ON BOUNDED LENGTH CONFIDENCE INTERVAL FOR THE REGRESSION COEFFICIENT BASED ON A CLASS OF RANK STATISTICS

by

Malay Ghosh and Pranab Kumar Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 680

April 1970
ON BOUNDED LENGTH CONFIDENCE INTERVAL FOR THE
REGRESSION COEFFICIENT BASED ON A CLASS OF RANK STATISTICS

BY MALAY GHOSH AND PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

SUMMARY. Bounded length (sequential) confidence intervals for the regression
coefficient (in a simple linear regression model) based on a class of robust
rank statistics are considered here, and their various asymptotic properties
are studied. In this context, several strong convergence results on some simple
and weighted empirical processes as well as on a class of rank order processes
are established. Comparison with an alternative procedure based on the least
squares estimators is also made.

1. INTRODUCTION

For the simple linear regression model \( X_i = \beta_0 + \beta c_i + e_i, \quad i=1,2,\ldots, \)
where the \( c_i \) are known regression constants and the \( e_i \) are independent and identically dis-
tributed random variables (iidrv) with an absolutely continuous (unknown) cumula-
tive distribution function (cdf) \( F(x) \), defined on the real line \((\mathbb{R}, \mathbb{C})\), we desire
to provide a robust confidence interval (of confidence coefficient \( 1-\alpha, \quad 0<\alpha<1 \))
for the regression coefficient \( \beta \), such that the length of this confidence interval
is bounded above by \( 2d \), for some specified \( d>0 \). The proposed procedure rests on
the use of a class of regression rank statistics [due to Hájek (1962, 1968)] for
the derivation of robust confidence intervals for \( \beta \) [cf. Sen (1969, section 4)],
as extended here to the sequential case along the lines of Chow and Robbins (1965).

Several results, needed for this sequential extension, are derived here.
First an elegant result of Jurečková (1969) on the weak convergence of a class of

*Work supported by the National Institutes of Health, Grant GM-12868.
rank order processes to some appropriate linear processes is strengthened here to almost sure (a.s.) convergence under more stringent regularity conditions on the $c_i$ and the score-functions underlying the rank statistics. This result along with a martingale (or semi-martingale) property of the regression rank statistics, deduced in section 4, guarantees the "asymptotic (as $d \to 0$) consistency" and "efficiency" [see Chow and Robbins (1965)] of the proposed sequential procedure, and also enables one to study its asymptotic relative efficiency (ARE) with respect to the procedure by Gleser (1965) and Albert (1966), which are based on the least squares estimators. In this context, several well-known rank statistics are considered and the allied ARE results are briefly presented. In particular, for the so-called normal scores statistic, it is shown that the ARE is bounded below by 1, uniformly in a broad class of $\{F\}$.

2. NOTATIONS, ASSUMPTIONS AND THE MAIN THEOREM

In accordance with the model of section 1, consider a sequence $\{X_1, X_2, \ldots\}$ of independent random variables for which

$$F_1(x) = P(X_1 \leq x) = F(x - \beta_0 - \beta c_i), \quad i = 1, 2, \ldots, \quad (2.1)$$

where $\beta_0$ is a nuisance parameter. We intend to determine a confidence interval $I_n = \{\beta: \hat{\beta}_L, n < \beta < \hat{\beta}_U, n\}$ (where $\hat{\beta}_L, n$ and $\hat{\beta}_U, n$ are sample statistics) such that

$$P(\beta \in I_n) = 1 - \alpha, \quad \text{the preassigned confidence coefficient}, \quad (2.2)$$

$$0 < \hat{\beta}_U, n - \hat{\beta}_L, n < 2d, \quad \text{for some predetermined } d(>0). \quad (2.3)$$

Since $F$ is not known, no fixed-sample size procedure sounds valid for all $F$. It is therefore desired to determine sequentially a stopping variable $N$ (a positive integer) and the corresponding $(\hat{\beta}_L, N, \hat{\beta}_U, N)$, such that (2.2) and (2.3)
hold. Our proposed procedure is based on the following class of regression
rank statistics.

In a sample \( X_n = (X_1, \ldots, X_n) \) of size \( n > 1 \), let \( R_{ni} = \sum_{j=1}^{n} u(X_i - X_j) \) [where
\( u(t) \) is 1 or 0 according as \( t \geq 0 \) or \( t < 0 \)] be the rank of \( X_i, 1 \leq i \leq n \). Let

\[
c_{ni}^* = (c_i - \overline{c}_n)/c_n, \quad 1 \leq i \leq n,
\]
where \( \overline{c}_n = n^{-1} \sum_{i=1}^{n} c_i \) and \( c_n^2 = \sum_{i=1}^{n} (c_i - \overline{c}_n)^2 \). \( \quad \) (2.4)

Then, as in Hájek (1962, 1968), a regression rank statistic is defined as

\[
T_n = T(X_n) = \sum_{i=1}^{n} c_{ni}^* J_n R_{ni} / (n+1),
\]
where the "scores" \( J_n(i/(n+1)), 1 \leq i \leq n \), are generated by a "score-function"

\( \{J(u); \ 0 < u < 1 \} \) in the following manner. We let \( J_n(u) = J_n(i/(n+1)), \) for \( (i-1)/n < u < i/n, \)

\( 1 \leq i \leq n \), and define

\[
J_n(i/(n+1)) \quad \text{as equal to} \quad J(i/(n+1)) \quad \text{or} \quad EJ(U_{ni}), \quad 1 \leq i \leq n,
\]
where \( U_{n1} < \ldots < U_{nn} \) are the \( n \) ordered random variables in a sample of size \( n \) from
the rectangular \((0,1)\) distribution. The score-function \( J(u) \) is defined as

\( \Psi^{-1}(u); \ 0 < u < 1 \), where \( \Psi(x) \) is an absolutely continuous cdf satisfying the two
conditions

\[
(1) \quad M(t) = \int_{-\infty}^{\infty} e^{tx} d\Psi(x) = \infty \quad \text{for all} \quad 0 < t < t_0,
\]

(2.7)

\[
(2) \quad \lim_{u \to 0} \left[ \Psi^{-1}(u) / u, \Psi^{-1}(1-u) / (1-u) \right] \geq K > 0,
\]

