tests based on robust rank order statistics is developed. The proposed tests terminate with probability one for square integrable score functions. Under more stringent regularity conditions and for local alternatives, the OC and ASN of the proposed tests are obtained, and the allied asymptotic relative efficiency results with respect to the sequential probability ratio and likelihood ratio tests are studied.

1. Introduction. The object of the present investigation is to develop a general class of sequential rank order tests (SROT) for the one and two sample location problems. We motivate the SROT by the same principle underlying the asymptotic sequential likelihood ratio tests (SLRT) of Bartlett (1946) and Cox (1963). It is shown that for square integrable score functions, the proposed SROT terminates with probability one. Under comparatively more stringent regularity conditions and for local alternatives, the OC and ASN of the proposed SROT are obtained, and the allied asymptotic relative efficiency (ARE) results with respect to the SLRT and the sequential probability ratio tests (SPRT) are studied.

Along with the preliminary notions, the SROT are proposed in Section 2. The main theorems dealing with the properties of the SROT are presented in Section 3, and their proofs are considered in Section 4. Relative performances of the different tests are studied in Section 5. The last section includes, by way of remarks, alternative tests by Albert (1966), Hall (1969), and others. Some of the proofs of certain results in the main body of the paper are sketched in the appendix. We may remark that in the classical two sample location problem, once observations are drawn in pairs at each stage of experimentation, and we work with their differences (which are distributed symmetrically about the difference of locations), the problem reduces to the corresponding one sample case. Hence, we shall deal specifically with the one sample location problem only, while the above remark takes care of the two sample case.

2. Preliminary notions and the proposed SROT. Let $\{X_i, i \geqq 1\}$ be a sequence of independent and identically distributed random variables (i.i.d. rv) with an

Received July 1971; revised June 1973.

[1] Work partially supported by the National Institutes of Health, Grant-GM 12868, and partially by the Aerospace Research Laboratories, Air Force Systems Command, Contract No. F33615-71-C-1927. Reproduction in whole or in part permitted for any purpose of the U.S. Government.

AMS 1970 *subject classifications.* 62L10, 62G10.

Key words and phrases. Asymptotic relative efficiency, asymptotic sequential tests, average sample number, OC function, termination with probability one, Wiener process approximation.

absolutely continuous distribution function (df) $F_\theta(x)$, where we assume that

$$F_\theta(x) = F(x-\theta)\,, \qquad \theta \text{ unknown and } F \in \mathscr{F}_0\,; \tag{2.1}$$

$\mathscr{F}_0$ is the class of all absolutely continuous df's symmetric about 0 for which both the density function and its derivative exist almost everywhere. Our basic problem is to test sequentially

$$H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta = \theta_1 = \theta_0 + \Delta\,, \qquad \text{where } \theta_0 \text{ and } \Delta\,(>0) \text{ are known;} \tag{2.2}$$

without any loss of generality, we let $\theta_0 = 0$, as otherwise, work with $X_i - \theta_0$, $i \geqq 1$. Wald's (1947) SPRT is applicable for testing (2.2) when F is completely specified. Bartlett (1946) and Cox (1963) extended SPRT to asymptotic SLRT when F is of specified form involving some nuisance parameter, say δ. Starting with an initial sample size $n_0(\Delta)$, at least moderately large, working with the usual likelihood ratio statistics, namely, at the nth stage with λ_n^*, the logarithm of the ratio of the two maxima of the likelihood function under the two (composite) hypotheses in (2.2), and following the same stopping rule as in the SPRT, these authors were able to show that by virtue of appropriate Wiener process approximations to the sequence $\{\lambda_n^*,\ n \geqq n_0(\Delta)\}$, their SLRT has asymptotically (as $\Delta \to 0$) all the properties of the SPRT (corresponding to the case where δ is known). We may remark that for both the SLRT and SPRT, the form of F is assumed to be given, while in practice, the form of F is rarely known. The properties of these procedures, including their optimality, are open to questions when the true and the assumed forms of F are not the same. Basically, these procedures are not very robust.

In the nonparametric setup, F is unknown and belongs to some class of df's, such as $\mathscr{F}_0$ in (2.1). We are interested in studying SROT which are based on robust rank statistics and remain valid for the entire class $\mathscr{F}_0$. For one sample rank order statistics, martingale property, invariance principles and almost sure (a.s.) convergence to Wiener processes have been studied by Sen and Ghosh (1971, 1973a, b) and Sen (1974). These results provide convenient tools for motivating and studying the various properties of the SROT which we formulate below.

For every $n\,(\geqq 1)$ and real $d\,(-\infty < d < \infty)$, let $R_{ni}(d) = \frac{1}{2} + \sum_{j=1}^{n} c(|X_i - d| - |X_j - d|)$ be the rank of $|X_i - d|$ among $|X_1 - d|, \cdots, |X_n - d|$, where $c(u)$ is 1, $\frac{1}{2}$ or 0 according as u is $>$ $=$ or < 0. Consider now a set of *scores* $J_n(i/(n+1)) = EJ(U_{ni})$ or $J(i/(n+1))$, $i = 1, \cdots, n.$, where $U_{n1} < \cdots < U_{nn}$ are the ordered random variables of a sample of size n from the rectangular (0, 1) df, and the *score function* $J(u)$ is defined by

$$J(u) = -g'(G^{-1}((1+u)/2))/g(G^{-1}((1+u)/2))\,, \qquad 0 \leqq u < 1,\ G \in \mathscr{F}_0\,, \tag{2.3}$$

where G is non-degenerate and strongly unimodal, so that $J(u)$ is $\uparrow$ in u: $0 \leqq u < 1$. Notable examples of G are the normal and the logistic df's; the corresponding scores are known as the normal and the Wilcoxon scores. Let then

$$T_n(d) = \sum_{i=1}^{n} J_n((n+1)^{-1}R_{ni}(d))s(X_i - d)\,; \qquad s(u) = 2c(u) - 1\,, \tag{2.4}$$

be the usual one sample rank order statistics. Under H_0 as well as for general alternatives, invariance principles for these statistics are studied by Sen and Ghosh (1973a, b) and Sen (1974). Also, under H_1 in (2.2) and for small Δ, as $n \to \infty$,

$$n^{-1}\nu^{-1}E_\Delta T_n(0) \to \Delta/\tau\,; \quad \tau = \nu/C(F)\,, \tag{2.5}$$

where

$$\nu^2 = \int_0^1 J^2(u)\,du\,, \qquad C(F) = \int_0^\infty (d/dx)J(H(x))\,dH(X) \qquad \text{and} \tag{2.6}$$
$$H(x) = F(x) - F(-x)\,, \quad x \geqq 0\,.$$

Hence, if τ were known, one could have employed the results of Dvoretzky, Kiefer and Wolfowitz (1953) on SPRT for Wiener processes to contruct asymptotic sequential tests based on $\{T_n(0),\, n \geqq n_0(\Delta)\}$. We therefore proceed to construct first a strongly consistent estimator of τ.

