

Research supported by National Institutes of Health, Grant No. GM 12868

ON BOUNDED LENGTH SEQUENTIAL CONFIDENCE INTERVALS
BASED ON ONE-SAMPLE RANK-ORDER STATISTICS

by

Pranab Kumar Sen and Malay Ghosh

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 648

November 1969

ON BOUNDED LENGTH SEQUENTIAL CONFIDENCE INTERVALS BASED
ON ONE-SAMPLE RANK-ORDER STATISTICS*

by

PRANAB KUMAR SEN and MALAY GHOSH

Department of Biostatistics
University of North Carolina at Chapel Hill

1. Introduction. The problem of finding a bounded length confidence band for the mean of an unknown distribution is studied by Anscombe (1952) and Chow and Robbins (1965). Farrell (1966) considers the problem for the p -quantile of a distribution. Sproule (1969) has extended the results of Chow and Robbins to the class of Hoeffding's (1948) U -statistics, and in the particular cases of the signed-rank and sign statistics (which are both U -statistics), Geertsema (1968) considers the problem based on rank estimates of the median.

In the present paper, we consider the problem of providing a (sequential) confidence interval for the median of a symmetric (but otherwise unknown) distribution based on a general class of one-sample rank-order statistics. Of particular interest is the procedure based on the so called one-sample normal scores statistics. This procedure is shown to be asymptotically (i.e., as the prescribed bound on the width of the confidence interval is made to converge to zero) at least as efficient as the Chow-Robbins procedure for a broad class of parent distributions.

In course of this study, several asymptotic results, having importance of their own, are derived. First, the elegant result of Bahadur (1966) on the behaviour of the empirical distribution in the neighbourhood of a quantile is extended to the entire real line (see Theorem 4.1). Second, the weak convergence results of Sen (1966, Theorem 1) and Jurečková (1969) are replaced by almost

*Research supported by National Institutes of Health, Grant No. GM 12868

sure (a.s.) convergence results, under slightly more restrictive conditions on the scores and the underlying distribution; see Theorem 4.3. It is also shown (see Theorem 4.6) that the usual one-sample rank order statistic possesses the martingale or the sub-martingale property according as the parent distribution is symmetric about the origin or not.

Section 2 of the paper deals with the preliminary notions and basic assumptions. The next section describes the proposed procedure and states the main theorem of the paper. Section 4 is concerned with the results stated in the preceding paragraph. The proof of the main theorem is supplied in Section 5, and the asymptotic efficiency results are studied in Section 6.

2. Preliminary Notions. Let $\{X_1, X_2, \dots \text{ ad inf}\}$ be a sequence of independent and identically distributed random variables (iidrv) having an absolutely continuous distribution function (d.f.) $F_\theta(x)$ with location parameter θ (unknown). It is desired to determine (sequentially) a confidence interval $I_N = \{\theta: \hat{\theta}_{L,N} < \theta < \hat{\theta}_{U,N}\}$, where for each positive integer n , $\hat{\theta}_{L,n}$ and $\hat{\theta}_{U,n}$ are two statistics (not depending on d) based on the first n observations, such that $\hat{\theta}_{U,n} - \hat{\theta}_{L,n} > 0$ a.s. and $\lim_{n \rightarrow \infty} P\{\theta \in I_n\} = 1 - \alpha$ (the desired confidence coefficient), while N is the stopping variable defined to be the first integer $n \geq n_0$ (some positive integer) such that $\hat{\theta}_{U,n} - \hat{\theta}_{L,n} \leq 2d$.

Our procedure for determining N and $(\hat{\theta}_{L,N}, \hat{\theta}_{U,N})$ rests on the following class of one-sample rank order statistics. Let $c(u) = 0$ or 1 according as $u < 0$ or not. Let then

$$(2.1) \quad R_{n\alpha} = \sum_{\beta=1}^n c(|X_\alpha| - |X_\beta|), \quad \alpha=1, \dots, n; \quad X_n = (X_1, \dots, X_n).$$

Define

$$(2.2) \quad T_n = T_n(X_n) = n^{-1} \sum_{\alpha=1}^n c(X_\alpha) J_n((n+1)^{-1} R_{n\alpha}),$$

where $\{J_n(u): 0 < u < 1\}$ is generated by a score-function $\{J(u): 0 \leq u < 1\}$ in either of the following two ways:

$$(a) \quad J_n(u) = J\left(\frac{i}{n+1}\right), \quad (i-1)/n < u \leq i/n, \quad \text{for } i=1, \dots, n;$$

$$(b) \quad J_n(u) = EJ(U_{ni}), \quad (i-1)/n < u \leq i/n, \quad 1 \leq i \leq n, \quad \text{where } U_{n1} \leq \dots \leq U_{nn}$$

are the n ordered random variables from the rectangular $(0,1)$ d.f. Also, we assume that, $J(u) = \Psi^{-1}\left(\frac{1+u}{2}\right)$, $0 \leq u < 1$, where $\Psi(x)$ is a df defined on $(-\infty, \infty)$ satisfying the conditions

$$(2.3) \quad (a) \quad \Psi(-x) + \Psi(x) = 1 \quad \text{for all } 0 \leq x < \infty,$$

$$(2.4) \quad (b) \quad -\log[1-\Psi(x)] \text{ is convex for all } x \geq x_0,$$

where $x_0 (>0)$ is some real number.

Note that by definition, $J(0) = 0$ and $J(u)$ is \uparrow in $u: 0 \leq u < 1$. Also, (2.4) implies that there exists a $t_0 (>0)$, such that

$$(2.5) \quad M(t) = \int_0^1 \exp[tJ(u)] du < \infty \quad \text{for all } t \leq t_0;$$

$$(2.6) \quad J(u) \leq K[-\log(1-u)], \quad 0 \leq u < 1;$$

$$(2.7) \quad J'(u) \leq K/(1-u), \quad 0 \leq u < 1, \quad 0 < K < \infty.$$

Note that (2.4) [and hence, (2.6) and (2.7)] hold for the normal, the logistic, double exponential and many other df's. The statistic T_n when Ψ is the standard normal df is termed the normal scores statistic, and when Ψ is uniform over $(-1,1)$, it is termed the signed-rank statistic.

We denote by \mathcal{F}_0 the class of all absolutely continuous df's $\{F(x)\}$ symmetric about 0 for which both the density function $f(x)$ and its first derivative $f'(x)$ are bounded for almost all x (a.a.x). Also, let $\mathcal{F}_0(J)$ be the class of all $F \in \mathcal{F}_0$ for which

$$(2.8) \quad \lim_{x \rightarrow \infty} f(x)J'[F(x)-F(-x)] \text{ is bounded.}$$

That is

$$(2.9) \quad \mathcal{F}_0(J) = \{F: F \in \mathcal{F}_0, (2.8) \text{ holds}\} .$$

Throughout the paper it will be assumed that $F_\theta(x) = F(x-\theta)$, where $F \in \mathcal{F}_0(J)$.

[For better insight on (2.8), the reader is referred to Lemma 7.2 of Puri (1964)].

