ASYMPTOTIC SEQUENTIAL TESTS FOR REGULAR FUNCTIONALS OF DISTRIBUTION FUNCTIONS

PRANAB KUMAR SEN

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 740

April 1971

ASYMPTOTIC SEQUENTIAL TESTS FOR REGULAR FUNCTIONALS OF DISTRIBUTION FUNCTIONS*

PRANAB KUMAR SEN

Summary. The theory of asymptotic sequential likelihood ratio tests for composite hypotheses, developed by Bartlett [3] and Cox [6], among others, is extended here to cover a broad class of regular functionals of distribution functions. Various properties of the proposed sequential tests are studied and compared with those of some alternative procedures. A few applications are sketched.

1. Introduction. Consider a sequence $\{X_1, X_2, \ldots\}$ of independent and identically distributed random vectors (i.i.d.r.v.) having a $p(\geq 1)$-variate distribution function (d.f.) $F(x)$, $x \varepsilon R^p$, the p-dimensional Euclidean space. Assuming that the form of F is specified and it involves only a set of unknown parameters, Bartlett [3] and Cox [6] considered large sample sequential likelihood ratio tests (SLRT) for a parameter while treating the others as nuisance parameters. Basically, these procedures are quasi-sequential; starting with an initial sample (of at least moderately large size), observations are drawn sequentially until a terminal decision is reached.

In a broad class of nonparametric problems, the form of F is not known; it is only assumed that F belongs to some suitable family $\mathcal{F}$ of d.f. on R^p. In this setup, (estimable) parameters are defined as functionals of F, defined on $\mathcal{F}$; we may refer to Halmos [7], von Mises [9] and Hoeffding [8], for details.

---

*Work supported by the National Institutes of Health, Grant GM-12868.

The fact that the functional form of F is unknown makes it difficult to adapt the probability ratio or the likelihood ratio principle underlying the classical sequential testing theory. Nevertheless, it is shown here that for testing a null hypothesis H_0: $\theta(F)=\theta_0$ against H_1: $\theta(F)=\theta_1$, where $\theta(F)$ is a regular functional (estimable parameter) of F, by the same motivation as in Bartlett [3] and Cox [6], sequential tests can be constructed, which for every pair (θ_0,θ_1): $\theta_0 \neq \theta_1$, of parameters, terminates with probability one, and which for θ_1 approaching to θ_0, enjoy all the basic properties of the SLRT.

Along with the preliminary notions, the proposed tests are described in section 2. Sections 3 and 4 deal with the properties of the proposed tests. In section 5, an alternative procedure based on the Anscombe [2] and the Chow and Robbins [5] theory of fixed-width (sequential) confidence intervals is considered and compared with the earlier ones. A few applications are briefly sketched in the last section.

2. Preliminary notions and the proposed tests. Consider a functional $\theta(F)$ defined on $\mathcal{F}$. $\theta(F)$ is said to have the degree $m(\geq 1)$, when m is the smallest positive integer for which a kernel $\phi(X_1,\ldots,X_m)$ unbiasedly estimates $\theta(F)$ i.e.,

$$\int\cdots\int_{R^{pm}} \phi(x_1,\ldots,x_m)dF(x_1)\ldots dF(x_m) = E\phi(X_1,\ldots,X_m) = \theta(F), \tag{2.1}$$

for all $F\varepsilon\mathcal{F}$, where we assume, without any loss of generality, that ϕ is symmetric in its arguments. For every $0\leq c\leq m$, define

$$\psi_c(x_1,\ldots,x_c) = E\phi(x_1,\ldots,x_c,X_{c+1},\ldots,x_m)-\theta(F), \ \psi_0 = 0; \tag{2.2}$$

$$\zeta_c(F) = E\psi_c^2(X_1,\ldots,X_c), \ \zeta_0(F) = 0. \tag{2.3}$$

It follows from the results of Hoeffding [8] that

$$0 \leq \zeta_c(F) \leq (c/d)\zeta_d \quad \text{for all} \quad 1 \leq c \leq d \leq m, \tag{2.4}$$

and, $\theta(F)$ is termed stationary of order zero whenever

$$\zeta_1(F) > 0 \quad \text{and} \quad \zeta_m(F) < \infty. \tag{2.5}$$

We desire to provide a sequential test for

$$H_0: \ \theta(F)=\theta_0 \quad \text{vs.} \quad H_1: \ \theta(F)=\theta_1=\theta_0+\Delta; \ \theta_0, \ \Delta \text{ known}. \tag{2.6}$$

Our proposed tests are based on the natural estimates of $\theta(F)$, considered earlier by von Mises [7] and Hoeffding [8], among others. For a sample $(X_1,\dots,X_n)$ of size n, define the empirical d.f.

$$F_n(x) = n^{-1} \sum_{i=1}^{n} c(x-X_i), \ x \in R^p, \tag{2.7}$$

where for a p-vector u, $c(u)=1$ if all the p coordinates of u are non-negative and $c(u)=0$, otherwise. Then, von Mises' differentiable statistical function $\theta(F_n)$ is defined as

$$\begin{aligned} \theta(F_n) &= \int\cdots\int_{R^{pm}} \phi(x_1,\dots,x_m)dF_n(x_1)\dots dF_n(x_m) \\ &= n^{-m} \sum_{i_1=1}^{n}\dots\sum_{i_m=1}^{n} \phi(X_{i_1},\dots,X_{i_m}), \end{aligned} \tag{2.8}$$

so that $\theta(F_n)$ is the corresponding regular functional of the empirical d.f. F_n. Note that $\theta(F_n)$ is not necessarily an unbiased estimator of $\theta(F)$ though F_n unbiasedly estimates F. Hoeffding [8] considered the unbiased estimator (U-statistic)

$$U_n = \binom{n}{m}^{-1} \sum_{C_{n,m}} \phi(X_{i_1},\ldots,X_{i_m}); \; C_{n,m} = \{1 \leq i_1 < \ldots < i_m \leq n\}. \tag{2.9}$$

We motivate our proposed sequential tests by the same principle underlying the asymptotic sequential likelihood ratio tests of Bartlett [3] and Cox [6]. If F is of specified form involving as parameters θ(under test) and δ(nuisance), these authors working with the ratio of the two maxima of the likelihood function under the two (composite) hypotheses H_o: $\theta=\theta_0$ vs. H_1: $\theta=\theta_1-\theta_0+\Delta$, θ_0, Δ known, were able to approximate the same, at the n-th stage, by

$$n\Delta I_{\theta\theta}(\hat{\theta}_n - \tfrac{1}{2}[\theta_0+\theta_1]), \tag{2.10}$$

where $\hat{\theta}_n$ is the (unrestricted) maximum likelihood estimator of θ, and $I_{\theta\theta}$ is the Fisher information for θ. The asymptotic reduction in (2.10) along with the normality and other properties of the maximum likelihood estimates enables us to claim that the SLRT has asymptotically (as $\Delta\to 0$) all the properties of the classical Wald [16] sequential probability ratio test (SPRT) where δ is assumed to be known.