When $\theta = 0$, $T_n(0)$ has a known df, symmetric about 0, and for large n, $n^{-\frac{1}{2}}T_n(0)/\nu$ has asymptotically the standard normal df. Hence, for every $\varepsilon > 0$, we can select a $T_{n,\varepsilon}$ such that

$$P_\theta\{|T_n(\theta)| \leqq T_{n,\varepsilon}\} = P_0\{|T_n(0)| \leqq T_{n,\varepsilon}\} \to 1 - \varepsilon \qquad \text{as } n \to \infty\,. \tag{2.7}$$

Let then

$$\theta_{L,n}^* = \sup\{d\colon T_n(d) > T_{n,\varepsilon}\}\,, \qquad \theta_{U,n}^* = \inf\{d\colon T_n(d) < -T_{n,\varepsilon}\}\,; \tag{2.8}$$

$$C_n^* = 2T_{n,\varepsilon}/n(\theta_{U,n}^* - \theta_{L,n}^*)\,; \qquad \hat{\tau}_n = \nu/C_n^*\,. \tag{2.9}$$

Then, it follows from Sen and Ghosh (1971) that as $n \to \infty$, C_n^* (or $\hat{\tau}_n$) converges a.s. to $C(F)$ (or $\hat{\tau}_n$); also, the distribution of C_n^* (or $\hat{\tau}_n$) does not depend on θ. The proposed SROT may be formulated then as follows.

Corresponding to preassigned strength (α, β), consider two positive numbers (A, B): $0 < B < 1 < A < \infty$, where $A \leqq (1-\beta)/\alpha$ and $B \geqq \beta/(1-\alpha)$. Starting with an initial sample of size $n_0\,(= n_0(\Delta))$, continue drawing observations one by one as long as

$$b\nu^2 < \Delta C_m^* T_m(\Delta/2) < a\nu^2 \qquad (\text{where } a = \log A \text{ and } b = \log B)\,; \tag{2.10}$$

if for the first time (2.10) is vitiated for $m = n$ and $\Delta C_n^* T_n(\Delta/2)$ is $\leqq b\nu^2$ (or $\geqq a\nu^2$), accept H_0 (or H_1) in (2.2); the corresponding stopping variable is denoted by $N_J(\Delta)$, so that $N_J(\Delta) \geqq n_0(\Delta)$.

3. Properties of the SROT. We have the following theorems.

THEOREM 3.1. *Under* (2.1) *and for square integrable non-decreasing score function, for every (fixed)* θ *and* Δ, $\lim_{n\to\infty} P_\theta\{N_J(\Delta) > n\} = 0$ *i.e., the process in* (2.10) *terminates with probability one.*

For the study of the OC and ASN of the proposed SROT, we confine ourselves to local alternatives. Here, for theoretical justifications, we consider the asymptotic case where we let $\Delta \to 0$. This is comparable to the asymptotic situation in the dual problem of bounded length (sequential) confidence interval for θ considered by Chow and Robbins (1965) and others. First, for $\Delta \to 0$, the excess over the boundaries in the classical SPRT is negligible, so that we can take

$$e^a = A = (1-\beta)/\alpha \quad \text{and} \quad e^b = B = \beta/(1-\alpha)\,; \quad 0 < \alpha, \beta < \tfrac{1}{2}\,. \tag{3.1}$$

Second, compared to the Wald (1947) requirement of finite moment generating function of $\log\{f(X_i - \Delta)/f(X_i)\}$, we assume as in Sen and Ghosh (1971) that

$$(0 \leqq)\, J'(u) \leqq K(1-u)^{-1}\,, \qquad 0 \leqq u < 1,\ 0 < K < \infty\,, \tag{3.2}$$

which implies that for some $t_0\,(> 0)$,

$$M(t) = \textstyle\int_0^\infty \exp[tJ(u)]\,du < \infty \qquad \text{for all } t \leqq t_0\,. \tag{3.3}$$

Third, as will be seen later on that the ASN of $N_J(\Delta)$ is $O(\Delta^{-2})$ as $\Delta \to 0$, and up to the first order of approximation, the OC and ASN are not affected by an increase in $n_0(\Delta)$ (with $\Delta \to 0$), provided it increases at a slower rate. On the other hand, as in Bartlett (1946) and Cox (1963), we need to have an initial sample of at least moderately large size so as to estimate $C(F)$ by C_n^* with reasonable accuracy. Hence, to be specific, we assume that for some γ: $0 < \gamma \leqq 1$,

$$\lim_{\Delta\to 0}\{\Delta^\gamma n_0(\Delta)\} = \infty \quad \text{and} \quad \lim_{\Delta\to 0}\{\Delta^2 n_0(\Delta)\} = 0\,. \tag{3.4}$$

Finally, for every (fixed) $\theta\,(\neq 0)$, using the strong convergence of $T_n(\Delta/2)/n$ and C_n^* (viz., Sen and Ghosh (1971)), it can be shown that the OC of the SROT converges (as $\Delta \to 0$) either to 0 or 1 depending on whether $\theta < 0$ or < 0. Hence, to tackle the asymptotic case and to avoid the limiting degeneracy, we assume that

$$\theta = \phi\Delta\,, \qquad \phi \in I\,, \qquad \text{where } I = \{\phi : |\phi - \tfrac{1}{2}| \leqq K_0\}\,, \tag{3.5}$$

for some $K_0\,(> \frac{1}{2})$. We may remark that (3.1), (3.4) and (3.5) are also implicit in Bartlett (1946) and Cox (1963) for a rigorous treatment of their asymptotic SLRT.