Introduce the following notations:

$$(2.10) \quad \bar{E}_n = n^{-1} \sum_{i=1}^n J_n(i/(n+1)), \quad A_n^2 = n^{-1} \sum_{i=1}^n J_n^2(i/(n+1));$$

$$(2.11) \quad \mu = \int_0^1 J(u)du \text{ and } A^2 = \int_0^1 J^2(u)du.$$

Note that if $\theta=0$, $T_n(X_n)$ has a distribution independent of F . Hence, there exists two (known) quantities $T_{n,\alpha}^{(1)}$ and $T_{n,\alpha}^{(2)}$ ($=\bar{E}_n - T_{n,\alpha}^{(1)}$), such that

$$(2.12) \quad P_{\theta=0} \{T_{n,\alpha}^{(2)} \leq T_n(X_n) \leq T_{n,\alpha}^{(1)}\} = 1-\alpha \rightarrow 1-\alpha \text{ as } n \rightarrow \infty.$$

For large n , it is known that

$$(2.13) \quad \lim_{n \rightarrow \infty} n^{1/2}(T_{n,\alpha}^{(2)} - \bar{E}_n) = -1/2 A \tau_{\alpha/2}, \quad \lim_{n \rightarrow \infty} n^{1/2}(T_{n,\alpha}^{(1)} - \bar{E}_n) = 1/2 A \tau_{\alpha/2},$$

where

$$(2.14) \quad \Phi(\tau_{\alpha/2}) = 1-\alpha/2; \quad \Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-1/2 t^2) dt.$$

It is also known (cf [3, Theorem 2]) that under the assumptions made earlier

$$(2.15) \quad n^{-1} \sum_{i=1}^n |J_n(R_{n\alpha}/(n+1)) - J(R_{n\alpha}/(n+1))| = o(n^{-1/2}),$$

which implies that the use of either of the scores will lead to the same asymptotic results.

3. ~~The procedure for obtaining I_n .~~ Since, $J(u) \uparrow u$, ($0 \leq u < 1$), $T_n(X_n - a1_n)$, (where $1_n = (1, \dots, 1)$) is \downarrow in a : $-\infty < a < \infty$. Define as in [13],

$$(3.1) \quad \hat{\theta}_{L,n} = \sup\{a: T_n(X_n - a1_n) > T_{n,\alpha}^{(1)}\},$$

$$(3.2) \quad \hat{\theta}_{U,n} = \inf\{a: T_n(X_n - a1_n) < T_{n,\alpha}^{(2)}\}.$$

It follows immediately that $\hat{\theta}_{U,n} - \hat{\theta}_{L,n} > 0$ a.s. and $P_\theta\{\hat{\theta}_{L,n} \leq \theta \leq \hat{\theta}_{U,n}\} = 1 - \alpha_n$.

The main theorem of the paper is as follows:

THEOREM 3.1. Under the assumptions of section 2,

$$(3.3) \quad N(=N(d)) \text{ is a non-increasing function of } d, \text{ } N(d) \text{ is finite a.s. and } EN(d) < \infty \text{ for all } d > 0; \lim_{d \rightarrow 0} N(d) = \infty \text{ a.s. and } \lim_{d \rightarrow 0} EN(d) = \infty;$$

$$(3.4) \quad \lim_{d \rightarrow 0} N(d)/v(d) = 1 \text{ a.s.};$$

$$(3.5) \quad \lim_{d \rightarrow 0} P\{\theta \in I_{N(d)}\} = 1 - \alpha;$$

$$(3.6) \quad \lim_{d \rightarrow 0} E[N(d)]/v(d) = 1,$$

where $v(d) = A^2 \tau_\alpha^2 / (4d^2 B^2(F))$ and

$$(3.7) \quad B(F) = \int_0^\infty \frac{d}{dx} J[2F(x) - 1] dF(x).$$

The proof of the theorem is postponed to section 5. In proving this theorem, several results are needed which are proved in section 4.

4. ~~Some convergence results on the empirical process and on~~ $\{T_n(X_n - a_{1/n}) : -\infty < a < \infty\}$. Since $F(x)$ is absolutely continuous, the random variables $Y_i = F(X_i)$, $i=1,2,\dots$ are distributed uniformly over $(0,1)$. Define the empirical process $[V_n(t) : 0 \leq t \leq 1]$ by

$$(4.1) \quad V_n(t) = n^{1/2}[G_n(t) - t]; \quad G_n(t) = n^{-1} \sum_{i=1}^n c(t - Y_i), \quad 0 \leq t \leq 1.$$

Note that $V_n(t) = 0$ for $t \notin (0,1)$. Let $g_k(n) = n^{-1/2}(\log n)^k$, $k \geq 1$. Define

$$(4.2) \quad K_n(t) = \sup\{|V_n(t) - V_n(t+a)| : |a| \leq g_k(n)\}, \quad 0 < t < 1,$$

$$(4.3) \quad K_n(*) = \sup_{0 < t < 1} (K_n(t)).$$

It is known that $K_n(*)$ weakly converges to zero in the Prohorov sense, viz. Hájek and Šidák (1967). However, for our results in section 5, we require the strong convergence of $K_n(*)$. This is accomplished here by extending a result of Bahadur (1966, p. 578) to the entire real line.

THEOREM 4.1. For every finite $s(>0)$, there exists a positive constant $c_s^{(1)}$ and a sample size n_s , such that for $n \geq n_s$,

$$(4.4) \quad P\{K_n(*) \geq c_s^{(1)} n^{-1/2} (\log n)^k\} \leq 4n^{-s}.$$

Hence, $K_n(*) = O(n^{-1/2} (\log n)^k)$ with probability 1 as $n \rightarrow \infty$.

Proof. Let $\xi_{j,n} = j[n^{1/2}]^{-1} (0 \leq j \leq [n^{1/2}]), [x]$ being the largest integer less than or equal to x . Since for the uniform df., the density function is equal to 1 ($0 \leq t \leq 1$), by the same technique as in Bahadur (1966; (13), p. 579), it follows that for n large,

$$(4.5) \quad P\{K_n(\xi_{j,n}) > c_1^* n^{-\frac{1}{4}} (\log n)^k\} \leq 4 \exp. (-v_n), \quad j=1,2,\dots,[n^{\frac{1}{2}}],$$

where $c_1^* > 0$ and

$$(4.6) \quad v_n = -\frac{1}{4} \log n + [c_1^{*2} n^{\frac{1}{2}} (\log n)^{2k}] / [2\{n^{\frac{1}{2}} (\log n)^k + c_1^* n^{\frac{1}{4}} (\log n)^k\}] \\ = -\frac{1}{4} \log n + c_1^{*2} (\log n)^k (1 + c_1^* n^{-\frac{1}{4}})^{-1}.$$

Since $k \geq 1$, by proper choice of $c_1^* = \frac{1}{3} c_s^{(1)}$, say, v_n can be made greater than $(s + \frac{1}{2}) \log n$ for any given $s (> 0)$. Thus, using the Bonferroni inequality, we obtain that for sufficiently large n ,

$$(4.7) \quad P\{K_n(\xi_{j,n}) > \frac{1}{3} c_s^{(1)} n^{-\frac{1}{4}} (\log n)^k \text{ for at least one } j=1,2,\dots,[n^{\frac{1}{2}}]\} \\ \leq 4[n^{\frac{1}{2}}] \exp(-v_n) \leq 4n^{-s}.$$

(4.4) will follow from (4.7) and (4.8) which we prove below.

$$(4.8) \quad K_n(*) \leq 3 \left[\max_{1 \leq j \leq [n^{\frac{1}{2}}]} K_n(\xi_{j,n}) \right].$$