In the nonparametric setup, U_n is the optimal (minimum variance) unbiased estimator of $\theta(F)$, and $\theta(F_n)$ is asymptotically equivalent, in probability, to U_n i.e., $n^{\frac{1}{2}}[U_n-\theta(F_n)]\overset{P}{\to}0$ as $n\to\infty$ (cf. [8]). Moreover, $n^{\frac{1}{2}}[U_n-\theta(F)]$ (or $n^{\frac{1}{2}}[\theta(F_n)-\theta(F)]$) is asymptotically normally distributed with zero mean and variance

$m^2\zeta_1(F)(>0)$, which may be consistently estimated as follows. For $n>m$, let

$$V_{n,i} = \binom{n-1}{m-1}^{-1} \sum\nolimits_i^* \phi(X_i, X_{i_2}, \ldots, X_{i_m}),\ 1\leq i\leq n, \tag{2.11}$$

where the summation $\sum_i^*$ extends over all possible $1\leq i_2<\ldots<i_m\leq n$ with $i_j\neq i$, $2\leq j\leq m$. Then, from Sen [10],

$$s_n^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n} [V_{n,i}-U_n]^2 \xrightarrow{P} \zeta_1(F) \text{ as } n\to\infty. \tag{2.12}$$

Thus, looking at (2.10) and the asymptotic SLRT, we are, at least heuristically, in favor of replacing $\hat{\theta}_n$ by U_n, $I_{\theta\theta}^{-1}$ by $m^2s_n^2$, and thereby, considering the following procedure.

Corresponding to our desired first and second kinds of error α, β (where we take $0<\alpha,\beta<\frac{1}{2}$), we consider two numbers $A(\leq(1-\beta)/\alpha)$ and $B(\geq\beta/(1-\alpha))$, so that $0<B<1<A<\infty$. Then, we start with an initial sample of size $n_0(=n_0(\Delta)$, moderately large for small $\Delta)$, and define a stopping variable $N(=N(\Delta))$ as the smallest positive integer $(\geq n_0)$ for which

$$m^2s_n^2 \log B < n\Delta[U_n-\tfrac{1}{2}(\theta_0+\theta_1)] < m^2s_n^2(\log A) \tag{2.13}$$

is vitiated. If for $N(\Delta)=N$, $n\Delta[U_n-\frac{1}{2}(\theta_0+\theta_1)] \leq m^2s_n^2 \log B$, re accept H_0: $\theta(F)=\theta_0$, while if $n\Delta[U_n-\frac{1}{2}(\theta_0+\theta_1)] \geq m^2s_n^2 \log A$, we accept H_1: $\theta(F)=\theta_1=\theta_0+\Delta$, $\Delta>0$.

A parallel sequential procedure based on $\theta(F_n)$ may be posed as follows. Let us define $V_{n,i}^* = n^{-(m-1)}\sum_{i_2=1}^{n}\ldots\sum_{i_m=1}^{n}\phi(X_i,X_{i_2},\ldots,X_{i_m}) = \int\cdots\int_{R^{p(m-1)}} \phi(X_i,x_2,\ldots,x_m)\, dF_n(x_2)\ldots dF_n(x_m)$, $1\leq i\leq n$, and let

$$(s_n^*)^2 = (n-1)^{-1} \sum_{i=1}^n [V_{n,i}^* - \theta(F_n)]^2. \tag{2.14}$$

Then, in the second procedure, we replace, in (2.13), U_n by $\theta(F_n)$ and s_n by s_n^*, while the rest remains the same. It will be seen later on that the two procedures have asymptotically (as $\Delta \to 0$) the same properties.

The heuristic proposal made above is justified theoretically in the next two sections.

3. Termination probability. We have the following theorem.

Theorem 3.1. Under (2.5), both the procedures terminate with probability one i.e., for every (fixed) $\theta(F)$ and $\Delta(>0)$, $P\{N(\Delta)>n|\theta(F)\} \to 0$ as $n\to\infty$.

Proof. By definition in (2.13), for every $n>n_0$,

$$P_\theta\{N(\Delta)>n\} = P_\theta\{\frac{m^2 s_n^2}{r\Delta} (\log B) < [U_r - \tfrac{1}{2}(\theta_0+\theta_1)] < \frac{m^2 s_r^2}{r\Delta} (\log A),\ n_0 \le r \le n\}$$

$$\le P_\theta\{m^2 s_n^2 (\log B)/n\Delta < [U_n - \tfrac{1}{2}(\theta_0+\theta_1)] < m^2 s_n^2 (\log A)/n\Delta\}, \tag{3.1}$$

where P_θ stands for the probability computed for $\theta(F)=\theta$. Now, we have three situations: (a) $\theta(F) > \frac{1}{2}(\theta_0+\theta_1)$, (b) $\theta(F) < \frac{1}{2}(\theta_0+\theta_1)$ and (c) $\theta(F) = \frac{1}{2}(\theta_0+\theta_1)$. We know that $\{U_n\}$ being a reverse martingale sequence (cf. Berk [4]) converges almost surely (a.s.) to $\theta(F)$ as $n\to\infty$. Also, Sproule [14] has strengthened (2.12) to $s_n^2 \to \zeta_1(F)$ a.s. as $n\to\infty$. Hence, $m^2 s_n^2(\log A)/n\Delta$ and $m^2 s_n^2(\log A)/n\Delta$ both a.s. converge to zero as $n\to\infty$; on the other hand, in cases (a) and (b), $[U_n - \frac{1}{2}(\theta_0+\theta_1)] \to [\theta(F) - \frac{1}{2}(\theta_0+\theta_1)] (\neq 0)$ a.s. as $n\to\infty$. Hence, (3.1) tends to 0 as $n\to\infty$. In case (c), $m^2 s_n^2(\log B)/(\Delta\sqrt{n})$ and $m^2 s_n^2(\log A)/(\Delta\sqrt{n})$ both a.s. converge to 0 as $n\to\infty$, whereas $n^{\frac{1}{2}}[U_n - \frac{1}{2}(\theta_0+\theta_1)]$ is asymptotically normally distributed with zero mean and a finite (positive) variance $m^2\zeta_1(F)$; hence, (3.1) again tends to 0 as $n\to\infty$. Finally, it follows from Sen [12] that $|\theta(F_n)-U_n|\to 0$ a.s. as $n\to\infty$, and hence, the proof for the second procedure follows on parallel lines. Q.E.D.

4. OC and ASN functions. For theoretical justifications, we now consider the asymptotic situation where $\Delta\to 0$; this is comparable to the asymptotic situation in (sequential) fixed-width confidence interval problems, treated in Chow and Robbins [5], and others. In practice, the results are valid whenever we have the access to choose small Δ. For $\Delta\to 0$, we concentrate ourselves to a range of possible values of $\theta(F)$ also contracting to 0 as $\Delta\to 0$. Specifically, we assume that $\theta(F)\varepsilon I_\Delta$ where

$$I_\Delta = \{\theta(F)=\theta_0+\phi\Delta:\ \ \phi\varepsilon I\},\ I = \{\phi:\ \ |\phi|<K\}, \tag{4.1}$$

where $K(>1)$ is a finite constant. We may remark that if $\theta(F)\notin I_\Delta$, then asymptotically the OC function will be either very close to zero or to one, and hence, will cease to be of any statistical interest. Further, looking at (2.13) and the a.s. convergence of (i) U_n to $\theta(F)$ and (ii) s_n^2 to $\zeta_1(F)$, we are confident that as $\Delta\to 0$, a terminal decision will not be reached (in probability) at an early stage. Consequently, there is no harm in letting $n_0(=n_0(\Delta))\to\infty$ as $\Delta\to\infty$, but at a slower rate as compared to the ASN of $N(\Delta)$. We shall see later on that the ASN of $N(\Delta)$ increases at a rate of Δ^{-2} as $\Delta\to 0$. Hence, we assume that

$$\lim_{\Delta\to 0} n_0(\Delta) = \infty \ \text{ but } \ \lim_{\Delta\to 0} \Delta^2 n_0(\Delta) = 0. \tag{4.2}$$

These assumptions are also implicitly needed to justify rigorously the procedure of Bartlett [3] and Cox [6].