On denoting by $L_J^{(F)}(\phi, \Delta)$ the OC of the SROT based on $N_J(\Delta)$ in (2.10) when $\theta = \phi\Delta$ and the underlying df is F, we have the following.

THEOREM 3.2. *Under the assumptions made above, for every* $\phi \in I$,

$$\begin{aligned} \lim_{\Delta\to 0} L_J^{(F)}(\phi, \Delta) &= (A^{1-2\phi} - 1)/(A^{1-2\phi} - B^{1-2\phi})\,, && \phi \neq \tfrac{1}{2}\,, \\ &= (\log A)/(\log A - \log B)\,, && \phi = \tfrac{1}{2}\,. \end{aligned} \tag{3.6}$$

Hence, asymptotically, the OC *of the* SROT *does not depend on* $F\ (\in \mathscr{F}_0)$ *and the score function* $J(u)$, $0 \leqq u < 1$, *and the* SROT *has the prescribed strength* (α, β). *That is, the* SROT *is asymptotically distribution-free and consistent.*

THEOREM 3.3. *Under the assumptions made earlier, for every* $\phi \in I$,

$$\lim_{\Delta\to 0}\{\Delta^2 E_\phi[N_J(\Delta)]\} = \psi(\phi, \tau)\,, \tag{3.7}$$

where E_ϕ refers to the expectation under $\theta = \phi\Delta$, and

$$\psi(\phi, \tau) = [P(\phi) \log B + \{1 - P(\phi)\} \log A][\tau^2/(\phi - \tfrac{1}{2})], \qquad \phi \neq \tfrac{1}{2}, \tag{3.8}$$
$$= -P'(\tfrac{1}{2})/[\log A - \log B], \qquad \phi = \tfrac{1}{2},$$

where τ is defined by (2.5), *$P(\phi)$ is the right-hand side of* (3.6), *and $P'(\frac{1}{2})$ is $(d/d\phi)P(\phi)$ at $\phi = \frac{1}{2}$.*

4. Proofs of the theorems. First, consider Theorem 3.1. By (2.10), we have for $n \geqq n_0(\Delta)$,

$$P\{N_J(\Delta) > n \mid \theta\} \leqq P\{\nu^2 b/(\Delta C_n^*) < T_n(\Delta/2) < \nu^2 a/(\Delta C_n^*) \mid \theta\} \tag{4.1}$$
$$= P\{\nu^2 b/(\Delta C_n^*) < T_n(\Delta/2 - \theta) < \nu^2 a/(\Delta C_n^*) \mid \theta = 0\},$$

where by Sen (1966) and Jurečková (1969), C_n^* converges in probability to $C(F)$ as $n \to \infty$. When $\theta = \Delta/2$, the proof follows directly by noting that under $\theta = 0$, $n^{-\frac{1}{2}}T_n(0)$ is asymptotically normal with mean 0 and variance ν^2. For $\theta \neq \Delta/2$, the asymptotic normality of $n^{-\frac{1}{2}}[T_n(\Delta/2 - \theta) - ET_n(\Delta/2 - \theta)]$ requires an additional condition on the positiveness of its variance which neither holds universally nor is needed here. Note that by Sen (1970), for every real δ, under $\theta = 0$, as $n \to \infty$,

$$n^{-1}T_n(\delta) \to_{\text{a.s.}} \xi(\delta) = 2 \textstyle\int_0^\infty J(F(x + \delta) - F(-x + \delta))\, dF(x + \delta) \tag{4.2}$$
$$- (\textstyle\int_0^\infty J(u)\, du),$$

where under (2.1) and (2.3), $\xi(\delta)$ is $<$, $=$ or > 0 according as δ is $>$, $=$ or < 0. Hence, the proof follows from (4.1) and (4.2). □

The following lemma (proved in Lemma 4.2 of Sen and Ghosh (1971)) will be used repeatedly in the sequel.

LEMMA 4.1. *Under* (2.1) *and* (3.2), *for every $s > 0$, there exist positive constants (c_{s1}, c_{s2}) and an integer n_s, such that for $n \geqq n_s$,*

$$P\{|C_n^*/C(F) - 1| > c_{s1} n^{-\frac{1}{2}}(\log n)^3\} \leqq c_{s2} n^{-s}. \tag{4.3}$$

Hence, $C_n^*/C(F) \to 1$ a.s. as $n \to \infty$.

Let us now consider the proof of Theorem 3.2. For every $\varepsilon > 0$, we let

$$a_{\varepsilon,i} = \log A_{\varepsilon,i} = a[1 + (-1)^i\varepsilon] \qquad \text{and} \tag{4.4}$$
$$b_{\varepsilon,i} = \log B_{\varepsilon,i} = b[1 + (-1)^i\varepsilon], \qquad i = 1, 2.$$

Suppose now that in the sequential procedure sketched in (2.10), we replace C_m^* by $C(F)$, a by $a_{\varepsilon,j}$ and b by $b_{\varepsilon,i}$; the corresponding stopping variables are denoted by $N_{J,\varepsilon}^{(ij)}(\Delta)$ and the OC function by $L_{J,ij}^{(F)}(\phi, \Delta)$, for $i, j = 1, 2$. Then, rewriting (2.10) as $b\nu^2/C_m^* < \Delta T_m(\Delta/2) < a\nu^2/C_m^*$, we obtain by virtue of (3.4), (3.5), Lemma 4.1 and some standard steps that for every $\varepsilon > 0$,

$$\lim_{\Delta\to 0} P_\phi\{N_{J,\varepsilon}^{(11)}(\Delta) \leqq N_J(\Delta) \leqq N_{J,\varepsilon}^{(22)}(\Delta)\} = 1 \qquad \text{for all } \phi \in I, \tag{4.5}$$

where P_ϕ stands for the probability computed under $\theta = \phi\Delta$, and also, for every