Note that if t and $t+a$ both belong to the same interval $(\xi_{j,n}, \xi_{j+1,n})$, as $\xi_{j+1,n} - \xi_{j,n} = n^{-\frac{1}{2}} < g_k(n)$, $|V_n(t+a) - V_n(t)| \leq 2K_n(\xi_{j,n})$. So, we need consider only the case where t and $t+a$ belong to two different intervals. Suppose now that $t \in [\xi_{j,n}, \xi_{j+1,n}]$ and $(t+a) \in [\xi_{r,n}, \xi_{r+1,n}]$. Then, $|a| \leq g_k(n) \Rightarrow \xi_{r,n} - \xi_{j+1,n} \leq g_k(n)$, whenever $r \geq j+1$ (otherwise, interchange j and r). Hence,

$$(4.9) \quad |V_n(t+a) - V_n(t)| \leq |V_n(t+a) - V_n(\xi_{r,n})| \\ + |V_n(\xi_{r,n}) - V_n(\xi_{j+1,n})| + |V_n(\xi_{j+1,n}) - V_n(t)| \\ \leq K_n(\xi_{r,n}) + 2K_n(\xi_{j+1,n}) \leq 3 \max_{1 \leq j \leq [n^{\frac{1}{2}}]} K_n(\xi_{j,n}),$$

for all $j, r=1, 2, \dots, [n^{\frac{1}{2}}]$. Hence (4.8) follows from (4.9). The second part of the theorem follows from (4.4) by letting $s > 1$.

Remark. If $F(x)$ is such that $\sup_x f(x) = f_0 < \infty$, we have $|F(x+a) - F(x)| \leq f_0 |a|$, $-\infty < x < \infty$, and hence from (4.4), we have for every $s > 0$, $n \geq n_s$,

$$(4.10) \quad P\left\{ \sup_x \sup_{|a| < g_k(n)} n^{\frac{1}{2}} |F_n(x+a) - F_n(x) - F(x+a) + F(x)| > c_s^{(1)} n^{-\frac{1}{4}} \log n \right\} \leq 4n^{-s},$$

where the empirical df $F_n(x)$ is defined by

$$(4.11) \quad F_n(x) = n^{-1} \sum_{i=1}^n c(x - X_i), \quad -\infty < x < \infty.$$

We also need a strong convergence result on $\sup_{0 < t < 1} |V_n(t)|$ which we state below:

LEMMA 4.2. $P\left\{ \sup_{0 < t < 1} |V_n(t)| = O((\log n)^{\frac{1}{2}}) \geq 1 - O(n^{-s}) \right\}$ for every $s > 0$.

The proof proceeds on the lines of the theorem 1 after using the same grid points as in the theorem, and hence is omitted.

In order to prove that $\sqrt{n} [\hat{\theta}_{U,n} - \hat{\theta}_{L,n}] \rightarrow A\tau_{\alpha/2}/B(F)$ a.s. as $n \rightarrow \infty$ (to be used in deriving (3.3)), we also require the theorem 4.3. The corresponding weak convergence result was proved by Sen (1966). Assume, without any loss of generality that $\theta=0$; the translation invariance of the estimates in (3.1) and (3.2) permits it. Then, define the process

$$(4.12) \quad W_n(a) = n^{\frac{1}{2}} [T_n(X_{[n]}) - T_n(X_{[n]-a1_n}) - aB(F)], \quad -\infty < a < \infty.$$

THEOREM 4.3. Under the assumptions of section 2, for every $s(>0)$, there exists positive constants $(k_s^{(1)}, k_s^{(2)})$, and a sample size n_s such that for all $n \geq n_s$, $k \geq 1$,

$$(4.13) \quad P\{\text{Sup}(|W_n(a)| : |a| < n^{-\frac{1}{2}}(\log n)^k) > k_s^{(1)} n^{-\frac{1}{4}}(\log n)^{k+1}\} < k_s^{(2)} n^{-s}$$

Hence, as $n \rightarrow \infty$, for all $|a| < n^{-\frac{1}{2}}(\log n)^k$, $W_n(a) \rightarrow 0$ a.s.

Proof. Define

$$(4.14) \quad H(x) = F(x) - F(-x), \quad H_a(x) = F(x+a) - F(-x+a) = P\{|X-a| \leq x\}$$

$$(4.15) \quad H_n(x) = F_n(x) - F_n(-x), \quad H_{n,a}(x) = F_n(x+a) - F_n(-x+a).$$

Note that as $F \in \mathfrak{F}_0$, for all $|a| < g_k(n)$, and n large,

$$(4.16) \quad H_a(x) - H(x) = O(n^{-1}(\log n)^{2k}), \text{ for all } -\infty < x < \infty.$$

Also, we may note that

$$(4.17) \quad \begin{aligned} & |H_{n,a}(x) - H_n(x) - H_a(x) + H(x)| \\ & \leq |F_{n,a}(x) - F_n(x) - F_a(x) + F(x)| \\ & \quad + |F_{n,a}(-x) - F_n(-x) - F_a(-x) + F(-x)|. \end{aligned}$$

We now prove the following lemma.

LEMMA 4.4. If $F \in \mathfrak{F}_0$, for all $|a| < n^{-\frac{1}{2}}(\log n)^k = g_k(n)$, and n large,

$$(4.18) \quad \sup_a \sup_x n^{\frac{1}{2}} |H_{n,a}(x) - H_n(x)| \leq 2 c_s^{(1)} (n^{-\frac{1}{4}}(\log n)^k) + O(n^{-\frac{1}{2}}(\log n)^{2k}),$$

with probability greater than $1 - 4n^{-s}$ for any $s(>0)$.

The proof of the lemma follows directly from theorem 4.1 and (4.14)-(4.17).

Returning now to the proof of the theorem we may write after using (2.15),

$$(4.19) \quad T_{n \sim n}(X_n) = \int_0^\infty J_n\left(\frac{n}{n+1}H_n(x)\right) dF_n(x) = \int_0^\infty J\left(\frac{n}{n+1}H_n(x)\right) dF_n(x) + o(n^{-\frac{1}{2}}),$$

and a similar expression for $T_n(X_n - a1_n)$. Hence,

$$\begin{aligned}
 (4.20) \quad & n^{\frac{1}{2}} [T_n(X_n) - T_n(X_n - a1_n)] \\
 &= n^{\frac{1}{2}} \left[\int_0^{\infty} \left\{ J\left(\frac{n}{n+1} H_n(x)\right) dF_n(x) - J\left(\frac{n}{n+1} H_{n,a}(x)\right) dF_n(x+a) \right\} \right] + o(1) \\
 &= I_{n1}(a) + I_{n2}(a) + o(1),
 \end{aligned}$$

where

$$(4.21) \quad I_{n1}(a) = n^{\frac{1}{2}} \int_0^{\infty} J\left(\frac{n}{n+1} H_n(x)\right) d[F_n(x) - F_n(x+a)],$$

$$(4.22) \quad I_{n2}(a) = n^{\frac{1}{2}} \int_0^{\infty} \left\{ J\left(\frac{n}{n+1} H_n(x)\right) - J\left(\frac{n}{n+1} H_{n,a}(x)\right) \right\} dF_n(x+a).$$