(2.8)

where \( \Psi(x) = d\Psi(x)/dx \). Thus, we have \( M(t) = \int_{0}^{1} e^{tJ(u)} du < \infty \) for all \( 0 < t < t_0 \), and

by (2.8),

\[
|J(u)| \leq K_o [-\log u(1-u)], \quad |J'(u)| \leq K_o [u(1-u)]^{-1}, \quad 0 < u < 1; \quad K_o < \infty.
\]

(2.9)
These assumptions appear to be more restrictive than those in Hájek (1962, 1968) and Jurečková (1969) [where the weak convergence results are only studied], but are somewhat needed for our purpose, as we require, in the sequel, to establish certain a.s. convergence results where the exponential bounds [to follow from (2.9)] are quite useful. It may be noted that (2.7) and (2.8) hold for the entire class of normal, logistic, double-exponential, exponential, rectangular and many other cdf's. In fact, if we use \( J(u) = u \) [i.e., \( \Psi \) rectangular cdf] or \( J(u) \) as the inverse of the standard normal cdf, we obtain respectively the Wilcoxon and the normal scores; the corresponding \( T_n \) are termed the Wilcoxon and the normal scores regression rank statistics. For both of these, (2.7) and (2.8) hold.

Regarding \( \{c_n\} \) [where \( c_n = (c_1, \ldots, c_n) \)], we make the following assumptions, where the first one is due to Hájek (1968):

\[
(i) \quad \max_{1 \leq i \leq n} |c_{ni}^*| = O(n^{-\frac{1}{2}}), \quad (2.10)
\]

\[
(ii) \quad \liminf_{n \to \infty} n^{-1} c_n^2 c_n^2 > 0, \quad (2.11)
\]

\[
(iii) \quad c_n^2 = Q(n) \text{ is a strictly increasing function of } n, \text{ such that for every } a > 0,
\]

\[
\lim_{n \to \infty} Q(a, n)/Q(n) = s(a) \text{ whenever } \lim_{n \to \infty} a_n = a, \quad (2.12)
\]

where \( s(a) \) is strictly monotone (increasing) with \( s(1) = 1 \). The condition (2.10) is again more stringent than the classical Noether-condition [cf. Hájek (1962, 1968)], but is satisfied in the majority of practical situations. (2.11) is less restrictive than the parallel condition: \( \lim_{n \to \infty} n^{-1} c_n^2 c_n^2 > 0 \), made by Gleser (1965) and Albert (1966) in connection with least squares theory. For example,
if $c_i = a+(i-1)h$, $h>0$, $i=1,2,...$, (2.10) and (2.11) hold but not the Gleser-condition. Also, in this case, $Q(n) = nh^2(n^2-1)/12$, so that $s(a) = a^3$.

Finally, regarding the cdf $F$ in (2.1), we assume that $F \in \mathcal{F}(\mathcal{Y})$, where $\mathcal{F}(\mathcal{Y})$ is the case of all absolutely continuous $F$ for which the density function $f(x)$ and its first derivative $f'(x)$ are bounded for almost all $x(a \leq a \leq x)$ and further

$$\lim_{x \to \infty} f(x) J'[F(x)]$$

are bounded. (2.13)

From (2.1), it follows that under $H_0: \beta=0$, implying that $X_1,..., X_n$ are iidrv, $T(X_n)$ has a completely specified distribution generated by the $n!$ equally likely realizations of $(R_{n1},..., R_{nn})$ over the permutations of $(1,..., n)$. Hence, there exists two known constants $T_n^{(1)}$ and $T_n^{(2)}$ (depending on $c_n$), and an $\alpha_n$ (known) such that

$$P(T_n^{(1)} \leq T(X_n) \leq T_n^{(2)} | H_0) = 1-\alpha_n (\rightarrow 1-\alpha \text{ as } n \to \infty)$$

(2.14)

[For small $n$, $\alpha_n$ may not be equal to $\alpha$]. If we let

$$A_n^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} [J_i \left(\frac{i}{n+1}\right) - \bar{J}_n]^2; \bar{J}_n = \frac{1}{n} \sum_{i=1}^{n} J_i \left(\frac{i}{n+1}\right),$$

(2.15)

then from the classical Wald-Wolfowitz-Noether-Hoeffding-Hájek permutational central limit theorem, (cf. [8, p. 160]), we have

$$\lim_{n \to \infty} \{T_n^{(2)}/A_n\} = \lim_{n \to \infty} \{-T_n^{(1)}/A_n\} = \tau_\alpha/2,$$

where $\Phi(\tau_\alpha) = 1-\epsilon$ and $\Phi(x)$ is the standard normal cdf.

It follows from Sen (1969, Section 6) that $T(X_n-a \cdot c_n)$ is $\permr$ in $a: -\infty < a < \infty$.

Hence, if we let

$$\hat{S}_{L,n} = \sup\{a: T(X_n-a \cdot c_n) > T_n^{(2)}\}, \hat{S}_{U,n} = \inf\{a: T(X_n-a \cdot c_n) < T_n^{(1)}\},$$

(2.17)

it follows as in Sen (1969, Section 4) that

$$P(\hat{S}_{L,n} < \beta \leq \hat{S}_{U,n} | \beta) = 1-\alpha_n (\rightarrow 1-\alpha \text{ as } n \to \infty).$$

(2.18)
We are now in a position to define our sequential procedure. For every $d > 0$, let $N(d)$, the stopping variable, be the smallest positive integer [$\geq n_o$, an initial sample size ($\geq 3$)] for which $\hat{\beta}_{U,N(d)} \leq \beta_{L,N(d)} \leq 2d$. Then our proposed confidence interval for $\beta$ is $I_{N(d)} = \{\beta: \beta_{L,N(d)} \leq \beta < \beta_{U,N(d)}\}$ and is based on the stopping variable $N(d)$. We justify the proposed procedure on the ground of its robustness (for outliers or gross errors etc) and its asymptotic properties considered in the following theorem, sketched in the same fashion as in Chow and Robbins (1965).