$\eta > 0$, there exists a $\Delta_0\,(> 0)$, such that

$$L_{J,21}^{(F)}(\phi, \Delta) - \eta \leqq L_J^{(F)}(\phi, \Delta) \leqq L_{J,12}^{(F)}(\phi, \Delta) + \eta\,, \qquad \text{for all } \Delta \leqq \Delta_0,\ \phi \in I\,. \tag{4.6}$$

For $\phi = \frac{1}{2}$, we note that $\{T_m(\Delta/2),\ m \geqq n_0(\Delta)\}$ has the same distribution as of $\{T_m(0),\ m \geqq n_0(\Delta)\}$, under $\theta = 0$. From the results of Sen and Ghosh (1973a), it follows that as $n \to \infty$, under $\theta = 0$,

$$\nu^{-1}T_n(0) = W(n) + o(n^{\frac{1}{2}}) \quad \text{a.s.}\,, \tag{4.7}$$

where $\{W(t),\ t \geqq 0\}$ is a standard Wiener process. Consequently, on using (4.7), standard results on the boundary crossing problems for Wiener processes, and a few routine steps, we obtain that

$$\lim_{\Delta\to 0} L_{J,ij}^{(F)}(\tfrac{1}{2}, \Delta) = [\log A_{\varepsilon,i}]/[\log A_{\varepsilon,i} - \log B_{\varepsilon,j}]\,, \qquad \text{for } i, j = 1, 2\,. \tag{4.8}$$

Since ε and η are arbitrary, from (4.4), (4.6) and (4.8), it follows that

$$\lim_{\Delta\to 0} L_J^{(F)}(\tfrac{1}{2}, \Delta) = (\log A)/[\log A - \log B]\,. \tag{4.9}$$

For $\phi \neq \frac{1}{2}$, under $\theta = \phi\Delta$, $\{T_m(\Delta/2),\ m \geqq n_0(\Delta)\}$ has the same distribution as that of $\{T_m((\frac{1}{2} - \phi)\Delta),\ m \geqq n_0(\Delta)\}$ under $\theta = 0$. We consider the following lemma.

LEMMA 4.2. *Under $\theta = 0$, (2.1) and (3.2), for every $\varepsilon > 0$, there exist a positve $c(\varepsilon)$ and an n_ε, such that for every $n \geqq n_\varepsilon$.*

$$P\{n^{-\frac{1}{2}}|T_n((\tfrac{1}{2} - \phi)\Delta) - n\xi((\tfrac{1}{2} - \phi)\Delta) - W_n(\phi, \Delta)| > \varepsilon\} \leqq c(\varepsilon)n^{-1-\delta}\,, \qquad \delta > 0\,, \tag{4.10}$$

where $\xi(\cdot)$ is defined by (4.2), $W_n(\phi, \Delta) = \sum_{i=1}^n Z_i(\phi, \Delta)$, and the $Z_i(\phi, \Delta)$ are i.i.d. rv'*s with mean 0 and a finite variance $\nu^2(\phi, \Delta)$, where*

$$\lim_{\Delta\to 0} \nu^2(\phi, \Delta) = \nu^2\,, \qquad \textit{for all } \phi \in I\,. \tag{4.11}$$

The proof of the lemma is sketched in the appendix.

Note that by (4.2), we have

$$\lim_{\Delta\to 0}\{\Delta^{-1}\xi((\tfrac{1}{2} - \phi)\Delta)\} = -(\tfrac{1}{2} - \phi)C(F)\,, \qquad \text{for all } \phi \in I\,. \tag{4.12}$$

Upon noting that the standard results of Wald (1947) are applicable for the sequence $\{W_n(\phi, \Delta) + n\xi((\frac{1}{2} - \phi)\Delta),\ n \geqq n_0(\Delta)\}$, we obtain from (4.10), (4.11), (4.12) and a few standard steps that

$$\lim_{\Delta\to 0} L_{J,ij}^{(F)}(\phi, \Delta) = [(A_{\varepsilon,i})^{1-2\phi} - 1]/[(A_{\varepsilon,i})^{1-2\phi} - (B_{\varepsilon,j})^{1-2\phi}] \qquad \text{for } \phi\,(\neq \tfrac{1}{2}) \in I\,, \tag{4.13}$$

and $i, j = 1, 2$. Again, as ε and η are arbitrary, by (4.4), (4.6) and (4.13), it follows that for every $\phi\,(\neq \frac{1}{2}) \in I$,

$$\lim_{\Delta\to 0} L_J^{(F)}(\phi, \Delta) = [A^{1-2\phi} - 1]/[A^{1-2\phi} - B^{1-2\phi}]\,, \tag{4.14}$$

and the proof of Theorem 3.2 is complete.

For the proof of Theorem 3.3, first note that $N_J(\Delta) \geqq n_0(\Delta)$, so that

$$E_\phi[N_J(\Delta)] = \Delta^2 n_0(\Delta) P\{N_J(\Delta) \geqq n_0(\Delta)\} + \Delta^2 \textstyle\sum_{n \geqq n_0(\Delta)} P_\phi\{N_J(\Delta) > n\}\,, \tag{4.15}$$

where by (3.4), $\Delta^2 n_0(\Delta) \to 0$ as $\Delta \to 0$. Also,

$$\begin{aligned} P_\phi\{N_J(\Delta) > n\} &= P_\phi\{N_J(\Delta) > n, |C_n^*/C(F) - 1| \leqq \varepsilon\} \\ &\quad + P_\phi\{N_J(\Delta) > n, |C_n^*/C(F) - 1| > \varepsilon\} \\ &= P_\phi\{N_J(\Delta) > n, |C_n^*/C(F) - 1| \leqq \varepsilon\} \\ &\quad + O(n^{-s})\,, \qquad \text{by Lemma 4.1,} \end{aligned} \tag{4.16}$$

where we let $s > 1$. Thus, using (4.15), (4.16) and the definitions of $N_J^{(11)}(\Delta)$ and $N_J^{(22)}(\Delta)$, made after (4.4), it suffices to show that for every $\varepsilon > 0$ there exists an $\eta > 0$, such that

$$\lim_{\Delta \to 0} |\Delta^2 E_\phi[N_J^{(ii)}(\Delta)] - \psi(\phi, \tau)| < \eta\,, \qquad \text{for } i = 1, 2 \text{ and } \phi \in I\,. \tag{4.17}$$

Let us define $E_\phi T_n(\Delta/2) = \mu(n, \phi, \Delta)$, $n \geqq 1$, $\Delta > 0$ and $\phi \in I$. Then, we have the following.