We shall only consider the case $0 < a < g_k(n)$ as the case of negative a follows on the same line. Then, we can write $I_{n1}(a)$ as

$$(4.23) \quad I_{n1}(a) = \begin{cases} 0 & \text{if } X_{(n)} \leq 0, \\ I_{n1}^{(1)}(a) + I_{n1}^{(2)}(a) & \text{if } X_{(n)} > 0, \quad (X_{(n)} = \max_{i < i < n} X_i) \end{cases}$$

where
$$I_{n1}^{(1)}(a) = n^{\frac{1}{2}} \int_0^{X_{(n)}} J(H(x)) d[F_n(x) - F_n(x+a)],$$

$$I_{n1}^{(2)}(a) = n^{\frac{1}{2}} \int_0^{X_{(n)}} \left\{ J\left(\frac{n}{n+1} H_n(x)\right) - J(H(x)) \right\} d[F_n(x) - F_n(x+a)].$$

Since $P\{X_{(n)} \leq 0\} = 2^{-n}$, we have on integration by parts and on using theorem 4.1,

$$\begin{aligned}
 (4.24) \quad I_{n1}^{(1)}(a) &= \int_0^{X_{(n)}} n^{\frac{1}{2}} [F_n(x+a) - F_n(x)] J'(H(x)) dH(x) \\
 &= \int_0^{X_{(n)}} n^{\frac{1}{2}} [F(x+a) - F(x)] J'(H(x)) dH(x) \\
 &\quad + O(n^{-\frac{1}{2}} (\log n)^k) \int_0^{X_{(n)}} J'[H(x)] dH(x),
 \end{aligned}$$

with probability $\geq 1-4n^{-s}-2^{-n}$. Now, using the fact that $H(x) = F(x)-F(-x) \leq F(x)$ and for $\delta > 0$, $P\{[1-F(X_{(n)})] < n^{-(2+\delta)}\} = 1-\{1-n^{-(2+\delta)}\}^n \leq n^{-(1+\delta)}$, we obtain that the second term on the right hand side of (4.24) is of $O(n^{-\frac{1}{4}}(\log n)^{k+1})$, with probability $\geq 1-n^{-s}-2^{-n}$, by letting $1+\delta > s$. Hence, with probability $\geq 1-5n^{-s}-2^{-(n-1)}$,

$$(4.25) \quad \begin{aligned} & |I_{n1}^{(1)}(a) - n^{\frac{1}{2}}aB(F)| \\ & \leq \left| \int_0^{X_{(n)}} n^{\frac{1}{2}}\{F(x+a)-F(x)-af(x)\} J'(H(x))dH(x) \right| \\ & \quad + \left| \int_{X_{(n)}}^{\infty} n^{\frac{1}{2}}af(x)J'[H(x)]dH(x) \right| + O(n^{-\frac{1}{4}}(\log n)^{k+1}). \end{aligned}$$

Now, using the fact that $F \in \mathfrak{F}_0(J)$ and (2.7) holds, we obtain that the first term on the right hand side of (4.24) is $O(n^{-\frac{1}{2}}(\log n)^{2k+1})$, while the second term is bounded by $[\limsup |h(x)J'[H(x)]|](\log n)^k[1-F(X_{(n)})]$, which by the fact that (2.8) holds and $P\{1-F(X_{(n)}) \geq cn^{-\frac{1}{2}}\} = (1-cn^{-\frac{1}{2}})^n \leq O(n^{-s})$ for any $s > 0$, is also $O(n^{-\frac{1}{2}}(\log n)^k)$, with probability $\geq 1-O(n^{-s})$. Hence, we obtain that as $n \rightarrow \infty$

$$(4.26) \quad P\left\{ \sup_{0 < a < g_k(n)} |I_{n1}^{(1)}(a) - n^{\frac{1}{2}}aB(F)| = O(n^{-\frac{1}{4}}(\log n)^{k+1}) \right\} \geq 1-O(n^{-s}),$$

for any $s > 0$. We now rewrite $I_{n1}^{(2)}(a)$ as

$$(4.27) \quad n^{-\frac{1}{2}} \sum_{i=1}^n \left\{ J\left(\frac{n}{n+1} H_n(|X_i|)\right) - J(H(|X_i|)) \right\} \{c(X_i) - c(X_i+a)\}.$$

Note that $c(X_i) - c(X_i+a) = 0$ unless $-a < X_i \leq 0$. Hence, using theorem 4.1, (4.27), lemma 4.2, and after noting that $J'(u)$ is bounded in the neighbourhood of zero, we have,

$$(4.28) \quad \sup_{0 < a < g_k(n)} |I_{n1}^{(2)}(a)| \leq |n^{\frac{1}{2}}[F_n(0) - F_n(-g_k(n))]| \sup_{0 < a < g_k(n)} |J[\frac{n}{n+1} H_n(a)] - J[H(a)]|$$

$$= O((\log n)^k) \cdot O(n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}) = O(n^{-\frac{1}{2}}(\log n)^{k+\frac{1}{2}}),$$

with probability $\geq 1 - O(n^{-s})$, as $n \rightarrow \infty$.

Thus, it remains to prove that $\sup_{0 < a < g_k(n)} |I_{2n}(a)| = O(n^{-\frac{1}{4}}(\log n)^{k+1})$, with probability $\geq 1 - O(n^{-s})$. Define x_n^* by $1 - H(x_n^*) = 4c_s^{(1)} n^{-3/4} (\log n)^k$, where $c_s^{(1)}$ is defined by (4.4). Then

$$(4.29) \quad I_{n2}(a) = \int_0^{x_n^*} + \int_{x_n^*}^{\infty} \{J(\frac{n}{n+1} H_n(x)) - J(\frac{n}{n+1} H_{n,a}(x))\} dF_n(x+a).$$

$$= I_{n2}^{(1)}(a) + I_{n2}^{(2)}(a).$$

We first consider the following lemma the proof of which follows directly from the theorem 1 of Hoeffding (1963) after noting that $H_n(x)$ and $H_{n,a}(x)$ are bounded valued random variables and (4.16) holds.

LEMMA 4.5. For every $s(>0)$, there exist positive constants $c_s^{(1)}(>0)$, $c_s^{(2)}(>0)$ such that for sufficiently large n ,

$$(4.30) \quad P\{|H_n(x_n^*) - H(x_n^*)| > c_s^{(1)} n^{-\frac{7}{8}} (\log n)^k\} < c_s^{(2)} n^{-s},$$

$$(4.31) \quad P\{|H_{n,a}(x_n^*) - H(x_n^*)| > c_s^{(1)} n^{-\frac{7}{8}} (\log n)^k\} < c_s^{(2)} n^{-s}.$$