Consider now a standard Brownian motion W_t: $0 \le t < \infty$, and denote by

$$P(\phi,\gamma) = P\left\{\begin{array}{l} W_t \text{ first crosses the line } \gamma^{-1}\log B + \gamma t(\tfrac{1}{2}-\phi) \\ \text{before crossing the line } \gamma^{-1}\log A + \gamma t(\tfrac{1}{2}-\phi):\quad 0 \le t < \infty \end{array}\right\}; \tag{4.3}$$

$$P(\phi) = \lim_{\gamma\to 0} P(\phi,\gamma):\ \phi\varepsilon I = \{\phi:\ |\phi| < K\}. \tag{4.4}$$

Let then the OC function of the sequential procedure in (2.13) be

$$L_\Delta(\phi) = P\{H_0:\ \theta(F)=\theta_0 \text{ is accepted } |\theta(F)=\theta_0+\phi\Delta\}. \tag{4.5}$$

Theorem 4.1. Under (2.5) and (4.2),

$$\lim_{\Delta\to 0} L_\Delta(\phi) = P(\phi) \text{ for all } \phi\varepsilon I; \tag{4.6}$$

$$\lim_{\Delta\to 0} L_\Delta(0) = 1-\alpha \text{ and } \lim_{\Delta\to 0} L_\Delta(1) = \beta. \tag{4.7}$$

Remark. The last equation specifies the asymptotic (as $\Delta\to 0$) consistency of the proposed sequential test.

Proof. For every $\varepsilon>0$, let us define

$$n_0^{(1)}(\Delta) = \max\{n_0(\Delta), [(\log AB^{-1})\varepsilon/\Delta^2]\},\ n_0^{(2)}(\Delta) = [\{(\log AB^{-1})/\varepsilon\Delta\}^2], \tag{4.8}$$

where [s] denotes the largest integer contained in s. Then, by (2.13),

$$P_\theta\{N(\Delta) \le n_0^{(1)}(\Delta)\}$$

$$= P_\theta\{\frac{m^2 s_r^2(\log B)}{\Delta r} < [U_r - \tfrac{1}{2}(\theta_0+\theta_1] < \frac{m^2 s_r^2(\log A)}{\Delta r},\ n_0(\Delta) \le r \le n^{(1)}(\Delta)\}. \tag{4.9}$$

Since, $s_n^2 \to \zeta_1(F)$ a.s. as $n\to\infty$ and (4.2) holds, we conclude that for every $\eta>0$, $P_\theta\{s_r^2 \le (1+\eta)\zeta_1(F),\ n_0(\Delta) \le r \le n^{(1)}(\Delta)\} \to 1$ as $\Delta\to 0$. Consequently, (4.9) is bounded above by

$$P_\theta\{|U_r-\tfrac{1}{2}(\theta_0+\theta_1)| \le \frac{m^2(1+\eta)}{\Delta r}\zeta_1(F)(\log AB^{-1}),\ n_0(\Delta)\le r\le n^{(1)}(\Delta)\} + \eta(\Delta), \quad (4.10)$$

where $\eta(\Delta)\to 0$ as $\Delta\to 0$. Now, for all $\theta(F)\varepsilon I_\Delta$, $|\theta(F)-\tfrac{1}{2}(\theta_0+\theta_1)| \le (K+1)\Delta$, and hence, for every $K(<\infty)$, we can always select an $\varepsilon>0$, such that $|\theta(F)-\tfrac{1}{2}(\theta_0+\theta_1)| \le (m^2/2\Delta r)\zeta_1(F)(1+\eta)\log AB^{-1}$ for all $r\le n_0^{(1)}(\Delta)$. [Note that for $r\le n_0^{(1)}(\Delta)$, $(r\Delta)^{-1} \ge \Delta(\varepsilon \log AB^{-1})^{-1}$ can be made large by proper choice of $\varepsilon>0$.] Hence, (4.10) is bounded above by

$$P_\theta\{|U_r-\theta(F)| \le \frac{m^2}{2\Delta r}(1+\eta)\zeta_1(F)(\log AB^{-1}),\ n_0(\Delta)\le r\le n_0^{(1)}(\Delta)\} + \eta(\Delta). \quad (4.11)$$

Now, using the reverse martingale property of U-statistics (cf. [4]), and thereby applying the Chow extension of the Hájek-Rènyi inequality (as in Sen [11], (3.6)), the first term of (4.11) can be easily shown to be bounded by

$$[4\Delta^2/m^2(1+\eta)(\log AB^{-1})\zeta_1(F)][Cm^2\zeta_m(F)\{n_0^{(1)}(\Delta)-n_0(\Delta)\}]$$

$$\le n_0^{(1)}(\Delta)[\zeta_m(F)/\zeta_1(F)][4C\Delta^2/(1+\eta)\log AB^{-1}]$$

$$\le C_0\varepsilon \le \varepsilon', \text{ where } C_0<\infty, \text{ independently of } \Delta, \varepsilon \text{ and } \eta. \quad (4.12)$$

Thus, for every $\delta>0$, there exists an $\varepsilon>0$, such that

$$\lim_{\Delta\to 0} P_\theta\{N(\Delta)<n_0^{(1)}(\Delta)\} < \tfrac{1}{2}\delta \quad \text{for all } \theta(F)\ \varepsilon I_\Delta. \quad (4.13)$$

Again, as in (3.1), for every $\theta(F)\ \varepsilon I_\Delta$,

$$P_\theta\{N(\Delta)>n_0^{(2)}(\Delta)\} \le P\{\frac{m(\log B)s_n}{\Delta n^{1/2}} < \frac{n^{1/2}[U_n-\tfrac{1}{2}(\theta_0+\theta_1)]}{ms_n} < \frac{m(\log A)s_n}{\Delta n^{1/2}}\}\Big|_{n=n_0^{(2)}(\Delta)}. \quad (4.14)$$

Now, $s_n^2 \to \zeta_1(F)$ a.s. as $n\to\infty$, and $n^{\frac{1}{2}}[U_n-\theta(F)]/m\zeta_1^{\frac{1}{2}}(F)$ is asymptotically normally distributed with zero mean and unit variance. Therefore, $n^{\frac{1}{2}}[U_n-\frac{1}{2}(\theta_0+\theta_1)]/ms_n$ is asymptotically normally distributed with mean $[\theta(F)-\frac{1}{2}(\theta_0+\theta_1)]/m[\zeta_1(F)]^{\frac{1}{2}}$ and unit variance. On the other hand, $m(\log AB^{-1})s_n/\Delta n^{\frac{1}{2}}$, for $n=n_0^{(2)}(\Delta)$, converges almost surely to $m\varepsilon[\zeta_1(F)]^{\frac{1}{2}}$, as $\Delta\to 0$. Hence, the right hand side of (4.14) is asymptotically bounded above by

$$(2\pi)^{-\frac{1}{2}}m\varepsilon[\zeta_1(F)]^{\frac{1}{2}} < \varepsilon', \tag{4.15}$$

which can be made smaller than $\frac{1}{2}\delta$ by proper choice of $\varepsilon(>0)$. This leads us to the following: for every $\theta(F)\varepsilon I_\Delta$, and for every $\delta>0$, there exists an $\varepsilon(>0)$, such that

$$\lim_{\Delta\to 0} P_\theta\{n_0^{(1)}(\Delta) \leq N(\Delta) \leq n_0^{(2)}(\Delta)\} \geq 1-\delta. \tag{4.16}$$