LEMMA 4.3. *For every $\phi\ (\neq \frac{1}{2}) \in I$, as $\Delta \to 0$,*

$$E_\phi[\mu(N_J^{(ii)}(\Delta), \phi, \Delta)] = \{E_\phi[N_J^{(ii)}(\Delta)]\}\{\xi((\tfrac{1}{2} - \phi)\Delta) + o(\Delta)\}\,.$$

The proof is sketched in the appendix.

Returning now to the proof of the theorem, we note that by virtue of Theorem 2.1 of Sen (1974), under $\theta = \phi\Delta$, $\{T_n(\Delta/2) - \mu(n, \phi, \Delta), n \geqq 1\}$ is a martingale with respect to an increasing sequence of σ-fields, and by definition, $n^{-1}T_n(d)$ is bounded for all n and every real d. For this martingale, the conditions of Lemma 1 of Chow, Robbins and Tiecher (1965) are easy to verify, and hence, we obtain that

$$E_\phi[T_{N_J^{(ii)}(\Delta)}(\Delta/2)] = E_\phi\{\mu(N_J^{(ii)}(\Delta), \phi, \Delta)\}\,, \qquad \text{for all } \phi \in I \text{ and } i = 1, 2\,. \tag{4.18}$$

On the other hand, neglecting the excess over the boundaries (permissible for $\Delta \to 0$), $T_{N_J^{(ii)}(\Delta)}(\Delta/2)$ can only assume the two values $b\nu^2[1 + (-1)^i\varepsilon]/\Delta C(F)$ and $a\nu^2[1 + (-1)^i\varepsilon]/\Delta C(F)$ with respective probabilities $L_{J,ii}^{(F)}(\phi, \Delta)$ and $1 - L_{J,ii}^{(F)}(\phi, \Delta)$, for $i = 1, 2$. Hence, as $\Delta \to 0$,

$$\begin{aligned} &\Delta E_\phi[T_{N_J^{(ii)}(\Delta)}(\Delta/2)] \\ &\quad \to [1 + (-1)^i\varepsilon]\nu^2[bP(\phi) + a\{1 - P(\phi)\}]/C(F)\,, \qquad \text{by (3.6).} \end{aligned} \tag{4.19}$$

Also, by (4.12) and Lemma 4.3, for $\phi \neq \frac{1}{2}$, as $\Delta \to 0$,

$$\begin{aligned} &\Delta^{-1} E_\phi[N_J^{(ii)}(\Delta)][\xi((\tfrac{1}{2} - \phi)\Delta) + o(\Delta)] \\ &\quad \to -(\tfrac{1}{2} - \phi)C(F)E_\phi[N_J^{(ii)}(\Delta)] + o(1)\,. \end{aligned} \tag{4.20}$$

From (4.19), (4.20) and the definition of $\psi(\phi, \tau)$ in (3.8), we get that as $\Delta \to 0$,

$$\Delta^2 E_\phi[N_J^{(ii)}(\Delta)] \to [1 + (-1)^i\varepsilon]\psi(\phi, \tau)\,, \qquad \text{for } \phi\ (\neq \tfrac{1}{2}) \in I\,. \tag{4.21}$$

Since ε is arbitrary, (4.17) holds and the theorem is proved for $\phi\ (\neq \frac{1}{2}) \in I$.

The above proof fails when $\phi = \frac{1}{2}$, as then Lemma 4.3 and (4.20) may not be used. However, noting that $P(\phi)$, defined by (3.6), has a continuous derivative $P'(\phi)$ at $\phi = \frac{1}{2}$, and considering a sequence of ϕ-values, say, $\frac{1}{2} \pm \varepsilon_r$, where $\varepsilon_r \to 0$ as $r \to \infty$, and using the above proof, we obtain by the L'Hôspital rule that

$$(4.22) \qquad \lim_{\Delta\to 0} \{\Delta^2 E_{\frac{1}{2}}[N_J(\Delta)]\} = -\tau^2 P'(\tfrac{1}{2}) \log AB^{-1}.$$

Hence, the proof of the theorem is complete.

5. ARE results. We shall compare the proposed SROT with the normal theory SLRT as well as the Wald SPRT. When F in (2.1) is assumed to be normal with an unknown variance σ^2, the Bartlett-Cox SLRT is based on the stopping variable $N_M(\Delta)$ defined below. Let $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$ and $s_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$, $n \geqq 2$. Then $N_M(\Delta)$ is the least positive integer ($\geqq n_0(\Delta)$) for which the inequality

$$(5.1) \qquad b s_m^2 < m\Delta(\bar{X}_m - \Delta/2) < a s_m^2 \qquad (m \geqq n_0(\Delta))$$

is vitiated; if for $N_M(\Delta) = n$, $n\Delta(\bar{X}_n - \Delta/2) \leqq b s_n^2$ (or $\geqq a s_n^2$), we accept H_0 (or H_1). We are interested in the properties of the test when the underlying F is not necessarily normal. Since the sample mean is a particular case of U-statistic or von Mises' differentiable statistical function for which asymptotic sequential tests are studied in detail by Sen (1973), omitting the details of the derivations, we may state the following results. First, if F possesses a finite second moment, the sequential test based on $N_M(\Delta)$ terminates with probability one. Second, by virtue of the a.s. behavior results of sample partial sums by Strassen (1967), under the same condition, the asymptotic OC function of the test based on $N_M(\Delta)$ agrees with (3.6). That is, like the SROT, this test is ADF (asymptotically distribution-free), but for the class of df's with finite second moment. For the ASN, we need to assume that F has a finite absolute moment of order r for some $r > 4$. Then, from the results of Sen (1973), it follows that

$$(5.2) \qquad \lim_{\Delta\to 0} \{\Delta^2 E_\phi[N_M(\Delta)]\} = \psi(\phi, \sigma), \qquad \phi \in I,$$

where $\psi(\phi, \tau)$ is defined by (3.8) and σ replaces τ.