A direct use of lemma 4.4 and lemma 4.5 leads to the fact that for all $x \leq x_n^*$, with probability $\geq 1 - O(n^{-s})$,

$$(4.32) \quad 1 - H_n(x) \geq \frac{1}{3} [1 - H_{n,a}(x)],$$

and hence for any θ : $0 < \theta < 1$ and $0 \leq x \leq x_n^*$,

$$(4.33) \quad 1-H_{n,a,\theta}(x) = 1-\{\theta H_n(x)+(1-\theta)H_{n,a}(x)\}$$

$$= \theta[1-H_n(x)] + (1-\theta)[1-H_{n,a}(x)] \geq \frac{1}{3}[1-H_{n,a}(x)],$$

with probability $\geq 1-O(n^{-S})$. Hence, by using Lemmas 4.4, 4.5, (2.7), (4.32) and (4.33), we have with probability $\geq 1-O(n^{-S})$

$$(4.34) \quad \sup_{0 < a < g_k(n)} |I_{n2}^{(1)}(a)| = \sup_{0 < a < g_k(n)} n^{\frac{1}{2}} \int_0^{x_n^*} \frac{n}{n+1} [H_n(x) - H_{n,a}(x)]$$

$$J' \left[\frac{n}{n+1} H_{n,a,\theta_n}(x) \right] dF_{n,a}(x), \text{ (for some } \theta_n: 0 < \theta_n < 1)$$

$$\leq O(n^{-\frac{1}{4}}(\log n)^k) \int_0^{x_n^*} J' \left(\frac{n}{n+1} H_{n,a,\theta_n}(x) \right) dF_{n,a}(x)$$

$$\leq O(n^{-\frac{1}{4}}(\log n)^k) 3K \int_0^{x_n^*} \left[1 - \frac{n}{n+1} H_{n,a}(x) \right]^{-1} dH_{n,a}(x)$$

$$\leq O(n^{-\frac{1}{4}}(\log n)^k) 3K \int_0^{\infty} \left[1 - \frac{n}{n+1} H_{n,a}(x) \right]^{-1} dH_{n,a}(x)$$

$$= O(n^{-\frac{1}{4}} \log n)^k 3Kn^{-1} \sum_{i=1}^n \left(1 - \frac{i}{n+1} \right)^{-1}$$

$$\leq O(n^{-\frac{1}{4}} \log n)^k 6K \sum_{i=1}^n i^{-1}$$

$$\leq O(n^{-\frac{1}{4}}(\log n)^k) (1 + \log n) = O(n^{-\frac{1}{4}}(\log n)^{k+1}).$$

Finally, we consider $I_{n2}^{(2)}(a)$, for which we have

$$(4.35) \quad \sup_{0 < a < g_k(n)} |I_{n2}^{(2)}(a)| \leq \sup_{0 < a < g_k(n)} \left[n^{\frac{1}{2}} \int_{x_n^*}^{\infty} J \left(\frac{n}{n+1} H_n(x) \right) dF_n(x+a) \right]$$

$$+ \sup_{0 < a < g_k(n)} \left[n^{\frac{1}{2}} \int_{x_n^*}^{\infty} J \left(\frac{n}{n+1} H_{n,a}(x) \right) dF_n(x+a) \right].$$

Now, for every $a \geq 0$,

$$(4.36) \quad n^{\frac{1}{2}} \int_{x_n^*}^{\infty} J\left(\frac{n}{n+1} H_n(x)\right) dF_n(x+a) \leq n^{\frac{1}{2}} \int_{x_n^*+a}^{\infty} J\left(\frac{n}{n+1} H_n(x)\right) dF_n(x) \\ \leq n^{\frac{1}{2}} \int_{x_n^*}^{\infty} J\left(\frac{n}{n+1} H_n(x)\right) dH_n(x),$$

as $dF_n(x) \leq dH_n(x)$ and $J(u) \uparrow$ in $u: 0 \leq u < 1$. By (2.6), the right hand side of (4.36) is bounded above by

$$(4.37) \quad n^{\frac{1}{2}} \int_{H_n(x_n^*)}^1 -K \log(1-u) du \leq n^{\frac{1}{2}} H_n(x_n^*) \{-K \log(1-H_n(x_n^*))\} \\ = O(n^{-\frac{1}{4}} (\log n)^{k+1}), \text{ with probability } \geq 1 - O(n^{-s}),$$

by lemma 4.5 and definition of $H(x_n^*)$. Similarly, the second term on the right hand side of (4.32) is also $O(n^{-\frac{1}{4}} (\log n)^{k+1})$ with probability $\geq 1 - O(n^{-s})$. The proof is now complete.

We may observe that the weak convergence of the process $\{W_n(a)\}$ can be established along the lines of Jurečková [10] under less stringent conditions. However, we need here a stronger result as given in Theorem 4.3.

We shall now prove that when $F(x)$ is symmetric about 0 and $J_n(u)$ is specified by (b) of section 2, then, $T_n^* = n(T_n - \frac{1}{2}\mu)$, $n=1,2,\dots$ forms a martingale sequence with respect to a non-decreasing sequence of σ -fields \mathfrak{T}_n defined as follows.

Let $\underline{c}_n = (c(X_1), \dots, c(X_n))$ and $\underline{R}_n = (R_{n1}, \dots, R_{nn})$, where the $R_{n\alpha}$ are defined in (2.1). \mathfrak{T}_n is the σ -field generated by $(\underline{c}_n, \underline{R}_n)$. Then, obviously $\mathfrak{T}_n \uparrow$ in n and we prove the following theorem to be used in proving the 'uniform continuity' (to be explained in section 5) of the estimates $\hat{\theta}_{L,n}$ and $\hat{\theta}_{U,n}$.

THEOREM 4.6. If $J_n(u) = EJ(U_{ni})$, $(i-1)/n < u < i/n$, and if $F(x) + F(-x) = 1$ for all real x , $\{T_n^*, T_n\}$ forms a martingale sequence.

Proof. By definition in (2.2),

$$\begin{aligned}
 (4.38) \quad E(T_{n+1} | T_n) &= (n+1)^{-1} \sum_{i=1}^{n+1} E\{c(X_i) J_{n+1}(\frac{R_{n+1, \alpha}}{n+2}) | T_n\} \\
 &= (n+1)^{-1} \left[\sum_{i=1}^n c(X_i) E\{J_{n+1}(\frac{R_{n+1, \alpha}}{n+2}) | T_n\} \right. \\
 &\quad \left. + E\{c(X_{n+1}) J_{n+1}(\frac{R_{n+1, n+1}}{n+2}) | T_n\} \right]
 \end{aligned}$$

Under the hypothesis of the theorem, $c(X_{n+1})$ is independent of $R_{n+1, n+1}$ (see e.g. [7], p. 40), and hence,

$$(4.39) \quad E\{c(X_{n+1}) J_{n+1}(\frac{R_{n+1, n+1}}{n+2}) | T_n\} = \frac{1}{2(n+1)} \sum_{j=1}^{n+1} J_{n+1}(\frac{j}{n+2}) = \frac{1}{2} \bar{E}_{n+1},$$

where \bar{E}_{n+1} is defined by (2.10). It is easy to see that

$$(4.40) \quad \bar{E}_n = n^{-1} \sum_{i=1}^n EJ(U_{ni}) = \int_0^1 J(u) du = \mu, \text{ for all } n.$$

Also, given R_n , $R_{n+1, \alpha}$ can either assume the value $R_{n\alpha}$ or $R_{n\alpha} + 1$, with respective conditional probability $\frac{n+1-R_{n\alpha}}{n+1}$ and $\frac{R_{n\alpha}}{n+1}$. Hence,

$$\begin{aligned}
 (4.41) \quad E\{J_{n+1}(\frac{R_{n+1, \alpha}}{n+2}) | T_n\} &= \frac{n+1-R_{n\alpha}}{n+1} J_{n+1}(\frac{R_{n\alpha}}{n+2}) + \frac{R_{n\alpha}}{n+1} J_{n+1}(\frac{R_{n\alpha}+1}{n+2}) \\
 &= J_n(R_{n\alpha}/(n+1)), \quad 1 \leq \alpha \leq n,
 \end{aligned}$$

after using the fact that by definition of $J_{n+1}(\frac{i}{n+2})$, $1 \leq i \leq n+1$,

$$(4.42) \quad \frac{n+1-i}{n+1} J_{n+1}\left(\frac{i}{n+2}\right) + \frac{i}{n+1} J_{n+1}\left(\frac{i+1}{n+2}\right) = J_n\left(\frac{i}{n+1}\right).$$