Consider now a stopping variable $N^*(\Delta)$ defined as the smallest positive integer $n(\geq n_0(\Delta))$ for which

$$m^2\zeta_1(F)(\log B)/\Delta < n[U_n-\tfrac{1}{2}(\theta_0+\theta_1)] < m^2\zeta_1(F)(\log A)/\Delta \tag{4.17}$$

is vitiated; if $n[U_n-\frac{1}{2}(\theta_0+\theta_1)] \leq m^2\zeta_1(F)(\log B)/\Delta$ accept H_0, and accept H_1 if $n[U_n-\frac{1}{2}(\theta_0+\theta_1)] \geq m^2\zeta_1(F)(\log A)/\Delta$. By the same technique as in above, it follows that (4.16) also holds for $N^*(\Delta)$. Further, $s_n^2\to\zeta_1(F)$ a.s. as $n\to\infty$, and hence, as $\Delta\to 0$, for all $n_0^{(1)}(\Delta)\leq n\leq n_0^{(2)}(\Delta)$, $|s_n^2-\zeta_1(F)|\to 0$, with probability one. Thus, if we denote by $L_\Delta^*(\phi)$, the OC function of the procedure in (4.17), then from (2.13), (4.17) and the above discussion, we conclude that

$$\lim_{\Delta\to 0} |L_\Delta(\phi)-L^*_\Delta(\phi)| = 0 \quad \text{for all} \quad \phi\varepsilon I. \tag{4.18}$$

Since for $n_0^{(1)}(\Delta)\leq n\leq n_0^{(2)}(\Delta)$, $n^{\frac{1}{2}}\Delta$ is bounded above by a positive constant $K_\varepsilon(<\infty)$, if we linearly interpolate $\Delta r[U_r-\theta(F)]/m[\zeta_1(F)]^{\frac{1}{2}}$ between two consecutive r, $n_0^{(1)}\leq r\leq n_0^{(2)}(\Delta)$, we may, after using the results in Sen [12], conclude that the process weakly converges to a Brownian movement process $\xi(t;\Delta)$, where $E\{\xi(t;\Delta)\} = 0, \forall t$ and $E\{\xi(s;\Delta)\xi(t;\Delta),\ s\leq t\} = s\Delta^2$, $0\leq s\leq t<\infty$. As such, if we let $\Delta' = \Delta/m[\zeta_1(F)]^{\frac{1}{2}}$, we have from (4.13), (4.17) and the above that

$$\lim_{\Delta\to 0} |L^*_\Delta(\phi)-P(\phi,\Delta')| = 0 \quad \text{for all} \quad \phi\varepsilon I, \tag{4.19}$$

and hence, by (4.4), (4.18) and (4.19)

$$\lim_{\Delta\to 0} L_\Delta(\phi) = \lim_{\Delta'\to 0} P(\phi,\Delta') = P(\phi) \quad \text{for all} \quad \phi\varepsilon I. \tag{4.20}$$

For testing a simple H_0: $\theta=\theta_0$ against a simple H_1: $\theta=\theta_1=\theta_0+\Delta$ (F known), consider the Wald sequential probability ratio test (SPRT) of strength (α,β). For small Δ, the excess over the boundaries can be neglected, so that $A = (1-\beta)/\alpha+O(\Delta)$, $B = \beta/(1-\alpha)+O(\Delta)$, and on denoting by $L^o_\Delta(\phi)$ the OC function of the SPRT, we have from Wald [16],

$$\lim_{\Delta\to 0} L^o_\Delta(0) = 1-\alpha \quad \text{and} \quad \lim_{\Delta\to 0} L^o_\Delta(1) = \beta. \tag{4.21}$$

On the other hand, the sequence of logarithm of Probability ratio forms a martingale sequence on which Theorem 4.4 of Strassen [15] yields the weak convergence to a Brownian movement process. Consequently, as in (4.19) and (4.20),

$$\lim_{\Delta\to 0} L_\Delta^o(\phi) = P(\phi) \quad \text{for all} \quad \phi\varepsilon I. \tag{4.22}$$

Thus, (4.7) follows directly from (4.20), (4.21) and (4.22). Q.E.D.

In passing, we may remark that by Sen [12], $n^{\frac{1}{2}}[\theta(F_n)-U_n]\to 0$ a.s. as $n\to\infty$ and $|s_n^2-(s_n^*)^2|\to 0$ a.s. as $n\to\infty$. Consequently, by some routine steps it follows that Theorem 4.1 remains valid for the alternative procedure based on $\theta(F_n)$ and s_n^*.

We now proceed to study the ASN of $N(\Delta)$. It is quite clear from (4.8) and (4.16) that as $\Delta\to 0$, $E_\theta\{N(\Delta)\}\to\infty$ for all $\theta(F)\varepsilon I_\Delta$. However, we shall see that under certain regularity conditions, $\Delta^2 E_\theta\{N(\Delta)\}$ converges to some limit (depending on $\theta(F)$ and F) as $\Delta\to 0$. This result is used in the next section to compare the asymptotic efficacy of the proposed procedures.

In addition to (2.5), we assume that for some $\delta>0$,

$$\nu(F) = E\{|\phi(X_1,\ldots,X_m)|^{4+\delta}\}<\infty \quad \text{for all} \quad F\varepsilon \mathcal{F}. \tag{4.23}$$

It may be remarked that in the classical SPRT, Wald [16], while computing the ASN assumed the existence of the moment generating function of the logarithm of the density ratio, and (4.23) is no more restrictive than that. Our next theorem relates to the rate of growth of the ASN when $\theta(F) \to \frac{1}{2}(\theta_0+\theta_1)$.

<u>Theorem 4.2</u>. <u>Under</u> (2.5), (4.2) <u>and</u> (4.23), <u>for every</u> $\phi(\neq\frac{1}{2})\varepsilon I$,

$$\lim_{\Delta\to 0}\{\Delta^2 E_\phi(N(\Delta))\} = \frac{m^2\zeta_1(F)\{P(\phi)\log B+[1-P(\phi)]\log A\}}{(\phi-\frac{1}{2})} = \psi(\phi), \tag{4.24}$$

<u>where</u> E_ϕ <u>stands for the expectation under</u> $\theta(F) = \theta_0+\phi\Delta$, $\phi\varepsilon I$.