We proceed to compare the SROT and the normal theory SLRT when the underlying df is not necessarily normal. For two sequential procedures R and Q for testing H_0 vs. H_1 in (2.2), if $N_R(\Delta)$ and $N_Q(\Delta)$ be the corresponding stopping variables, and if both have the same limiting OC function (specified by (3.6)), then the ARE of Q with respect to R when $\theta = \phi\Delta$ and F is the true df is given by

$$(5.3) \qquad \begin{aligned} e(Q, R, \phi, F) &= \lim_{\Delta\to 0} \{E_\phi N_R(\Delta)/E_\phi N_Q(\Delta)\} \\ &= \lim_{\Delta\to 0} \{[\Delta^2 E_\phi N_R(\Delta)]/[\Delta^2 E_\phi N_Q(\Delta)]\}. \end{aligned}$$

Thus, from Theorem 3.3 and (5.2), the ARE of the proposed SROT with respect to the normal theory SLRT when the actual df is F and $\theta = \phi\Delta$ is equal to

$$(5.4) \qquad e(J, M, \phi, F) = \psi(\phi, \sigma)/\psi(\phi, \tau) = \sigma^2/\tau^2 = \sigma^2 C^2(F)/\nu^2, \qquad \text{for all } \varepsilon \in I.$$

Now, (5.4) agrees with the Pitman-ARE of the general rank order test for location with respect to the t-test (cf. [6]). In particular, when $J(u) = u$: $0 \leqq u \leqq 1$, i.e., we use the Wilcoxon signed rank statistics for the SROT, (5.4) reduces to

$$e(W, M, \phi, F) = e(W, M, F) = 12\sigma^2(\textstyle\int_{-\infty}^{\infty} f^2(x)\,dx)^2 , \tag{5.5}$$

which equals to $3/\pi$ for normal F, is $\geqq 0.864$ for all F for which $\sigma^2 < \infty$, and exceeds one for many non-normal F, including the class of df's with "heavy tails." Again, when g in (2.3) is taken as the normal density, i.e., we use the normal scores statistics for the SROT, (5.4) is bounded below by 1, where the lower bound is attained iff F is normal. *This clearly indicates the asymptotic supremacy of the* SROT *based on the normal scores statistics over the usual normal theory* SLRT *when the underlying F is not necessarily normal.*

Let us now consider the Wald SPRT when the form of F (apart from θ) is specified. We denote by $h(x) = -f'(x)/f(x)$, $-\infty < x < \infty$, and assume that (i) $h(x)$ is square integrable (with respect to the Lebesgue measure), and (ii) $h'(x)$ is uniformly continuous in x $(-\infty < x < \infty)$. Then

$$I(f) = \textstyle\int_{-\infty}^{\infty} [f'(x)/f(x)]^2\,dF(x) = \int_{-\infty}^{\infty} h'(x)\,dF(x) < \infty . \tag{5.6}$$

The SPRT is based on the stopping variable $N_\lambda(\Delta)$, defined to be the smallest positive integer (m) for which the following inequality is vitiated:

$$b < \lambda_m < a \qquad \text{where} \quad \lambda_m = \textstyle\sum_{i=1}^{m} Z_i(\Delta) ; \tag{5.7}$$

$$Z_i(\Delta) = \log\{f(X_i - \Delta)/f(X_i)\} , \qquad i \geqq 1 ; \tag{5.8}$$

if for $N_\lambda(\Delta) = n$, λ_n is $\leqq b$ (or $\geqq a$), we accept H_0 (or H_1). Since the $Z_i(\Delta)$ are i.i.d. rv, and by (5.8), $Z_i(\Delta) = \log\{f(X_i - \Delta/2 - \Delta/2)\} - \log\{f(X_i - \Delta/2 + \Delta/2)\} = h(X_i - \Delta/2) + \frac{1}{8}\Delta^2[h'(X_i - \Delta/2 - \gamma_1\Delta/2) - h'(X_i - \Delta/2 + \gamma_2\Delta/2)]$ (where $0 < \gamma_1$, $\gamma_2 < 1$), by the uniform continuity of $h'(x)$, we have under $\theta = \phi\Delta$, as $\Delta \to 0$,

$$\begin{aligned} E_\phi Z_i(\Delta) &= \Delta^2(\phi - \tfrac{1}{2})I(f) + o(\Delta^2) \qquad \text{and} \\ V_\phi[Z_i(\Delta)] &= \Delta^2 I(f) + o(\Delta^2) . \end{aligned} \tag{5.9}$$

Hence, based on the well-known Wiener process approximations to $\{\lambda_n - E\lambda_n, n \geqq 1\}$ (viz., Theorem 4.4 of Strassen (1967)), it follows by using the results of Section 3 of Dvoretzky, Kiefer and Wolfowitz (1953) that the OC function of the SPRT has the same limiting (as $\Delta \to 0$) form (3.6). As such, by (5.9) and the fundamental result of Wald (1947, page 53), we readily obtain that

$$\lim_{\Delta\to 0} [\Delta^2 E_\phi\{N_\lambda(\Delta)\}] = \psi(\phi, [I(f)]^{-\frac{1}{2}}) . \tag{5.10}$$

Therefore by Theorem 3.3 and (5.10), the ARE of the SROT with respect to the SPRT is

$$\begin{aligned} e(J, \lambda, \phi, F) &= \psi(\phi, [I(f)]^{-\frac{1}{2}})/\psi(\phi, \tau) = [I(f)\tau^2]^{-1} \\ &= C^2(F)/[\nu^2 I(f)] = \rho^2 \\ &= (\textstyle\int_0^1 J(u)\gamma(u)\,du)^2/[(\int_0^1 J^2(u)\,du)(\int_0^1 \gamma^2(u)\,du)] , \end{aligned} \tag{5.11}$$

where $\gamma(u)$ is defined by (2.3) with g replaced by f. Thus, when $g = f$, i.e.,

$\rho = 1$, (5.11) equals to 1, so that if the assumed density g in (2.3) has the same form as of the true density f (apart from location or scale changes), the proposed SROT based on $N_J(\Delta)$ shares asymptotically (as $\Delta \to 0$) the optimality of SPRT. As a result, the SROT based on the normal scores and Wilcoxon signed rank statistics are respectively asymptotically optimal when the underlying df F are normal and logistic.