Hence, from (4.39) through (4.41), we obtain that

$$(4.43) \quad \begin{aligned} E(T_{n+1} | \mathbf{T}_n) &= (n+1)^{-1} \left\{ \sum_{i=1}^n c(X_i) J_n\left(\frac{R_{n\alpha}}{n+1}\right) + \frac{1}{2}\mu \right\} \\ &= \frac{n}{n+1} T_n + \frac{1}{2(n+1)} \mu, \quad n=1,2,\dots \end{aligned}$$

This implies that for $T_n^* = n(T_n - \frac{1}{2}\mu)$,

$$(4.44) \quad E(T_{n+1}^* | \mathbf{T}_n) = T_n^*, \quad n \geq 1. \quad \text{q.e.d.}$$

Remark. The theorem may not hold when $J_n(u)$ is defined by (a) of section 2. However, if $J(u)$ is convex, then it can be shown by the same technique that $\{nT_n\}$ forms a sub-martingale sequence with respect to \mathbf{T}_n even when $J_n(u)$ is defined by (a) of section 2. Also, if $F(x)$ is not symmetric about 0, the martingale property does not hold. However, as $\bar{E}_n > 0$, it follows by the same technique that $\{nT_n\}$ forms a sub-martingale sequence with respect to \mathbf{T}_n when $J_n(u)$ is defined by (b) of section 2.

5. Proof of the theorem 3.1. We do it in several steps. First, let us prove the following lemmas.

LEMMA 5.1. For every $s(>0)$, there exists positive constants $c_s^{(1)}$ and $c_s^{(2)}$ and a sample size n_s , such that for $n \geq n_s$

$$(5.1) \quad P\{n^{\frac{1}{2}}(\hat{\theta}_{L,n} - \theta) + \tau_{\alpha/2} A/2B(F) < -c_s^{(1)} (\log n)^2\} \leq c_s^{(2)} n^{-s},$$

$$(5.2) \quad P\{n^{\frac{1}{2}}(\hat{\theta}_{U,n} - \theta) - \tau_{\alpha/2} A/2B(F) > c_s^{(1)} (\log n)^2\} \leq c_s^{(2)} n^{-s}.$$

Proof. We only prove (5.1), as (5.2) follows exactly on the same line. Now

$$\begin{aligned}
 (5.3) \quad & P\{n^{\frac{1}{2}}(\hat{\theta}_{L,n}^{-\theta}) + \tau_{\alpha/2} A/2B(F) < -c_s^{(1)}(\log n)^2\} \\
 & = P\{\hat{\theta}_{L,n} < \theta - n^{-\frac{1}{2}}\tau_{\alpha/2} A/2B(F) - n^{-\frac{1}{2}}c_s^{(1)}(\log n)^2\} \\
 & = P_{\theta=0}\{T_n(X_n + \{n^{-\frac{1}{2}}[\tau_{\alpha/2} A/2B(F) + c_s^{(1)}(\log n)^2]\}) \leq T_{n,\alpha}^{(1)}\},
 \end{aligned}$$

where $T_{n,\alpha}^{(1)} = \frac{1}{2}\mu + \frac{1}{2}An^{-\frac{1}{2}}\tau_{\alpha/2} + o(n^{-\frac{1}{2}})$. Let us define $\tilde{T}_n = T_n(X_n + n^{-\frac{1}{2}}(A\tau_{\alpha/2}/2B(F))) \big|_{\sim n}$ where $\theta=0$, without any loss of generality, and let $\hat{T}_n = T_n(X_n + \{n^{-\frac{1}{2}}(\tau_{\alpha/2} A/2B(F) + c_s^{(1)}(\log n)^2)\}) \big|_{\sim n}$, Then, by theorem 4.2, with probability $\geq 1 - k_s^{(2)}n^{-s}$, $n^{\frac{1}{2}}(\hat{T}_n - \tilde{T}_n) = B(F)c_s^{(1)}(\log n)^2 + O(n^{-\frac{1}{2}}(\log n)^3)$. Hence, it suffices to prove that with probability $\geq 1 - O(n^{-s})$, $\sqrt{n} |\tilde{T}_n - T_{n,\alpha}^{(1)}| \leq B(F)c_s^{*(1)}(\log n)^2$, where $c_s^{*(1)} < c_s^{(1)}$. We now write $a = n^{-\frac{1}{2}}A\tau_{\alpha/2}/2B(F)$, and define $H_a(x)$, $H_{n,a}(x)$, $F_n(x+a)$ as in section 4. Then $\tilde{T}_n = \int_0^\infty J(\frac{n}{n+1} H_{n,a}(x)) dF_n(x+a)$, and it can be shown by some standard computations as in section 4 that $T_{n,\alpha}^{(1)} = \int_0^\infty J(H_a(x)) dF(x+a) + o(n^{-\frac{1}{2}})$. Hence, it is enough to show that for every $s > 0$, there exists an n_s such that

$$\begin{aligned}
 (5.4) \quad & P\{n^{\frac{1}{2}} \left| \int_0^\infty J(\frac{n}{n+1} H_{n,a}(x)) dF_n(x+a) - \int_0^\infty J[H_a(x)] dF_a(x) \right| > c_s^{*(1)}(\log n)^2\} \\
 & \leq c_s^{(2)}n^{-s}, \text{ for all } n \geq n_s.
 \end{aligned}$$

Now, we rewrite the first term on the left hand side of (5.4) as

$$\begin{aligned}
 (5.5) \quad & n^{\frac{1}{2}} \int_0^\infty J(\frac{n}{n+1} H_{n,a}(x)) d[F_n(x+a) - F(x+a)] \\
 & + n^{\frac{1}{2}} \int_0^\infty \{J[\frac{n}{n+1} H_{n,a}(x)] - J[H_a(x)]\} dF(x+a) = I_1 + I_2 \text{ (say)}.
 \end{aligned}$$

Then, by lemma 4.2 and (2.7) we have,

$$\begin{aligned}
 (5.6) \quad n^{-\frac{1}{2}}|I_1| &= \left| \frac{n}{n+1} \int_0^\infty [F_n(x+a) - F(x+a)] J' \left[\frac{n}{n+1} H_{n,a}(x) \right] dH_{n,a}(x) \right| \\
 &\leq \frac{1}{n+1} \sum_{i=1}^n |F_n(X_i+a) - F(X_i+a)| \left| J' \left[\frac{R_{ni}^{(a)}}{n+1} \right] \right| \\
 &\leq \sup_{1 \leq i \leq n} |F_n(X_i+a) - F(X_i+a)| \frac{1}{n+1} \sum_{i=1}^n J' \left(\frac{i}{n+1} \right) \\
 &\leq O(n^{-\frac{1}{2}} \sqrt{\log n}) \cdot \frac{K}{n+1} \sum_{i=1}^n \frac{n+1}{n+1-i} \\
 &= O(n^{-\frac{1}{2}} (\log n)^{3/2}), \text{ with probability } \geq 1 - O(n^{-s}),
 \end{aligned}$$

where $R_{ni}^{(a)} = \text{Rank}\{|X_i - a| : |X_1 - a|, \dots, |X_n - a|\}$. Again, on the same line as in (4.29) through (4.37), it follows that

$$(5.7) \quad |I_2| = \varepsilon_s (n^{-\frac{1}{2}} (\log n)^{3/2}), \text{ with probability } \geq 1 - O(n^{-s}),$$

where ε_s is some fixed positive number. This completes the proof of the lemma.