Proof. We only consider the case of $\phi>\frac{1}{2}$; the case of $\phi<\frac{1}{2}$ follows similarly. For some arbitrarily small $\varepsilon(>0)$, we define

$$n_1(\Delta) = [\varepsilon\Delta^{-2}] \quad \text{and} \quad n_2(\Delta) = [m^2\zeta_1(F)(\log AB^{-1})\Delta^{-2}\varepsilon^{-1}(\phi-\tfrac{1}{2})^{-1}]. \tag{4.25}$$

Then, we have

$$\Delta^2 E_\phi N(\Delta) = \Delta^2[\textstyle\sum_{n\le n_1(\Delta)}+\sum_{n_1(\Delta)<n\le n_2(\Delta)}+\sum_{n>n_2(\Delta)} nP_\phi\{N(\Delta)=n\}], \tag{4.26}$$

where obviously,

$$\Delta^2 \textstyle\sum_{n\le n_1(\Delta)} nP_\phi\{N(\Delta)=n\}<\varepsilon \quad \text{for all} \quad \phi\varepsilon I. \tag{4.27}$$

Now, by using the fact that $\sum_{n>k} nP\{N=n\} = (k+1)P\{N>k\} + \sum_{k+1}^{\infty} P\{N>n\}$, we have

$$\Delta^2 \textstyle\sum_{n>n_2(\Delta)} nP_\phi\{N(\Delta)=n\} = \{n_2(\Delta)+1\}P_\phi\{N(\Delta)>n_2(\Delta)\} + \sum_{n_2(\Delta)+1}^{\infty} P\{N(\Delta)>n\}. \tag{4.28}$$

Let now $\eta(>0)$ be an arbitrarily small positive number. Then, upon noting that $(\Delta n)^{-1}m^2(\log AB^{-1})(1+\eta)\zeta_1(F)<\Delta\varepsilon(\phi-\frac{1}{2})$ for all $n\ge n_2(\Delta)$, and proceeding as in (4.14), we have for $n\ge n_2(\Delta)$,

$$\begin{aligned}
P_\phi\{N(\Delta)>n\} &\le P\{m^2(\log B)s_n^2<n\Delta[U_n-\tfrac{1}{2}(\theta_0+\theta_1)]<m^2(\log A)s_n^2\} \\
&\le P\{m^2(\log B)(1+\eta)\zeta_1(F)<n\Delta[U_n-\tfrac{1}{2}(\theta_0+\theta_1)]<m^2(\log A)(1+\eta)\zeta_1(F)\} \\
&\qquad + P\{s_n^2>(1+\eta)\zeta_1(F)\} \\
&\le P\{[U_n-\theta(F)]>(\phi-\tfrac{1}{2})\Delta(1-\varepsilon)\} + P\{s_n^2>(1+\eta)\zeta_1(F)\}.
\end{aligned} \tag{4.29}$$

Now, from Sen [13], for every $r \geq 2$, if $E|\phi|^r < \infty$,

$$E\{|[U_n - \theta(F)]|^r\} \leq C_r n^{-r/2}, \; C_r < \infty, \tag{4.30}$$

and therefore by the Markov inequality, under (4.23),

$$\begin{aligned} &P\{[U_n - \theta(F)] > C\phi - \tfrac{1}{2})\Delta(1-\varepsilon)\} \\ &\leq P\{|U_n - \theta(F)| > (\phi - \varepsilon)\Delta(1-\varepsilon)\} \\ &\leq C[(\phi - \tfrac{1}{2})\Delta(1-\varepsilon)]^{-4-\delta} n^{-2-\delta/2}, \; C < \infty. \end{aligned} \tag{4.31}$$

Again, it has been shown by Sproule [14] that s_n^2 in (2.12) can be expressed as a linear combination of several U-statistics, for each of which, moment of the order $2+\delta/2$ exists, for some $\delta > 0$ (implied by (4.23)). Hence, using (4.30) and the Markov inequality,

$$P\{s_n^2 > (1+\eta)\zeta(F)\} \leq C(\eta) \; n^{-1-\delta'}, \tag{4.32}$$

where $\delta' = \delta/4 > 0$ and $C(\eta) < \infty$ for every $\eta > 0$. From (4.25), (4.28), (4.29), (4.31) and (4.32), we have by routine computations that

$$\lim_{\Delta \to 0} \sum_{n > n_2(\Delta)} nP_\phi\{N(\Delta) = n\} < \varepsilon', \quad \text{for all} \quad \phi \varepsilon I, \tag{4.33}$$

where $\varepsilon'(>0)$ can be made arbitrarily small, depending on $\varepsilon(>0)$. Therefore, it suffices to show that for $\phi > \tfrac{1}{2}(\varepsilon I)$,

$$\lim_{\Delta\to 0}\{\Delta^2[\sum_{n_1(\Delta)+1}^{n_2(\Delta)} P_\phi\{N(\Delta)>n\}]\} = \psi(\phi). \tag{4.34}$$

For this, we define

$$Z_n = m\sum_{i=1}^{n} \phi_1(X_i) \quad \text{and} \quad W_n = U_n - n^{-1}Z_n, \tag{4.35}$$

and let $N_i(\Delta)$, $i=1,2$, be two stopping variables, defined to be the smallest positive integer $n(\geq n_0(\Delta))$ for which the following

$$m^2(1+(-1)^i\eta)\zeta_1(F)(\log B)<\Delta[Z_n - \frac{mn}{2}(\theta_0+\theta_1)]<m^2(1+(-1)^i\eta)\zeta_1(F)(\log A) \tag{4.36}$$

is vitiated, where $\eta(>0)$ is arbitrarily small, and for $N_i(\Delta)=n$, if $\Delta[Z_n - \frac{mn}{2}(\theta_0+\theta_1)] \leq m^2(1+(-1)^i\eta)\zeta_1(F) \log B$, we accept H_0: $\theta(F)=\theta_0$, while if $\Delta[Z_n - \frac{mn}{2}(\theta_0+\theta_1)] \geq m^2(1+(-U^i\eta)\zeta_1(F) \log A$, we accept H_1: $\theta(F)=\theta_1$; $i=1,2$. Now, the Z_n involve summations of i.i.d.r.v., and for small Δ, the excess over the boundaries can be neglected. Further, proceeding precisely on the same line as in the proof of Theorem 4.1, it follows that the OC function for each $N_i(\Delta)$ satisfies (4.6) and (4.7), $i=1,2$. Hence, by the fundamental identity of Wald [16], we have for $\phi>\frac{1}{2}(\varepsilon I)$,

$$\lim_{\Delta\to 0}[\Delta^2 E_\phi\{N_i(\Delta)\}] = (1+(-1)^i\eta)\psi(\phi), \; i=1,2. \tag{4.37}$$

Also, by arguments similar to in (4.25) through (4.33),

$$\lim_{\Delta\to 0}\{\Delta^2[E_\phi N_i(\Delta)-\sum_{n_1(\Delta)+1}^{n_2(\Delta)} P_\phi\{N_i(\Delta)>n\}]\} = 0, \; i=1,2, \tag{4.38}$$

for all $\phi > \frac{1}{2}(\varepsilon I)$. Let us now consider the following.

Lemma 4.3. For every $\eta_1(>0)$, under (2.5),

$$P\{\max_{n\leq n_2(\Delta)} |\Delta n W_n| > \eta_1\} \leq \nu_1(\Delta,\eta_1); \lim_{\Delta\to 0} \nu_1(\Delta,\eta_1) = 0. \tag{4.39}$$