6. Some additional remarks. There are quite a few research papers in the area of one and two sample sequential rank tests ([3], [4], [5], [13], [14], [15], [23], [26], [27]). Most of these tests, however, are based on Lehmann alternatives, and as a result, do not apply for the location problem. Besides, with the exception of the termination with probability one and finiteness of the moment generating function of the stopping variables, other characteristics like the OC and ASN have not been adequately studied. For the two sample location problem, Hall (1969) has suggested an asymptotic sequential test based on the Wilcoxon statistics. His procedure achieves asymptotically (as $\Delta \to 0$) the prescribed strength (α, β). On the contrary, in the specification of the alternative hypothesis, it involves a functional of the unknown distribution function, and thereby, demands its knowledge; this, however, does not appear to be very realistic.

Albert (1966) has considered an alternative sequential test for H_0 vs. H_1 in (2.2). His procedure is based on the theory for the dual problem of the bounded length confidence interval for the parameter under test, and has the strength $\alpha(\Delta)$, $\beta(\Delta)$, with $\alpha(\Delta) \leqq \beta(\Delta)$. A closely related test of asymptotic strength (α, β) may be formulated as follows. Define $\theta^*_{L,n}$ and $\theta^*_{U,n}$ as in (2.8) and $C_n{}^*$ as in (2.9). Consider a sequential procedure where the stopping variable N is the first positive integer n ($\geqq n_0(\Delta)$) for which

$$\theta^*_{U,n} - \theta^*_{L,n} \leqq \Delta\tau_\varepsilon/(\tau_\alpha + \tau_\beta)\,; \tag{6.1}$$

accept H_1 if $N^{-\frac{1}{2}}T_N(0) \geqq \nu\tau_\alpha$, and accept H_0, otherwise. In (6.1), ε is defined as in (2.7) and τ_α is the upper $100\alpha\%$ point of the standard normal df. Using Lemma 4.1 and the weak convergence of one-sample rank order statistics (viz., Sen (1974)), it follows by standard arguments that the procedure in (6.1) as well as the procedure of Albert (1966) are asymptotically consistent in the sense that they have asymptotically (as $\Delta \to 0$) the strength (α, β). However, they suffer from the drawback that their ASN are the same as of the corresponding fixed sample size procedures had $C(F)$ or σ^2 been known. In general, these ASN are considerably higher than the ASN of the SROT or the SPRT, and hence, these tests are not generally efficient.

7. Appendix. (i) *The proof of Lemma* 4.2. Let $\{Z_i, i \geqq 1\}$ be i.i.d. rv with df $\Pi(z)$, $-\infty < z < \infty$, and let

$$\mu^* = 2 \int_0^\infty J(\Pi^*(z))\, d\Pi(z) - \int_0^1 J(u)\, du$$

where $\Pi^*(z) = \Pi(z) - \Pi(-z)$, $z \geqq 0$. Based on $Z_1, \cdots, Z_n$, we define T_n as in (2.4) with $d = 0$. Then, it follows from the results of Sen and Ghosh (1973 b) that under the original conditions of Chernoff and Savage (1958), $n^{-\frac{1}{2}}(T_n - n\mu^*) = n^{-\frac{1}{2}}[\sum_{i=1}^n B(Z_i)] + o(1)$ a.s., where

$$(7.1) \qquad B(Z_i) = \{J(\Pi^*(|Z_i|)) \operatorname{sgn} Z_i - EJ(\Pi^*(|Z_i|)) \operatorname{sgn} Z_i\} + \int_{-\infty}^{\infty} \operatorname{sgn} x[u(|x| - |Z_i|) - \Pi^*(|x|)]J'(\Pi^*(|x|))\, d\Pi(x) .$$

Since our (3.2) is stronger than the first derivative condition of Chernoff and Savage (1958), their second derivative condition is not needed for our purpose. Also, noting that everywhere in Sen and Ghosh (1973b), the Borel–Cantelli lemma has been used to establish the a.s. convergence, we may virtually, by repeating their steps, show that for every $\varepsilon > 0$, there exist a positive $c(\varepsilon)$ and an n_ε, such that for $n \geqq n_\varepsilon$,

$$(7.2) \qquad P\{n^{-\frac{1}{2}}|T_n - n\mu^* - \textstyle\sum_{i=1}^n B(Z_i)| > \varepsilon\} \leqq c(\varepsilon)n^{-1-\delta} , \qquad \delta > 0 .$$

Let us now take $Z_i = X_i - \Delta/2$, so that under $\theta = \phi\Delta$, $\Pi(z) = F(z + (\frac{1}{2} - \phi)\Delta)$. Therefore, $\mu^* = \xi((\frac{1}{2} - \phi)\Delta)$ as defined by (4.2). Also, by (7.1), $V[B(Z_i)] < \infty$, for all $\phi \in I$ and $\Delta > 0$. Finally, noting that $\lim_{\Delta\to 0} F(x + (\frac{1}{2} - \phi)\Delta) = F(x)$, $-\infty < x < \infty$, and $\phi \in I$, it readily follows by routine steps that $\lim_{\Delta\to 0} V[B(Z_i)] = \nu^2 = \int_0^1 J^2(u)\, du$. □

(ii) *The proof of Lemma* 4.3. With the definition of Z_i and T_n as in the proof of Lemma 4.2, for every $n \geqq 1$, let

$$(7.3) \qquad h_{n,r} = 2\binom{n-1}{r-1} \int_0^\infty [\Pi^*(z)]^{r-1}[1 - \Pi^*(z)]^{n-r}\, d\Pi(z) - 1/n , \qquad \text{for } r = 1, \cdots, n ;$$

$$(7.4) \qquad a_n = \textstyle\sum_{r=1}^n J_n(r/(n+1))h_{n,r} , \qquad n \geqq 1 .$$