Theorem 2.1. Under the assumptions made above, $N(d)$ is a non-increasing function of $d(>0)$, it is finite a.s., $EN(d) < \infty$ for all $d > 0$, $\lim_{d \to 0} N(d) = \infty$ a.s., and $\lim_{d \to 0} EN(d) = \infty$. Further,

$$\lim_{d \to 0} N(d)/Q^{-1}(v(d)) = 1 \text{ a.s.,}$$  \hspace{1cm} (2.19)

$$\lim_{d \to 0} P\{\beta \in I_{N(d)}\} = 1 - \alpha \text{ for all } F \in \mathcal{F}(\mathcal{Y}),$$  \hspace{1cm} (2.20)

$$\lim_{d \to 0} \frac{EN(d)}{Q^{-1}(v(d))} = 1,$$  \hspace{1cm} (2.21)

where

$$v(d) = \left[\frac{1}{2}(d/\alpha)^2\right][dB(F)]^2, \quad B(F) = \int_{-\infty}^{\infty} (d/dx)J[F(x)]dF(x), \quad (2.22)$$

$$A^2 = \int_{0}^{1} J^2(u)du - \mu^2 \quad \text{and} \quad \mu = \int_{0}^{1} J(u)du.$$  \hspace{1cm} (2.23)

The proof of the theorem is postponed to section 5; certain other results needed in this context and having importance of their own are derived in the next two sections.

3. ASYMPTOTIC BEHAVIOUR OF SOME EMPIRICAL PROCESSES

It has been shown by Jurečková (1969) that under $H_o: \beta = 0$, for all real and finite $b$, denoting $W(X_n,b) = T(X_n) - T(X_n - bc_n) - bB(F)$,
our primary concern is to strengthen this statement to almost sure convergence, specifying the order of convergence as well as extending the domain of \( b \) to 

\[
|b| \leq C \log n \quad k \geq 1, \quad 0 < C < \infty. \quad \text{We define now}
\]

\[
H_n(x; b) = \frac{1}{n+1} \sum_{l=1}^{n} u(x + bc_n, X_l) - X_l, \quad S^*(x; b) = \sum_{l=1}^{n} c_n u(x + bc_n, X_l)
\]

for \(-\infty < x < \infty\) and \(-\infty < b < \infty\). Then, by (2.5) and (3.2), we have 

\[
T(X_n - bc^*_n) = \int_{-\infty}^{\infty} [H_n(x; b)] dS^*(x; b).
\]

It will be convenient for us to study first the asymptotic behaviour of the two processes in (3.2).

Let \( \{Y_1, Y_2, \ldots\} \) be a sequence of iidrv having the rectangular \((0, 1)\) distribution. For every \( t: \quad 0 < t < 1, \) define 

\[
d_{n_i}(t; b) = \mathbf{P}(F^{-1}(t) + bc_n) - t, \quad 1 \leq i \leq n, \quad -\infty < b < \infty.
\]

It follows that for every \( i: \quad 1 \leq i \leq n \) and \( b: \quad -\infty < b < \infty, \)

(i) \( 0 < t + d_{n_i}(t; b) \leq 1; \quad t + d_{n_i}(t; b) \) is \( \uparrow \) in \( t: \quad 0 < t < 1, \)

(ii) \( d_{n_i}(t; b) \) and \( bc_n \) have the same sign,

(iii) \( \bar{d}_n(t; b) = n^{-1} \sum_{l=1}^{n} d_{n_l}(t, b) = b^2(0(n^{-1}) \mathbf{P} F \mathbf{F}(\Psi).
\]

Consider then the two stochastic processes 

\[
G^*_n(t; b) = \sum_{l=1}^{n} c_n u(t + d_{n_l}(t; b) - Y_l) - \sum_{l=1}^{n} c_n d_{n_l}(t; b), \quad -\infty < b < \infty, \quad 0 < t < 1.
\]
\[ L^*_n(t; b) = (n+1)^{-\frac{k}{2}} \sum_{i=1}^{n} c^*_n d_n^i (t; b) - Y_i, \quad -\infty < b < \infty, \quad 0 < t < 1. \] (3.9)

Simple computations yield that

\[ \zeta_n(t; b) = \sum_{i=1}^{n} c_n^i d_n^i (t; b) = b f(F^{-1}(t)) + b^2 O(n^{-k}); \] (3.10)

\[ G_n^*(t; b) = \sum_{i=1}^{n} G_n^i (F^{-1}(t); b) - b f(F^{-1}(t)) + b^2 O(n^{-k}); \] (3.11)

\[ L_n^*(t; b) = (n+1)^{-\frac{k}{2}} H_n (F^{-1}(t); b), \quad 0 < t < 1, \quad -\infty < b < \infty. \] (3.12)

**Theorem 3.1.** For every \( h > 0 \), there exists two positive constants \( (K_1, K_2) \) and an \( n^* \) (which all may depend on \( h \)), such that for \( n > n^* \),

\[ P \left\{ \sup_{0 < t < 1} \sup_{b \in I_n^*} \left| G_n^*(t; b) - G_n^*(t; 0) \right| \geq K_1 n^{-\delta} (\log n)^k \leq K_2 n^{-h} \right\}, \] (3.13)

\[ P \left\{ \sup_{0 < t < 1} \sup_{b \in I_n^*} \left| L_n^*(t; b) - L_n^*(t; 0) \right| \geq K_1 n^{-\delta} (\log n)^k \leq K_2 n^{-h} \right\}, \] (3.14)

where \( \delta > 0, \, k > 1 \) and \( I_n^* = \{ b : |b| < C (\log n)^k, \, 0 < C < \infty \} \).