Proof. It follows from Berk [4] that $\{U_n, n\geq m\}$ forms a reverse martingale sequence, and the same proof applies to $\{n^{-1}Z_n, n\geq 1\}$; both these are defined with respect to a common sequence of σ-fields. Hence, by (4.35), $\{W_n, n\geq m\}$ is also a reverse martingale sequence. Now, for $\{W_n, n_0(\Delta)<n\leq n_2(\Delta)\}$, we reverse the ordering of the index set $\{i\}$ to $\{n_2(\Delta)-i+1,\ 1\leq i\leq n_2(\Delta)-n_0(\Delta)\}$, and thereby, convert the sequence into a (forward) martingale sequence. Then, we use the Chow extension of the Hájek-Rènyi inequality (cf. [11, (3.6)]), and obtain that

$$\begin{aligned} P\{\max_{n\leq n_2(\Delta)} |\Delta n W_n| > \eta_1\} &\leq \Delta^2\eta_1^{-2}\{n_2^2(\Delta) E W^2_{n_2(\Delta)} + \\ &\qquad \sum_{j=m}^{n_2(\Delta)-1} (n_2(\Delta)-j)^2 [E W^2_{n_2(\Delta)-j} - E W^2_{n_2(\Delta)-j+1}]\} \\ &\leq \Delta^2\eta_1^{-2} C_1 \zeta_m(F)\{1 + \sum_m^{n_2(\Delta)-1} \{1-[(n_2(\Delta)-j)/(n_2(\Delta)-j+1)]^2\} \\ &\leq \Delta^2\eta_1^{-2} C_1 \zeta_m(F)\{1 + 2\log[n_2(\Delta)/m]\} \\ &\leq \Delta^2\eta_1^{-2} C_1 \zeta_m(F)\{1 + 2\log C_2 + 2\log(\varepsilon^{-2}\Delta^{-2})\},\quad 1<C_2<\infty, \end{aligned}$$

as some routine computations yield that $E\{W_n^2\} \leq C_1\zeta_1(F)n^{-2}$ and $E\{W_n^2 - W_{n+1}^2\} \leq C_1\zeta_1(F)[n^{-2}-(n+1)^{-2}]$ (where $C_1<\infty$). Since the right hand side of (4.40) converges to 0 as $\Delta\to 0$ (for every fixed $\eta>0$), the lemma follows.

Also, by (4.32), for every $\eta_1>0$, there exists a positive $C(\eta_1)$ $(<\infty)$, such that

$$P\{\max_{n_0(\Delta)<n\leq n_2(\Delta)} s_n^2 > (1+\eta_1)\zeta_1(F)\} \leq C(\eta_1)[n_0(\Delta)]^{-\delta'}$$

$$= \nu_2(\Delta,\eta_1) \quad \text{where} \quad \lim_{\Delta\to 0} \nu_2(\Delta,\eta_1) = 0. \tag{4.41}$$

Therefore, by (4.35), (4.39), (4.41) and the definition of $N(\Delta)$, for every n $\varepsilon[n_1(\Delta),n_2(\Delta)]$, as $\Delta\to 0$,

$$P_\phi\{N(\Delta)>n\} \leq P_\phi\{m^2(\log B)s_r^2 < r\Delta[U_r-\tfrac{1}{2}(\theta_0+\theta_1)]$$

$$< m^2(\log A)s_r^2,\ n_0(\Delta)\leq r<n\}$$

$$\leq P_\phi\{m^2(\log B)(1+\eta)\zeta_1(F) < \Delta[Z_r-\frac{rm}{2}(\theta_0+\theta_1)]$$

$$< m^2(\log A)(1+\eta)\zeta_1(F),\ n_0(\Delta)\leq r<n\} + \nu(\Delta,\eta)$$

$$= P_\phi\{N_2(\Delta)>n\} + \nu(\Delta,\eta),\ \eta=2\eta_1, \tag{4.42}$$

where $\lim_{\Delta\to 0} \nu(\Delta,\eta)=0$. Essentially, by similar steps,

$$P_\phi\{N(\Delta)>n\} \geq P_\phi\{N_1(\Delta)>n\} - \nu(\Delta,\eta), \tag{4.43}$$

for all n $\varepsilon[n_1(\Delta),n_2(\Delta)]$ and $\Delta\to 0$. Consequently, by (4.34), (4.37) and (4.38), we get that

$$(1-\eta)\psi(\phi) - \lim_{\Delta\to 0}\{\Delta^2\nu(\Delta,\eta)[n_2(\Delta)-n_1(\Delta)]\} \leq$$

$$\lim_{\Delta\to 0} \Delta^2 \, E_\phi N(\Delta) \leq (1+\eta)\psi(\phi) + \lim_{\Delta\to 0}\{\Delta^2\nu(\Delta,\eta)[n_2(\Delta)-n_1(\Delta)] \ . \tag{4.44}$$

Since, by (4.25), (4.40) and (4.41), $\lim_{\Delta\to 0} \Delta^2\nu(\Delta,\eta)n_2(\Delta)\to 0$ and $\eta(>0)$ is arbitrary, the proof follows.

A similar proof follows for the alternative procedure based on $\{\theta(F_n)\}$ and $\{s_n^*\}$.

The above proof fails when $\phi=\frac{1}{2}$. However, if

$$(d/d\phi)P(\phi)\Big|_{\frac{1}{2}-o} = (d/d\phi)P(\phi)\Big|_{\frac{1}{2}+o} = P'(\phi) \text{ exists,} \tag{4.45}$$

by using (4.24) and the L'Hospital's rule, we obtain that

$$\lim_{\Delta\to 0}[\Delta^2 E_{\frac{1}{2}}\{N(\Delta)\}] = m^2\zeta_1(F)P'(\tfrac{1}{2})[\log\ (AB^{-1})]. \tag{4.46}$$

5. *An alternative sequential procedure*. Derived from the theory of bounded-length (sequential) confidence intervals of Chow and Robbins [5], Albert [1] has considered a second type of sequential tests in linear models, which may be extended as follows.

Had $\zeta_1(F)$ been known, $n^{\frac{1}{2}}[U_n-\theta(F)]/m\zeta_1^{\frac{1}{2}}(F)$ would have asymptotically a normal distribution with zero mean and unit variance. Hence, for a test for (2.6), based on U_n, if for small Δ, we want to have the strength (α,β), we could have used a fixed-sample size procedure with n specified by

$$n \geq n_\Delta(\alpha,\beta) = [m^2\zeta_1(F)\{\tau_\alpha+\tau_\beta\}^2/\Delta^2], \tag{5.1}$$

where τ_α is the upper 100α% point of the standard normal distribution. Since s_n^2, defined by (2.12), converges a.s. to $\zeta_1(F)$ as $n\to\infty$, we may consider the following sequential procedure:

define a stopping variable $N^*(\Delta)$ as the smallest positive integer n for which $s_n^2 m^2(\tau_\alpha+\tau_\beta)^2/\Delta \le n$; accept H_0: $\theta(F)=\theta_0$ if $n^{1/2}[U_n-\theta_0]/ms_n < \tau_\alpha$ and reject H_0, otherwise.

Along the same line as in Sproule [14], who extended the Chow-Robbins [5] theory of bounded-length confidence intervals to U-statistics, it follows that

$$\lim_{\Delta\to 0}\{E_\phi\{N^*(\Delta)\}/n_\Delta(\alpha,\beta)\} = 1,\ \forall\phi\varepsilon I, \tag{5.2}$$

(in fact, $E_\phi N^*(\Delta)$ does not depend on ϕ). Thus, on comparing (4.24), (4.46), (5.1) and (5.3), we obtain that the asymptotic relative efficiency (A.R.E.) of $N^*(\Delta)$ with respect to $N(\Delta)$ is

$$e(\phi) = \lim_{\Delta\to 0}[E_\phi N(\Delta)]/[E_\phi N^*(\Delta)]$$

$$= \begin{cases} \dfrac{P(\phi)\log B+[1-P(\phi)]\log A}{(\phi-\frac{1}{2})\{\tau_\alpha+\tau_\beta\}^2} & \text{if } \phi\neq\frac{1}{2} \\ P'(\frac{1}{2})(\log AB^{-1})/\{\tau_\alpha+\tau_\beta\}^2, & \text{if } \phi=\frac{1}{2}. \end{cases} \tag{5.3}$$

For $\phi=0$ or 1, $P(\phi)$ is known, and hence, $e(\phi)$ can be easily computed. For example, when $\alpha=\beta=.05$ and $\phi=0$(or 1), $e(\phi)$ equals to 0.48, which is considerably less than one. In fact, the table on page 57 in Wald [16] is quite appropriate here and reveals the supremacy of the earlier procedure over the one considered in this section. Because of the remark made before (4.45), the two procedures of section 2 are asymptotically equally efficient too.