LEMMA 5.2. For every $s(>0)$, there exists an n_s , such that for $n \geq n_s$,

$$(5.8) \quad P\{B(F)n^{\frac{1}{2}}(\hat{\theta}_{U,n} - \hat{\theta}_{L,n})/A_{\tau_{\alpha/2}} = 1 + O(n^{-\frac{1}{4}}(\log n)^3) \geq 1 - O(n^{-s})\} \geq 1 - O(n^{-s}).$$

Proof. By virtue of lemma 5.1, we have with probability $\geq 1 - O(n^{-s})$,

$$\begin{aligned}
 (5.9) \quad \theta - n^{-\frac{1}{2}}\{(A_{\tau_{\alpha/2}}/2B(F)) + c_s^{(1)}(\log n)^2\} &\leq \hat{\theta}_{L,n} \leq \hat{\theta}_{U,n} \\
 &\leq \theta + n^{-\frac{1}{2}}\{(A_{\tau_{\alpha/2}}/2B(F)) + c_s^{(1)}(\log n)^2\}.
 \end{aligned}$$

Hence, the proof directly follows from theorem 4.3.

LEMMA 5.3. $\{\hat{\theta}_{L,n}\}$ and $\{\hat{\theta}_{U,n}\}$ are uniformly continuous with respect to $\{n^{-1/2}\}$ in the sense that for every ϵ and η (both positive), there exists a $\delta(>0)$, such that as $n \rightarrow \infty$,

$$(5.10) \quad P\left\{ \sup_{|n'-n| < \delta n} |\sqrt{n} (\hat{\theta}_{L,n'} - \hat{\theta}_{L,n})| > \eta \right\} < \epsilon ;$$

$$(5.11) \quad P\left\{ \sup_{|n'-n| < \delta n} |\sqrt{n} (\hat{\theta}_{U,n'} - \hat{\theta}_{U,n})| > \eta \right\} < \epsilon .$$

Proof. Write $n^{1/2}(\hat{\theta}_{L,n'} - \hat{\theta}_{L,n}) = (n/n')^{1/2} \sqrt{n'} (\hat{\theta}_{L,n'} - \theta) - n^{1/2}(\hat{\theta}_{L,n} - \theta)$. By lemma 5.1, theorem 4.3 and (2.13), we have with probability $\geq 1 - O(n^{-s})$,

$$\begin{aligned} \sqrt{n} (\hat{\theta}_{L,n'} - \hat{\theta}_{L,n}) &= \frac{1}{2}(A\tau_{\alpha/2}/B(F)) [(n'/n)^{1/2} - 1] \\ &\quad + B^{-1}(F) [n^{1/2}\{T_n(X_{\tilde{n}} - \theta 1_{\tilde{n}}) - T_{n'}(X_{\tilde{n}'} - \theta 1_{\tilde{n}'})\}] + o(1). \end{aligned}$$

By lemma 5.1, theorem 4.3 and (2.13), we have, with probability $\geq 1 - O(n^{-s})$,

$$\begin{aligned} \sqrt{n} (\hat{\theta}_{L,n'} - \hat{\theta}_{L,n}) &= \frac{1}{2}(A\tau_{\alpha/2}/B(F)) [(n'/n)^{1/2} - 1] \\ &\quad + B^{-1}(F) [\sqrt{n}\{T_n(X_{\tilde{n}} - \theta 1_{\tilde{n}}) - T_{n'}(X_{\tilde{n}'} - \theta 1_{\tilde{n}'})\}] + o(1). \end{aligned}$$

Thus it suffices to prove that

$$(5.12) \quad P\left\{ \sup_{|n'-n| < \delta n} \sqrt{n} |T_n(X_{\tilde{n}}) - T_{n'}(X_{\tilde{n}'})| > \eta | \theta = 0 \right\} < \epsilon .$$

Define T_n^* as in theorem 4.6. Then, routine computation yields that $E(T_n^*) = 0$, $E(T_n^{*2}) = nA_n^2/4$, where A_n^2 is defined by (2.10). As $\{T_n^*, T_n\}$ forms a martingale sequence (cf. theorem 4.6), by using the Kolmogorov inequality for martingales (see Loève (1965, p. 386)), we get,

$$(5.13) \quad P\left\{ \sup_{1 \leq k \leq [\delta n]} |T_{n+k}^* - T_n^*| > t \right\} \leq t^{-2} \{E(T_{n+[\delta n]}^{*2} - T_n^{*2})\} \\ = \frac{1}{2} t^{-2} \{(n+[\delta n])A_{n+[\delta n]}^2 - nA_n^2\} .$$

Put $t = \eta_1 n^{\frac{1}{2}}$ and note that $A_n^2 = A^2 + o(1)$. Then,

$$(5.14) \quad P\left\{ \sup_{1 \leq k \leq [\delta n]} \left| \frac{n+k}{n} T_{n+k} - T_n + \frac{k}{2n} \mu \right| > \eta_1 n^{-\frac{1}{2}} \right\} \leq \frac{1}{2} \eta_1^{-2} \{\delta A^2 + o(1)\} .$$

Since, $k/n \leq \delta$ and $0 \leq T_{n+k} < \mu$, we get,

$$(5.15) \quad P\left\{ \sup_{n < n' < n+[\delta n]} \sqrt{n} |T_n(X_n) - T_{n'}(X_{n'})| > \eta | \theta = 0 \right\} < \epsilon/2 .$$

Proceeding similarly when $n - [\delta n] < n' < n$, we get (5.12). The proof of (5.11) is quite analogous.

We may observe that (5.10) and (5.11) remain true whether (a) or (b) is used in defining $T_n(X_n)$ in (2.2). This follows immediately from (2.15).

LEMMA 5.4. $\lim_{n \rightarrow \infty} P\left\{ \frac{\sqrt{n} (\hat{\theta}_{L,n}^{-\theta}) B(F)}{A/2} + \tau_{\alpha/2} \leq x \right\} = \Phi(x)$, where Φ is the standard normal d.f.

Proof. See Sen [13].

Define $n_i(d) = [v(d)\{1+(-1)^i \epsilon\} + \frac{1+(-1)^i}{2}]$, $i=1,2$. We next prove the following lemma.

LEMMA 5.5. $\sum_{n=n_2}^{\infty} P\{N(d) > n\} < \infty$.