Then, by Theorem 2.1 of Sen (1974), we have

$$(7.5) \qquad n\mu_n^* = \textstyle\sum_{k=1}^n a_k .$$

Since $\binom{n-1}{r-1} \int_0^1 u^{r-1}(1-u)^{n-r}\, du = n^{-1}$, we may rewrite $h_{n,r} = \binom{n-1}{r-1} \int_0^\infty [\Pi^*(z)]^{r-1} \cdot [1 - \Pi^*(z)]^{n-r}\, d[2\Pi(z) - 1 - \Pi^*(z)]$. Hence, integrating by parts and using the fact that under $\theta = \phi\Delta$, $2\Pi(z) - 1 - \Pi^*(z) = (\frac{1}{2} - \phi)\Delta[\pi(z) - \pi(-z)] + o(\Delta)$, where $\pi(z) = \Pi'(z)$, it follows by some standard steps that $a_n = \xi((\frac{1}{2} - \phi)\Delta) + o(\Delta)$ for $n \geqq n_0(\Delta)$. As a result, by (7.5),

$$(7.6) \qquad \mu_n^* = \xi((\tfrac{1}{2} - \phi)\Delta) + o(\Delta) \qquad \text{for all } n \geqq n_0(\Delta) .$$

Since $\phi\ (\neq \frac{1}{2}) \in I$, by (4.18) and (4.19), the existence of $E_\phi\, \mu(N_J^{(i)}(\Delta), \phi, \Delta)$ is insured, for $i = 1, 2$. Hence, we have on using (4.12) that for $\phi\ (\neq \frac{1}{2}) \in I$,

$$(7.7) \qquad \begin{aligned} E_\phi[\mu(N_J^{(i)}(\Delta), \phi, \Delta)] &= \textstyle\sum_{n=n_0(\Delta)}^\infty P_\phi\{N_J^{(i)}(\Delta) = n\}\mu(n, \phi, \Delta) \\ &= \textstyle\sum_{n=n_0(\Delta)}^\infty nP_\phi\{N_J^{(i)}(\Delta) = n\}\{\xi((\tfrac{1}{2} - \phi)\Delta) + o(\Delta)\} \\ &= E_\phi[N_J^{(i)}(\Delta)]\{\xi((\tfrac{1}{2} - \phi)\Delta) + o(\Delta)\} . \qquad \square \end{aligned}$$

8. Acknowledgment. Thanks are due to the referee for his useful comments on the paper.

REFERENCES

[1] ALBERT, A. (1966). Fixed size confidence ellipsoids for regression parameters. *Ann. Math. Statist.* **37** 1602–1630.

[2] BARTLETT, M. S. (1946). The large sample theory of sequential tests. *Proc. Cambridge Philos. Soc.* **42** 239–244.

[3] BERK, R. H. and SAVAGE, I. R. (1968). The information in rank order and the stopping time of some associated SPRT's. *Ann. Math. Statist.* **39** 1661–1674.

[4] BRADLEY, R. A., MARTIN, D. C. and WILCOXON, F. (1965). Sequential rank tests, I. Monte Carlo studies of the two-sample procedure. *Technometrics* **7** 463–483.

[5] BRADLEY, R. A., MARTIN, S. D. and WILCOXON, F. (1966). Sequential rank tests, II. A modified two-sample procedure. *Technometrics* **8** 615–623.

[6] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[7] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed width sequential confidence interval for the mean. *Ann. Math. Statist.* **36** 457–463.

[8] CHOW, Y. S., ROBBINS, H. and TEICHER, H. (1965). Moments of randomly stopped sums. *Ann. Math. Statist.* **36** 789–799.

[9] COX, D. R. (1963). Large sample sequential tests of composite hypotheses. *Sankhyā Ser. A* **25** 5–12.

[10] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1953). Sequential decision problems for processes with continuous time parameter: testing hypotheses. *Ann. Math. Statist.* **24** 254–264.

[11] HALL, W. J. (1969). A sequential Wilcoxon test (abstract). *Ann. Math. Statist.* **40** 1879.

[12] JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

[13] MILLER, R. G. JR. (1970). A sequential signed rank test. *J. Amer. Statist. Assoc.* **65** 1554–1561.

[14] MILLER, R. G. JR. (1972). Sequential rank tests—One sample case. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 97–108. Univ. of California Press.

[15] SAVAGE, I. R. and SETHURAMAN, J. (1966). Stopping time of a rank order sequential probability ratio test based on Lehmann alternatives. *Ann. Math. Statist.* **37** 1154–1160.

[16] SEN, P. K. (1966). On a distribution-free method of estimating the asymptotic efficiency of a class of nonparametric tests. *Ann. Math. Statist.* **37** 1759–1770.

[17] SEN, P. K. (1970). On some convergence properties of one-sample rank order statistics. *Ann. Math. Statist.* **41** 2140–2143.

[18] SEN, P. K. (1973). Asymptotic sequential tests for regular functionals of distribution functions. *Theor. Probability Appl.* **18** 235–249.

[19] SEN, P. K. (1974). The invariance principle for one sample rank order statistics. *Ann. Statist.* **2** 49–62.

[20] SEN, P. K. and GHOSH, M. (1971). On bounded length sequential confidence intervals based on one sample rank order statistics. *Ann. Math. Statist.* **42** 189–203.

[21] SEN, P. K. and GHOSH, M. (1973 a). A law of iterated logarithm for one-sample rank order statistics and an application. *Ann. Statist.* **1** 568–576.

[22] SEN, P. K. and GHOSH, M. (1973 b). A Chernoff–Savage representation of rank-order statistics for stationary ϕ-mixing processes. *Sankhyā Ser. A* **35** 157–172.

[23] SETHURAMAN, J. (1970). Stopping time of a rank order sequential probability ratio test based on Lehman alternatives—II. *Ann. Math. Statist.* **41** 1322–1333.

[24] STRASSEN, V. (1967). Almost sure behavior of sample sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 315–343. Univ. of California Press.

[25] WALD, A. (1947). *Sequential Analysis.* Wiley, New York.

[26] WEED, H. D. JR. and BRADLEY, R. A. (1971). Sequential one sample grouped signed rank tests for symmetry. *J. Amer. Statist. Assoc.* **66** 321–326.

[27] WILCOXON, F., RHODES, L. J. and BRADLEY, R. A. (1963). Two sequential two-sample grouped rank tests with applications to screening experiments. *Biometrics* **19** 58–84.

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA 27514

INDIAN STATISTICAL INSTITUTE
CALCUTTA, INDIA