6. A few applications.

(i) Sequential test for variance of a distribution. Let $\{X_1, X_2, \ldots\}$ be iidrv with a univariate d.f. $F(x)$, defined on the real line $(-\infty,\infty)$, and let

$$\mu(F) = \int_{-\infty}^{\infty} x dF(x),\ \sigma^2(F) = \int_{-\infty}^{\infty} x^2 dF(x) - \mu^2(F), \tag{6.1}$$

where we assume that $0<\sigma(F)<\infty$. We want to test

$$H_0:\ \sigma^2(F) = \sigma_0^2 \text{ (known) vs. } H_1:\ \sigma^2(F) = \sigma_0^2+\Delta,\ \Delta \text{ known.} \tag{6.2}$$

Since for the kernel $\phi(X,Y) = \frac{1}{2}(X-Y)^2$, $E\phi = \sigma^2(F)$, by (2.8) and (2.9), we have

$$\sigma^2(F_n) = \sum_{i=1}^n (X_i-\bar{X}_n)^2/n \text{ and } U_n = [n/(n-1)]\sigma^2(F_n), \text{ where } X_n = \frac{1}{n}\sum_{i=1}^n X_i. \tag{6.3}$$

Also, by (2.11), (2.12) and some routine steps, we get that

$$s_n^2 = [n/(n-1)]^3\ \tfrac{1}{4}\{\frac{1}{n}\sum_{i=1}^n (X_i-\bar{X}_n)^4 - [\sigma^2(F_n)]^2\}, \tag{6.4}$$

$$(s_n^*)^2 = \frac{1}{4n}\sum_{i=1}^n (X_i-\bar{X}_n)^4 - \tfrac{1}{4}[\sigma^2(F_n)]^2. \tag{6.5}$$

Hence, we have no problem in following the sequential procedures in (2.13) or the alternative one after (2.14). Incidentally, here $\zeta_1(F) = \frac{1}{4}\{E(X-\mu)^4 - \sigma^4(F)\}$, and hence, the computations in (4.24) or (4.46) pose no problem.

(ii) Sequential tests for independence. Let $X_i = (X_i^{(1)}, X_i^{(2)})$, $i=1,2,\ldots$, be iidrv with a bivariate d.f. $F(x_1,x_2)$, $-\infty<x_1,x_2<\infty$. We want to test the hypothesis that $X^{(1)}$, $X^{(2)}$ are stochastically independent i.e.,

$$H_0: \quad F(x_1,x_2) = F(x_1,\infty)F(\infty,x_2) \;\forall (x_1,x_2)\varepsilon R^2. \tag{6.6}$$

Various alternative hypotheses may be framed to test for departure from independence. This may be the usual correlation between $X^{(1)}$, $X^{(2)}$, or the grade correlation {for definition etc., see [8]}, or some sort of association of the two variates. We consider here a simple case; the other cases follow on parallel lines.

Let us define the <u>probability of concordance</u> by

$$\Pi^{(C)}_{(F)} = P\{X_1^{(1)}<X_2^{(1)}, X_1^{(2)}<X_2^{(2)}\} + P\{X_1^{(1)}>X_2^{(1)}, X_1^{(2)}>X_2^{(2)}\}, \tag{6.7}$$

where we assume that F is continuous. Then, under H_0 in (6.6), $\Pi^{(C)}_{(F)}=\frac{1}{2}$, while $\Pi^{(C)}_{(F)}>(\text{or } <)\frac{1}{2}$ indicates a positive (or negative) association between the two variates. In fact, the tau correlation coefficient, proposed by Kendall and others, is given by $\tau(F) = 2\Pi^{(C)}(F) - 1$. Thus, we want to test for H_0 in (6.6) against an alternative that $\tau(F) = \Delta$, $\Delta>0$.

Let $c(u_1,u_2)$ be 1 if both u_1 and u_2 are non-negative, and let $c(u_1,u_2)=0$, otherwise. Then,

$$U_n = 2\binom{n}{2}^{-1} \sum_{1\leq i<j\leq n} c(X_i^{(1)}-X_j^{(1)}, X_i^{(2)}-X_j^{(2)})-1 \tag{6.8}$$

unbiasedly estimates $\tau(F)$. In a sample of n observations, let C_{ni} be the number of X_j, $1\leq j(\neq i)\leq n$, which are concordant with X_i, $1\leq i\leq n$, so that $U_n = \frac{4}{n(n-1)} \sum_{i=1}^{n} C_{ni}-1$. Then, by (2.11) and (2.12), we have here

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \{2(n-1)^{-1} C_{ni} - U_n\}^2. \tag{6.9}$$

Thus, we can again proceed as in section 2. Since under (6.7), $\zeta_1(F) = \frac{1}{9}$, for small Δ, we may replace s_n^2 by $\frac{1}{9}$ in the definition of $N(\Delta)$.

(iii) Sequential tests based on generalized U-statistics. Let $\{X_1, X_2, \ldots\}$ be iidrv with a d.f. $F(x)$, and let $\{Y_1, Y_2, \ldots\}$ be an independent sequence of iidrv with a d.f. $G(x)$. Consider a functional

$$\theta(F,G) = \int \cdots \int \phi(x_1, \ldots, x_{m_1}, y_1, \ldots, y_{m_2}) dF(x_1) \ldots dF(x_{m_1}) dG(y_1) \ldots dG(y_{m_2}) \tag{6.10}$$

of degree (m_1, m_2); $m_1 \geq 1$, $m_2 \geq 1$. For every $n \geq \max[m_1, m_2]$, let

$$U_n = \binom{n}{m_1}^{-1} \binom{n}{m_2}^{-1} \sum_{C_{n,m_1}} \sum_{C_{n,m_2}} \phi(X_{i_1}, \ldots, X_{i_{m_1}}, Y_{j_1}, \ldots, Y_{j_{m_2}}) \tag{6.11}$$

be the U-statistic corresponding to $\theta(F,G)$, while $\theta(F_n, G_n)$ is defined by replacing F and G in (6.10) by $F_n(x) = n^{-1} \sum_{i=1}^{n} c(x - X_i)$ and $G_n(y) = n^{-1} \sum_{j=1}^{n} c(y - Y_j)$, respectively. Also, let

$$V_{n,io} = \binom{n-1}{m_1 - 1}^{-1} \binom{n}{m_2}^{-1} \sum_i^* \sum_{C_{n_1 m_2}} \phi(X_i, X_{i_2}, \ldots, X_{i_{m_1}}, Y_{j_1}, \ldots, Y_{j_{m_2}}), \tag{6.12}$$

$$V_{n,oj} = \binom{n}{m_1}^{-1} \binom{n-1}{m_2 - 1}^{-1} \sum_{C_{n,m_1}} \sum_j^* \phi(X_{i_1}, \ldots, X_{i_{m_1}}, Y_j, Y_{j_2}, \ldots, Y_{i_{m_2}}), \tag{6.13}$$

where the summation $\sum_i^*$ $(\sum_j^*)$ extends over all possible $1 \leq i_2 < \ldots < i_{m_1} \leq n$ with $i_j \neq i$, $j = 2, \ldots, m_1$ $(1 \leq i_2 < \ldots < i_{m_2} \leq n$ with $j_i \neq j$, $i = 2, \ldots, m_2)$, and finally set

$$v_n^2 = m_1^2 s_{n,1}^2 + m_2^2 s_{n,2}^2; \tag{6.14}$$

$$s_{n,1}^2 = \frac{1}{n-1} \sum_{i=1}^n [V_{n,io} - U_n]^2, \quad s_{n,2}^2 = \frac{1}{n-1} \sum_{j=1}^n [V_{n,oj} - U_n]^2. \tag{6.15}$$

Then, we may consider a sequential procedure for testing

$$H_0: \ \theta(F,G)=\theta_0 \quad \text{vs.} \quad H_1: \ \theta(F,G)=\theta_1=\theta_0+\Delta, \ \Delta>0, \tag{6.16}$$

based on the same stopping variable as in (2.13), wherein we replace $m^2 s_n^2$ by v_n^2. A similar procedure follows for $\theta(F_n, G_n)$.