**Proof.** Let \( a_n = C (\log n)^k \), \( r_n = [n^{-\delta}] \) and \( n_{r,n} = n_{a_n} / r_n \), \( r = 0, 1, \ldots, \lfloor r_n \rfloor \), where \( 0 < \delta < 1 \) and \( [s] \) denotes the integral part of \( s > 0 \). Then, \( n_{r+1,n} - n_{r,n} = O(n^{-\delta} (\log n)^k) \), for all \( r = -r_n, \ldots, -1 \). Also, let \( s_n = [n^{\frac{k}{2}+\delta}] \) and \( n_{s,n} = s_n / n \), \( s = 0, 1, \ldots, s_n \). Now, for \( b \in [n_{r,n}, n_{r+1,n}] \), we have

\[ c_n^i [u(t+d_n^i (t; n_{r,n}) - Y_i) - u(t-Y_i)] \leq c_n^i [u(t+d_n^i (t; b) - Y_i) - u(t-Y_i)] \]

\[ \leq c_n^i [u(t+d_n^i (t; n_{r+1,n}) - Y_i) - u(t-Y_i)], \quad 1 \leq i \leq n, \] (3.15)

\[ \zeta_n(t; n_{r,n}) \leq \zeta_n(t; b) \leq \zeta_n(t; n_{r+1,n}) \] (3.16)
\[ \zeta_n(t; \eta_{r+1}, n) - \zeta_n(t; \eta_r, n) = O(n^{-\delta_1 (\log n)^k}), \quad (3.17) \]

for all \( r = -r_n, \ldots, r_n - 1, 0 < t < 1 \). Hence, for \( b \in [\eta_r, n; \eta_{r+1}, n] \)

\[
\left| G_n(t; b) - G_n(t; 0) \right| \leq \max_{r, r+1} \left| G_n(t; \eta_{j, n}) - G_n(t; 0) \right| \\
\quad + O(n^{-\delta_1 (\log n)^k}), \quad -r_n \leq r < r_n - 1. \quad (3.18)
\]

We denote by \( S_1(n) = \{ i : c_{ni}^* \geq 0, 1 \leq i \leq n \}, S_2(n) = \{ i : c_{ni}^* < 0, 1 \leq i \leq n \} \),

\[
V_{b,n}^{(s)} = \sum_{i \in S_1(n)} c_{ni}^*[u(\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b) - Y_i) - u(\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b) - Y_i)] \text{ and}
\]

\[
V_{b,n}^{(s)} = \sum_{i \in S_1(n)} c_{ni}^*[\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b) - \xi_{s,n} - d_{ni}(\xi_{s,n}; b)], \text{ for } s = 0, 1, \ldots, s_n - 1,
\]

\( b \in I_n \). Then, we get after some simple algebraic manipulations that

\[
\sum_{i=1}^{n} c_{ni}^*[u(\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b) - Y_i) - (\xi_{s+1,n} + d_{ni}(\xi_{s,n}; b))] - V_{b,n}^{(s)} \leq \sum_{i=1}^{n} c_{ni}^*[u(t + d_{ni}(t; b) - Y_i) - (t + d_{ni}(t; b))]
\]

\[
\leq \sum_{i=1}^{n} c_{ni}^*[u(\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b) - Y_i) - (\xi_{s+1,n} + d_{ni}(\xi_{s+1,n}; b))] \\
\quad + V_{b,n}^{(s)} + U_{b,n}^{(s)}.
\]

Thus,

\[
\left| \sum_{i=1}^{n} c_{ni}^*[u(t + d_{ni}(t; b) - Y_i) - (t + d_{ni}(t; b))] \right| \\
\leq \max_{j=s, s+1, \ldots, s_n} \left| \sum_{i=1}^{n} c_{ni}^*[u(\xi_{j,n} + d_{ni}(\xi_{j,n}; b) - Y_i) - (\xi_{j,n} + d_{ni}(\xi_{j,n}; b))] \right| \\
\quad + V_{b,n}^{(s)} + U_{b,n}^{(s)}.
\quad (3.19)
\]

Hence, by (3.18) and (3.19), we have,
\[ \sup_{0 \leq t \leq 1} \sup_{b \leq a_n} |G_n^*(t; b) - G_n^*(t; 0)| \leq \max_{0 \leq s \leq a_n} \max_{\|r\| \leq r_n} |G_n^*(\xi_{s,n}; \eta_{r,n}) - G_n^*(\xi_{s,n}; 0)| + \max_{0 \leq s \leq a_n} \max_{-r_n \leq r \leq -r_{n-1}} |V(s)| + \max_{0 \leq s \leq a_n} \max_{-r_n \leq r \leq -r_{n-1}} |U(s)| \leq \delta n^{-1} (\log n)^k. \] (3.20)

Hence, it suffices to prove that the first three terms on the right hand side of (3.20) are bounded above by \(0(n^{-h}(\log n)^k)\) with probability \(\geq 1 - O(n^{-h})\) for every fixed \(h > 0\).

Consider any fixed \(r\) (say \(\geq 0\)) and any fixed \(s\). Then, \(G_n^*(\xi_{s,n}; \eta_{r,n}) - G_n^*(\xi_{s,n}; 0) = \sum_{i=1}^{n} (Z_{ni} - EZ_{ni})\), where, \(Z_{ni} = c_{ni}^* [u(\xi_{s,n}; \eta_{r,n}) - Y_i] + u(\xi_{s,n} - Y_i)\), \(i = 1, 2, \ldots, n\) are independent and \(Z_{ni}\) can assume the values 0 and \(c_{ni}^*\) respectively with probability \(1 - d_{ni}(\xi_{s,n}; \eta_{r,n})\) and \(d_{ni}(\xi_{s,n}; \eta_{r,n})\) respectively. Hence, for every \(g > 0\), and \(h > 0\),

\[ P(\sum_{i=1}^{n} (Z_{ni} - EZ_{ni}) > g) \leq \exp[-h(g + \sum_{i=1}^{n} EZ_{ni})] \leq \prod_{i=1}^{n} P(\exp(h Z_{ni})) \leq \exp[-h(c_{ni}^*)] \prod_{i=1}^{n} [1 + (\exp[h |c_{ni}^*|] - 1)d_{ni}(\xi_{s,n}; \eta_{r,n})] = \chi_n(g,h), \text{ (say).} \] (3.21)

Upon taking now \(g = g_n = K_n \delta (\log n)^k\), \(h = h_n = n^\delta\), where \(0 < \delta < \delta_1 < k, k \geq 1\), \(0 < K_n < \infty\), we obtain after some simplifications that
\[ \log \chi_n(g_n, h_n) = -K_1 n^{-\delta_1} (\log n)^{k} [1 + o(n^{-\delta_1} (\log n)^{-k})] \]  
\hfill (3.22)\]

uniformly in \( s=0,1,\ldots,s_n \) and \( |r| \leq r_n \). Thus, for \( n \) adequately large, we have, uniformly in \( r \) and \( s \),

\[
P\{ G^*(\xi_n, n; r_n) - G^*(\xi_n, n; 0) > K_1^{\delta_1} (\log n)^k \}
\leq \exp\{-K_1^{\delta_1} (\log n)^k[1+o(1)]\},
\]
\hfill (3.23)

and a similar bound is easily obtained for the left hand tail. Thus, for large \( n \), \( 0 \leq s \leq s_n \), \( |r| \leq r_n \),

\[
P\{ |G^*(\xi_n, n; r_n) - G^*(\xi_n, n; 0)| > K_1^{\delta_1} (\log n)^k \}
\leq 2 \exp\{-K_1^{\delta_1} (\log n)^k[1+o(1)]\}.
\]
\hfill (3.24)