Proof. $\sum_{n=n_2}^{\infty} P\{N(d) > n\} = \sum_{n=n_2}^{\infty} P\{\sqrt{r}(\hat{\theta}_{U,r} - \hat{\theta}_{L,r}) > 2\sqrt{r}d, \text{ for all } r=1, \dots, n\}$ \\ $\leq \sum_{n=n_2}^{\infty} P\{\sqrt{n}(\hat{\theta}_{U,n} - \hat{\theta}_{L,n}) > 2\sqrt{n}d\} .$

Since for $n \geq n_2(d)$, $2d \sqrt{n} \geq 2d \sqrt{n_2(d)} \geq [A\tau_{\alpha/2}/B(F)](1+\epsilon')$, where $\epsilon/3 < \epsilon' < \epsilon/2$ and as by lemma 5.2, $|\sqrt{n}(\hat{\theta}_{U,n} - \hat{\theta}_{L,n}) - A\tau_{\alpha/2}/B(F)| = O(n^{-1/2}(\log n)^3)$, with probability $\geq 1 - O(n^{-s})$ as $n \rightarrow \infty$, it follows that there exists some $n_0 \geq n_2(d)$ such that $\sum_{n=n_0}^{\infty} P\{N(d) > n\} \leq \sum_{n=n_0}^{\infty} O(n^{-s}) < \infty$ for $s > 1$. The lemma follows.

Proof of the main theorem. It follows from lemma 5.2 and definition of N that N is finite a.s. for all $d > 0$ and is a nonincreasing function of d . Lemma 5.5 proves that $EN < \infty$. The remaining statements of (3.3) follow from the definition of N and Monotone Convergence Theorem. We obtain (3.4) from lemma 5.2 and the definition of $v(d)$. (3.5) is a consequence of theorem 1 of Anscombe [1], lemmas 5.2, 5.3 and 5.4.

To prove (3.6), consider

$$(5.16) \quad v^{-1}(d)E\{N(d)\} = v^{-1}(d)\{\Sigma_1 + \Sigma_2 + \Sigma_3 \text{ n } P\{N(d) = n\}\},$$

where Σ_1 extends over all $n \leq n_1(d)$, Σ_2 over all n such that $n_1(d) < n < n_2(d)$ and Σ_3 over all $n \geq n_2(d)$. Since $\lim_{d \rightarrow 0} v(d) = \infty$ and $\lim_{d \rightarrow 0} N(d)/v(d) = 1$ a.s., for every $\epsilon > 0$, there exists a value of d , say d_0 , such that for all $0 < d \leq d_0$, $P\{n_1(d) \leq N(d) \leq n_2(d)\} \geq P\{|\frac{N(d)}{v(d)} - 1| < \epsilon\} \geq 1 - \eta$, η being arbitrarily small. Hence, for $d \leq d_0$,

$$(5.17) \quad v^{-1}(d)\Sigma_1 \text{ n } P\{N(d)=n\} \leq (1-\epsilon)P\{N(d) \leq n_1(d)\} < \eta(1-\epsilon).$$

Also, for all $n_1(d) \leq n \leq n_2(d)$, $|\frac{n}{v(d)} - 1| < \epsilon$, and hence,

$$(5.18) \quad |v^{-1}(d)\Sigma_2 \text{ n } P\{N(d)=n\} - 1| \leq \epsilon \Sigma_2 P\{N(d)=n\} + \eta \leq \epsilon + \eta.$$

Finally,

$$(5.19) \quad v^{-1}(d) \sum_3^n P\{N(d)=n\} = v^{-1}(d) \sum_3 P\{N(d)>n\} + v^{-1}(d) n_2(d) P\{N(d) = n_2(d)\} \\ \leq v^{-1}(d) \sum_3 P\{N(d)>n\} + (1+\epsilon+v^{-1}(d)) P[N(d)=n_2(d)].$$

Since, $v(d) \rightarrow \infty$ as $d \rightarrow 0$, using lemma 5.5, the first term on the right hand side of (5.19) $\rightarrow 0$ as $d \rightarrow 0$. Using the same argument as used in the lemma 5.5, $P[N_2(d)=n_2(d)] \leq P\{N(d) \geq n_2(d)\} \rightarrow 0$ as $d \rightarrow 0$. (3.6) is now proved. Q.E.D.

6. Asymptotic Relative Efficiency (ARE). Suppose we have two bounded length confidence interval procedures A and B for estimating the median of a symmetric population by means of an interval of length $\leq 2d$ ($d > 0$); if N_A and N_B denote the stopping variables and P_A and P_B the coverage probabilities of the two procedures A and B respectively, then we define the ARE of the procedure A with respect to the procedure B by

$$(6.1) \quad e_{A,B} = \lim_{d \rightarrow 0} (EN_B/EN_A),$$

provided $\lim_{d \rightarrow 0} P_A = \lim_{d \rightarrow 0} P_B$ and either of the limit exists.

Let s and c stand for the procedures suggested by us and that by Chow and Robbins [4] respectively. Using (3.5) and (3.6) of theorem 3.1 and the corresponding results of [4], we get under the assumption that $\sigma^2 = \text{Var}(X_1) < \infty$,

$$(6.2) \quad e_{s,c} = \sigma^2 B^2(F) / A^2.$$

The above is the Pitman-efficiency of a general rank order test with respect to Student's t-test. In the particular case when $J(u) = \phi^{-1}(\frac{1+u}{2})$, where ϕ is the d.f. of a standard normal variable, (2.5)-(2.7) are satisfied and it has been proved by Chernoff and Savage (1958) that $\sigma^2 B^2(F) \geq A^2$ for all d.f. F with a density f and a finite second moment, equality being attained when

and only when F is a normal $(0, \sigma^2)$ d.f. Hence, in that case, the ARE of our procedure with respect to the Chow and Robbins procedure (see [4]) is ≥ 1 , equality being attained if and only if the parent df is normal.

In the particular case, when $J(u)=u$, $e_{sc} = 12\sigma^2 \left[\int_{-\infty}^{\infty} f^2(x) dx \right]^2$ and this includes Geertsema's [6] $e(W, M)$ expression as a particular case of the sequential procedure suggested by us.

REFERENCES

- [1] ANSCOMBE, F. J. (1952). Large-sample theory of sequential estimation. Proc. Camb. Phil. Soc. 48 600-607.
- [2] BAHADUR, R. R. (1966). A note on quantiles in large samples. Ann. Math. Statist. 37 577-580.
- [3] CHERNOFF, H. AND SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Statist. 29 972-994.
- [4] 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.
- [5] FARRELL, R. H. (1966). Bounded length confidence intervals for the p-point of a distribution function, III. Ann. Math. Statist. 37 586-592.
- [6] GEERTSEMA, J. C. (1968). Sequential confidence intervals based on rank tests. Unpublished doctoral dissertation, University of California, Berkeley.
- [7] HÁJEK, J. AND ŠIDÁK, Z. (1967). Theory of Rank Tests. Academic Press, New York.
- [8] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. Ann. Math. Statist. 19 293-325.
- [9] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. 58 13-29.
- [10] JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. Ann. Math. Statist. 40, in press.
- [11] LOÈVE, M. (1963). Probability Theory (3rd Edition) Van Nostrand, Princeton.
- [12] PURI, M. L. (1964). Asymptotic efficiency of a class of c-sample tests. Ann. Math. Statist. 35 102-121.
- [13] SEN, P. K. (1966). On a distribution-free method of estimating asymptotic efficiency of a class of nonparametric tests. Ann. Math. Statist. 37 1759-1770.
- [14] SPROULE, R. N. (1969). A sequential fixed-width confidence interval for the mean of a U-statistic. Unpublished doctoral dissertation, University of North Carolina.