Let then $\phi_{10}(x_i) = E\{\phi(x_i, X_{i_2}, \ldots, X_{i_{m_1}}, Y_{j_1}, \ldots, Y_{j_{m_2}})\}$, $\phi_{01}(y_i) = E\{\phi(X_{i_1}, \ldots, X_{i_{m_1}}, y_j, Y_{j_2}, \ldots, Y_{i_{m_2}})\}$, and let

$$\nu^2(F,G) = m_1^2 \ E\{\phi_{10}(X_i) - \theta(F,G)\}^2 + m_2^2 \ E\{\phi_{01}(Y_j) - \phi(F,G)\}^2. \tag{6.17}$$

Then, by direct generalization of Theorems 4.1 and 4.2, it can be shown that (4.6), (4.7), (4.24) and (4.46) hold; the only change is to replace $m^2 \zeta_1(F)$ by $\nu^2(F,G)$. The above results also can be easily extended to functionals of $c(\geq 2)$ independent distributions.

We conclude this section with an example of $\theta(F,G)$ having a lot of practical importance. Consider the functional

$$\theta(F,G) = P\{X<Y\} = \int F(x)dG(x), \tag{6.18}$$

and suppose we want to test the hypotheses in (6.16). There are two possible ways of constructing sequential tests for the same.

(a) The Wald SPRT. Define r_i as 1 or 0 according as $X_j<Y_i$ or $X_i \geq Y_i$, $i \geq 1$. Thus, $P\{r_i=1\} = E\{r_i\} = \theta(F,G)$, and as the r_i are iidrv with a binomial distribution, we may directly use the Wald [16] SPRT. Thus, theorem 4.1 holds here, while the ASN, as given in Wald [16, p.57] reduces (as $\Delta \to 0$) for $\theta(F,G) = \theta_0+\phi\Delta$, $\phi \varepsilon I$, to

$$\frac{P(\phi)\log B + \{1-P(\phi)\}\log A}{\{\theta_0+\phi\Delta\}\log\{1+\theta_0^{-1}\Delta\} + \{1-\theta_0-\varepsilon\Delta\}\log\{1-(1-\theta_0)^{-1}\Delta\}}$$

$$= \frac{\theta_0(1-\theta_0)}{(\phi-\frac{1}{2})\Delta^2} \{P(\phi)\log B + \{1-P(\phi)\}\log A\}\{1+O(\Delta)\}. \tag{6.19}$$

(b) Asymptotic sequential test based on the Wilcoxon statistic. The generalized U-statistic corresponding to $\theta(F,G)$ is

$$\begin{aligned} U_n &= n^{-2} \sum_{i=1}^n \sum_{j=1}^n \phi(X_i,Y_j) \\ &= n^{-2} \sum_{j=1}^n \{\text{No of } X_i<Y_j,\ 1\leq i\leq n\}. \\ &= n^{-1} \sum_{j=1}^n F_n(Y_j), \end{aligned} \tag{6.20}$$

where $F_n(x)$ is the empirical d.f. for $X_1,\ldots,X_n$. By definition, in (6.12)-(6.15), we have

$$\nu_n^2 = \frac{1}{n-1} \{\sum_{i=1}^n [1-G_n(X_i)-U_n]^2 + \sum_{j=1}^n [F_n(Y_j)-U_n]^2\}, \tag{6.21}$$

where G_n is the empirical d.f. for $Y_1,\ldots,Y_n$. Thus, we may proceed as in (2.13) with the change suggested after (6.17).

Note that by (6.17)

$$\nu^2(F,G) = E[1-G(X_1)]^2 + E[F(Y_1)]^2 - 2\theta^2(F,G), \tag{6.22}$$

so that the ASN as $\Delta \to 0$ is given by

$$\frac{\nu^2(F,G)}{(\phi-\frac{1}{2})\Delta^2}\{P(\phi)\log B + [1-P(\phi)]\log A\}\{1+o(1)\}, \tag{6.23}$$

where $\theta(F,G) = \theta_0 + \phi\Delta$, $\phi \varepsilon I$. From (6.19) and (6.23), we obtain the A.R.E. of the Wald SPRT with respect to the proposed test equal to

$$\nu^2(F,G)/\theta_0(1-\theta_0). \tag{6.24}$$

In particular, when we want to test $F \equiv G$ against shift alternative, $\theta_0 = \frac{1}{2}$, $\nu^2(F,G) = \frac{1}{6}$, so that (6.24) equals to $\frac{2}{3}$, which clearly indicates the supremacy of the second procedure. In general, (6.24) is less than unity; this is because of the fact that whereas U_n compares every X_i with every Y_j, the Wald SPRT does not do so, and hence, looses some information.

REFERENCES

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

[2] F. J. Anscombe, Large sample theory of sequential estimation. Proc. Cambridge Phil. Soc. 48 (1952), 600-607.

[3] M. S. Bartlett, The large sample theory of sequential tests. Proc. Cambridge Phil. Soc. 42 (1946), 239-244.

[4] R. H. Berk, Limiting behavior of posterior distributions when the model is incorrect. Ann. Math. Statist. 37 (1966), 51-58.

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

[6] D. R. Cox, Large sample sequential tests of composite hypotheses. Sankhya, Ser. A 25 (1963), 5-12.

[7] P. R. Halmos, The theory of unbiased estimation. Ann. Math. Statist. 17 (1946), 34-43.

[8] W. Hoeffding, A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19, (1948), 293-325.

[9] R. V. Mises, On the asymptotic distribution of differentiable statistical functions. Ann. Math. Statist. 18 (1947), 309-348.

[10] P. K. Sen, On some convergence properties of U-statistics. Calcutta Statist. Assoc. Bull. 10 (1960), 1-18.

[11] P. K. Sen, The Hájek-Rènyi inequality for sampling from a finite population. Sankhya, Ser. A 32 (1970), 181-188.

[12] P. K. Sen, Convergence of sequences of regular functionals of empirical distributions to processes of Brownian motion. Inst. Statist., Univ. N. Carolina Mimeo Report No. 726, January 1971.

[13] P. K. Sen, On L^p-convergence of U-statistics. Ann. Math. Statist. (submitted, 1971).

[14] R. N. Sproule, A sequential fixed-width confidence interval for the mean of a U-statistic. Ph.D. Thesis, Univ. N. Carolina, Chapel Hill, 1969.

[15] V. Strassen, Almost sure behavior of sums of independent random variables and martingales. Proc. 5th Berkeley Symp. Math. Statist. Prob. 2 (1967), 315-343.

[16] A. Wald, Sequential Analysis. John Wiley, New York, 1947.