Also, it follows obviously that

\[
\gamma_{(s)}\eta_r^{(n)} = O(n^{-\delta_1}) \text{ uniformly in } |r| \leq r_n, 0 \leq s \leq s_n.
\]
\hfill (3.25)

By similar techniques as in (3.21), we get after putting \( h = n^{\delta_1} \) that for fixed \( r \) and \( s \),

\[
P\{ u_{(s)}^{(r_n, n)} > g_n \} \leq \exp\{-K_1^{\delta_1} (\log n)^k(1+o(n^{-\delta_1} (\log n)^{-k}))\}.
\]
\hfill (3.26)

From (3.24)-(3.26), we obtain (by use of the Bonferroni inequality) that for large \( n \), the left hand side of (3.20) is bounded above by \( K_1^{\delta_1} (\log n)^k \) with probability greater than or equal to
1 - 2(s_n + 1)2r_n \exp\{-K_1 n^{\delta_1 - \delta}(\log n)^k (1 + o(1))\}. \quad (3.27)

Since, \( s_n = \lfloor n^{\delta_1} \rfloor, r_n = n^{\delta_1} \) and \( \delta_1 > \delta, K_1 \) can always be so selected that for \( n \geq n^\star \), (3.27) is bounded above by \( 4 \exp\{-h(\log n)^k\} \leq 4n^{-h} \) (since, \( k \geq 1 \)). This completes the proof of (3.13).

It can be shown by the same technique as in above that

\[
\sup_{0 < t < 1} \sup_{\|b\| \leq a_n} |L_n^\star(t;b) - L_n^\star(t;0)|
\leq \max_{0 < s \leq s_n} \max_{|r| \leq r_n} |L_n^\star(\xi_{s,n}; r, n) - L_n^\star(\xi_{s,n}; 0)|
+ \max_{0 < s \leq s_{n-1}} \left[ |L_n^\star(\xi_{s+1,n}; 0) - L_n^\star(\xi_{n,s}; n)| \right]
+ \max_{0 < s \leq s_n} \max_{|r| \leq r_n} |W_n(s)|, \quad (3.28)
\]

where \( W_n(s) = (n+1)^{\frac{1}{2}} \sum_{i \in S_2} n^{ni}(u(\xi_{s,n} + d_i(\xi_{s,n}; n); -\gamma_i) - u(\xi_{s,n} + d_i(\xi_{s,n}; n); -\gamma_i)) \).

Using (3.9) we obtain that for every \( s_n > 0, h_n > 0, \)

\[
\log P\{L_n^\star(\xi_{s,n}; r, n) - L_n^\star(\xi_{s,n}; 0) > g_n\}
\leq -(n+1)^{\frac{1}{2}} h_n g_n + \left\{ \sum_{i \in S_1(n)} + \sum_{i \in S_2(n)} \log E \exp[h_n \{u(\xi_{s,n} + d_i(\xi_{s,n}; n); -\gamma_i) - u(\xi_{s,n}; -\gamma_i)\}]
\right\}
- \left\{ \sum_{i \in S_1(n)} \log[1 + (e^{h_n - 1})^{d_i(\xi_{s,n}; n; r, n)}]
+ \sum_{i \in S_2(n)} \log[1 - (1 - e^{h_n})(-d_i(\xi_{s,n}; n; r, n))] \right\}. \quad (3.29)
\]
On using the inequality that for $|x| < 1$, $\log(1 \pm |x|) \leq \pm |x|$, and on taking $h_n = n^{-\delta_1}$, $g_n = K_1 n^{-\delta} (\log n)^k$, $(0 < \delta < \delta_1 < \gamma)$, we get for adequately large $n$ that,

$$\log P\{L^*_n(\xi_{s,n}; \eta_{r,n}) - L^*_n(\xi_{s,n}; 0) > K_1 n^{-\delta} (\log n)^k\}$$

$$\leq -K_1 n^{\frac{1}{2} - \delta - \delta_1} (\log n)^k + n^{\delta_1} \{ \sum_{i=1}^{n} d_{ni}(\xi_{s,n}; \eta_{r,n}) \}$$

$$+ o(n^{-\frac{1}{2}}) \sum_{i=1}^{n} |d_{ni}(\xi_{s,n}; \eta_{r,n})|$$

$$= -K_1 n^{\frac{1}{2} - \delta - \delta_1} (\log n)^k + n^{\delta_1} \{ o(n^{-\frac{1}{2}}) (\log n)^{2k} \}$$

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

by (2.10), (3.4) and (3.7).

(3.30)

Since $\delta < \delta_1$, if we let $\delta_1 < \gamma$, the right hand side of (3.30) is

$$-K_1 n^{\frac{1}{2} - \delta - \delta_1} (\log n)^k [1 + o(n^{-\frac{1}{2} + \delta} (\log n)^k) + o(n^{-\frac{1}{2} - \delta})],$$

(3.31)

and hence, for $n$ sufficiently large,

$$P\{L^*_n(\xi_{s,n}; \eta_{r,n}) - L^*_n(\xi_{s,n}; 0) > K_1 n^{-\delta} (\log n)^k\}$$

$$\leq \exp\{-K_1 n^{-\frac{1}{2} - \delta_1} (\log n)^k \}, \quad 0 < K_1 < \infty,$$

(3.32)

and the same bound holds for the left hand tail. Hence, by the Bonferroni inequality, for large $n$,
\[
\begin{align*}
\mathbb{P}_n \left\{ \max_{0 \leq s \leq s_n} \max_{|r| \leq r_n} \left| L_n^*(\xi_{s,n}; r, n) - L_n^*(\xi_{s,n}; 0) \right| > K_n^{-\delta} (\log n)^k \right\} & \\
& \leq 4r_n(s_n + 1) \exp\left\{ -K'_n (\log n)^{k-1} \right\} \\
& \leq 4n^{-h}, \text{ for } k \geq 1 \text{ and } n \geq n^*, \tag{3.33}
\end{align*}
\]

where \( h \) is any (fixed) positive number.

Since, \( L_n^*(\xi_{s+1,n}; 0) - L_n^*(\xi_{s,n}; 0) = (n+1)^{-1/2} \sum_{i=1}^{n} [u(\xi_{s+1,n} - Y_i) - u(\xi_{s,n} - Y_i)] \) involves iid \((0,1)\) valued random variables, proceeding as in (3.21)-(3.24) and using the Bonferroni inequality, we obtain

\[
\begin{align*}
\mathbb{P}_n \left\{ \max_{0 \leq s \leq s_n - 1} \left| L_n^*(\xi_{s+1,n}; 0) - L_n^*(\xi_{s,n}; 0) \right| > K_2 n^{-\delta} (\log n)^k \right\} & \\
& \leq 2s_n \exp\left\{ -K_2 n^{-1} (\log n)^{k(1+o(1))} \right\} \\
& \leq 2n^{-h} \text{ for any fixed } h > 0 \text{ by proper choice of } K_2 \text{ (since } k \geq 1\right). \tag{3.34}
\end{align*}
\]

Similarly it can be shown that

\[
\begin{align*}
\mathbb{P}_n \left\{ |W_{r,n}(s)| > K_3 n^{-\delta} (\log n)^k \right\} & \\
& \leq 2 \exp\left\{ -K_3 n^{-1} (\log n)^{k(1+o(1))} \right\}, \tag{3.35}
\end{align*}
\]

uniformly in \( |r| \leq r_n \) and \( 0 \leq s \leq s_n \). As before, on using the Bonferroni inequality, we get

\[
\begin{align*}
\mathbb{P}_n \left\{ \max_{0 \leq s \leq s_n} \max_{|r| \leq r_n} |W_{r,n}(s)| > K_3 n^{-\delta} (\log n)^k \right\} & \\
& \leq 4r_n(s_n + 1) \exp\left\{ -K_3 n^{-1} (\log n)^{k(1+o(1))} \right\} \\
& \leq 4n^{-h}, \text{ for any fixed } h > 0, \text{ as } k \geq 1. \tag{3.36}
\end{align*}
\]
(3.14) then follows from (3.28), (3.33), (3.34) and (3.36). Q.E.D.

Theorem 3.1 will be utilized to prove the following basic result which
strengthens (3.1) to a.s. convergence.

**Theorem 3.2.** Under the assumptions of Section 2, for every \( s(>0) \), there exists
positive constants \( (c_s^{(1)}, c_s^{(2)}) \) and a sample size \( n_s \), such that \( \beta = 0 \) and all
\[ n > n_s \]

\[ P\left( \sup_{b \in I^*_n} |W(X_n, b)| > c_s^{(1)} n^{-\delta} (\log n)^{k+1} < c_s^{(2)} n^{-s} \right) \]

Hence, \( W(X_n, b) \to 0 \) a.s., as \( n \to \infty \).

**Proof.** Note that if we let \( J_n \left( \frac{i}{n+1} \right) = \text{E}J(U_{ni}), 1 \leq i \leq n \) [cf. (2.6)], we have

\[ \left| \sum_{i=1}^{n} c_n \frac{R_{ni}}{n!} \right| \left[ J_n \left( \frac{ni}{n+1} \right) - J \left( \frac{ni}{n+1} \right) \right] \]

\[ \leq \sum_{1 \leq i \leq n} \left| c_n \frac{R_{ni}}{n!} \right| \left[ J_n \left( \frac{i}{n+1} \right) - J \left( \frac{i}{n+1} \right) \right] \]

\[ = o(n^{-k}) \cdot o(n^2) = o(1), \quad (3.37) \]

by (2.10) and theorem 2 of Chernoff and Savage (1958). Hence, for our purpose,
it suffices to work with \( J_n \left( \frac{i}{n+1} \right) = J \left( \frac{i}{n+1} \right), 1 \leq i \leq n \). If we define

\[ R_{ni}^{(b)} = \sum_{j=1}^{n} u(x_j - bc_{n1}^*) - [x_j - bc_{n1}^*], 1 \leq i \leq n, -\infty < b < \infty, \]

we have

\[ T_{n1}(x) - T_{n1}(x - bc_{n1}^*) = \int_{-\infty}^{\infty} J[H_n(x, 0)]dS^*(x; 0) \]

\[ - \int_{-\infty}^{\infty} J[H_n(x; b)]dS^*(x; b) \]

\[ = I_{n1}(b) + I_{n2}(b), \quad (3.38) \]
where

$$I_{n1}(b) = \int_{-\infty}^{\infty} J[H_n(x;0)]d[S^*(x;0) - S_n^*(x;b)],$$

$$I_{n2}(b) = \int_{-\infty}^{\infty} \{J[H_n(x;0)] - J[H_n(x;b)]\}dS_n^*(x;b).$$

In (2.1), without any loss of generality, we may let $\beta_0 = 0$ and assume that $0 < F(0) < 1$. We select $x_n^{(1)}$ and $x_n^{(2)}$ such that

$$F(x_n^{(1)}) = 1 - F(x_n^{(2)}) = 4K_1 n^{-(\lambda + \delta)} (\log n)^k,$$

where $K_1$ is defined in (3.14). Then, for $n$ sufficiently large, say for $n > n_0$, we have $x_n^{(1)} < 0 < x_n^{(2)}$, and

$$I_{n2}(b) = \int_{-\infty}^{x_n^{(1)}} + \int_{x_n^{(1)}}^{0} + \int_{0}^{x_n^{(2)}} + \int_{x_n^{(2)}}^{\infty} \{J[H_n(x;0)] - J[H_n(x;b)]\}dS_n^*(x,b)$$

$$= I_{n2}^{(1)}(b) + I_{n2}^{(2)}(b) + I_{n2}^{(3)}(b) + I_{n2}^{(4)}(b), \text{ say.}$$

Since, by (3.2), $|S_n^*(x;b)| \leq \{(n+1) \max_{1 \leq i \leq n} |c_{ni}^*| \}H_n(x;b)$, we obtain by some standard computations that

$$|I_{n2}^{(3)}(b)| \leq (n+1) \max_{1 \leq i \leq n} |c_{ni}^*| \int_{0}^{x_n^{(2)}} |H_n(x;0) - H_n(x;b)| \left| J'[H_n(x;b, \theta_n)] \right| dH_n(x,b),$$

where $H_n(x;b, \theta_n) = \theta H_n(x;0) + (1-\theta) H_n(x;b)$, $0 < \theta < 1$. On using (3.12) and (3.14), we obtain for large $n$
\[ \sup_{-\infty < x < \infty} \sup_{b \in I^*_n} |H_n(x; b) - H_n(x; 0)| = O(n^{-1+\delta} \log n^k), \]  \hspace{1cm} (3.44)

with probability \( \geq 1 - O(n^{-S}) \), for every (fixed) \( s(>0) \). Hence, on using (3.44), (2.9) and (2.10), we obtain that

\[ |I_{n2}^{(3)}(b)| \leq O(n^{-\delta} \log n^k) \int_0^x \{H_n(x; b, \theta_n)[1-H_n(x; b, \theta_n)]\}^{-1} dH_n(x, b), \]  \hspace{1cm} (3.45)

with probability \( \geq 1 - O(n^{-S}) \). Let us define \( \bar{F}_n(x; b) = (n+1)^{-1} \sum_{i=1}^n F(x+bc_i) \), so that \( \bar{F}_n(x; 0) = n/(n+1)F(x) \). Note that by (2.10)-(2.13)

\[ \sup_{b \in I^*_n} |\bar{F}_n(x; b) - \bar{F}_n(x; 0)| = O(n^{-1}(\log n)^{2k}). \]  \hspace{1cm} (3.46)

Since \( H_n(x; 0) = (n+1)^{-1} \sum_{i=1}^n u(x-x_i) \) involves average of iid (bounded valued) rv's, using the results of Hoeffding (1963), we obtain that for large \( n \),

\[ P \{|H_n(x_n^{(2)}; 0) - \bar{F}_n(x_n^{(2)}; 0)| > K_1 n^{-3/4-\delta/2} (\log n)^k \} \leq K_2 n^{-S}, \]  \hspace{1cm} (3.47)

where \( K_1 \) and \( K_2 \) are the same as in (3.14). Using then (3.12), (3.14), (3.41), it follows that for all \( x \leq x_n^{(2)} \),

\[ 1-H_n(x; 0) \geq (1/3)[1-H_n(x; b)], \ \forall b \in I^*_n, \]  \hspace{1cm} (3.48)

with probability \( \geq 1 - O(n^{-S}) \). Again, by using the same results of Hoeffding (1963), we have \( P\{|I_n^*(F(0); 0) - \sqrt{n} F(0)| > K_1 \sqrt{\log n} \} \leq K_2 n^{-S} \), for every \( s(>0) \), and hence, by (3.12), (3.14) and the fact that \( F(0) > 0 \), we obtain that for all \( x > 0 \), as \( n \to \infty \),

\[ H_n(x; b) \geq F(0) - K_1 n^{-k}(\log n)^k [1+O(n^{-\delta})] \geq g_0 > 0, \ \forall b \in I^*_n, \]  \hspace{1cm} (3.49)
with probability \( \geq 1-O(n^{-g}) \). From the definition of \( H_n(x;b,\theta_n) \) and (3.48) - (3.49), it follows that for all \( 0 < x < x_n^{(2)} \),

\[
H_n(x;b,\theta_n)[1-H_n(x;b,\theta_n)] \geq C[1-H_n(x;b)],
\]

(3.50)

(where \( C > 0 \)), with probability \( \geq 1-O(n^{-g}) \). From (3.45) and (3.50), for large \( n \), with probability \( \geq 1-O(n^{-g}) \),

\[
\sup_{b \in I_n^*} |I_n^{(3)}(b)| \leq 0(n^{-\delta}(\log n)^k) \{ \sup_{b \in I_n^*} \int_0^{x_n^{(2)}} \{1-H_n(x;b)\}^{-1} dH_n(x;b) \}
\]

\[
\leq 0(n^{-\delta}(\log n)^k) \{ \sup_{b \in I_n^*} \int_{-\infty}^{\infty} \{1-H_n(x;b)\}^{-1} dH_n(x;b) \}
\]

\[
= 0(n^{-\delta}(\log n)^k) \{ \sup_{b \in I_n^*} (n+1)^{-1} \prod_{i=1}^{n+1} (1 - \frac{i}{n+1})^{-1} \}
\]

\[
= 0(n^{-\delta}(\log n)^k) \{1 + \frac{1}{n} + \ldots + \frac{1}{n} \} = O(n^{-\delta}(\log n)^{k+1}).
\]

(3.51)

It follows similarly that for large \( n \), with probability \( \geq 1-O(n^{-g}) \), \( \sup_{b \in I_n^*} |I_n^{(3)}(b)| = O(n^{-\delta}(\log n)^{k+1}) \). Again

\[
|I_n^{(4)}(b)| \leq \{(n+1) \max_{1 \leq i \leq n} |c_{ni}^*| \} \{ \int_{x_n^{(2)}}^{\infty} |J[H_n(x;b)]| dH_n(x;b) \}
\]

\[
+ \int_{x_n^{(2)}}^{\infty} |J[H_n(x;0)]| dH_n(x;b) \} = I_n^{(4)}(b)_1 + I_n^{(4)}(b)_2, \text{ say.}
\]

(3.52)

On using (2.9), (2.10), (3.41), (3.47) and (3.48), we obtain that for large \( n \), with probability \( \geq 1-O(n^{-g}) \),
\[
\sup_{b \in I_{n}^{*}} |I_{n}^{(4)}(b)| \leq 0(n^{-\frac{k}{2}}) \sum_{b \in I_{n}^{*}} \left\{ \sup_{x \in I_{n}^{*}} \int_{x}^{(2)} \{-\log u(1-u)\} du \right\}
\]

\[
\leq 0(n^{-\frac{k}{2}}) \sum_{b \in I_{n}^{*}} \int_{1-0(n^{-\frac{k-\delta}{2}})}^{\infty} \{-\log(1-u)\} du
\]

\[
= 0(n^{-\frac{k}{2}}) \cdot 0(n^{-\frac{k-\delta}{2}}(\log n)^{k+1}) = 0(n^{-\delta}(\log n)^{k+1}). \quad (3.53)
\]

Also, \(I_{n}^{(4)}(b) = 0(n^{-\frac{k}{2}}) \int_{x_{n}}^{\infty} |J[H_{n}(x;0)]| dH_{n}(x;0) + 0(n^{-\frac{k}{2}}) \int_{x_{n}}^{\infty} |J[H_{n}(x;0)]| d[H_{n}(x;0) - H_{n}(x;0)] \). By (3.53), the first term is \(0(n^{-\delta}(\log n)^{k+1})\); with probability \(\geq 1-0(n^{-\delta})\) (note that it does not depend on \(b \in I_{n}^{*}\), while integrating by parts and using the same trick as in the earlier cases, it follows that the second term is also \(0(n^{-\delta}(\log n)^{k+1})\) for all \(b \in I_{n}^{*}\), with probability \(\geq 1-0(n^{-\delta})\). Thus, as \(n \to \infty\),

\[
\sup_{b \in I_{n}^{*}} I_{n}^{(4)}(b) = 0(n^{-\delta}(\log n)^{k+1}) \text{ with probability } \geq 1-0(n^{-\delta}). \quad (3.54)
\]

Similarly, it follows that for large \(n\), with probability \(\geq 1-0(n^{-\delta})\),

\[
\sup_{b \in I_{n}^{*}} |I_{n}^{(1)}(b)| = 0(n^{-\delta}(\log n)^{k+1}). \quad (3.55)
\]

In (3.39), we now write \(I_{n1}(b) = I_{n1}^{(1)}(b) + I_{n1}^{(2)}(b)\), where

\[
I_{n1}^{(1)}(b) = \int_{-\infty}^{\infty} J[F(x)] d[S_{n}(x;0) - S_{n}(x;0)], \quad (3.56)
\]

\[
I_{n1}^{(2)}(b) = \int_{-\infty}^{\infty} [J[H_{n}(x;0)] - J[F(x)]] d[S_{n}(x;0) - S_{n}(x;0)]. \quad (3.57)
\]
By (2.10), (3.2) and (3.57), we have

\[
\sup_{b \in \mathcal{I}^*} I_{\mathcal{I}_n^*}^{(2)}(b) = \sup_{b \in \mathcal{I}^*} \sup_{i=1}^{n_1} c_i \{ J[H_n(X_i;0)] - J[F(X_i)] \} \{ u(-X_i) - u(b c_i - X_i) \} \\
\leq \{ \max_{1 \leq i \leq n_1} |c_i| \} \{ \sup_{x \in \mathcal{I}_n^*} |J[H_n(C_n^{-1/k}(\log n)^k;0)] - J[F(C_n^{-1/k}(\log n)^k)]| \} (n+1)^\ast \\
\times \{ |H_n(C_n^{-1/k}(\log n)^k;0) - H_n(C_n^{-1/k}(\log n)^k;0)| \},
\]

(3.58)

where \( C(0) \) is finite. On using the well-known result that

\[
P(\sup_{-\infty < x < \infty} \frac{1}{\sqrt{n}} |H_n(x;0) - F(x)| > C_1 \sqrt{\log n} \leq C_2 n^{-\gamma},
\]

(3.59)

where \( C_1, C_2 \) are finite (positive) constants, and that in the neighbourhood of \( 0 \) (where \( 0 < F(0) < 1 \), \( |J'(F(x))| \) is bounded [by (2.9)], and further that

\[
F(C_n^{-1/k}(\log n)^k) - F(-C_n^{-1/k}(\log n)^k) = O(n^{-1/k}(\log n)^k),
\]

we obtain from (3.58) that

\[
\sup_{b \in \mathcal{I}^*} I_{\mathcal{I}_n^*}^{(2)}(b) \leq [O(n^{-1/k})][O(n^{-1/k}(\log n)^k)][O(n)][O(n^{-1/k}(\log n)^k)]
\]

\[
= O(n^{-1/k}(\log n)^{k+\delta}), \text{ with probability } \geq 1 - O(n^{-\gamma}).
\]

(3.60)

Further, on denoting by \( c_n^* = \max_{1 \leq i \leq n_1} |c_i| \), we have

\[
I_{\mathcal{I}_n^*}^{(1)}(b) = \left\{ \begin{array}{ll}
X(n) + \frac{b}{c_n^*} & X(n) + \frac{b}{c_n^*} \\
X(1) - \frac{b}{c_n^*} & X(1) - \frac{b}{c_n^*}
\end{array} \right.
\]

\[
\times \int [J[F(x)]d[S_n^*(x;0) - S_n^*(x;b)]]
\]

(3.61)
where \( X(1) = \min_{1 \leq i \leq n} X_i \) and \( X(n) = \max_{1 \leq i \leq n} X_i \). Hence, by (2.22) and (3.61),

\[
I^{(1)}_{n1} (b) - bB(F) = \int_{X(1) - |b|/c^*_n}^{X(n) + |b|/c^*_n} \left[ S_n^{(x;b)} - S_n^{(x;0)} - bf(x) \right] J'[F(x)] dF(x)
\]

\[
- \int_{-\infty}^{X(1) - |b|/c^*_n} bf(x) J'[F(x)] dF(x) + \int_{X(n) + |b|/c^*_n}^{\infty} bf(x) J'[F(x)] dF(x). \tag{3.62}
\]

Now, \( P(F(X(1)) \geq cn^{-\frac{1}{2}}) \geq [1 - cn^{-\frac{1}{2}}] n \leq O(n^{-s}) \), for every (fixed) \( s > 0 \), and similarly, \( P(1 - F(X(n)) \geq cn^{-\frac{1}{2}}) \leq O(n^{-s}) \). Hence, by (2.13) the last term on the right hand side of (3.62) is bounded (for all \( b \in I_n^* \)) above by \( O(n^{-\frac{1}{2}}(\log n)^{k+1}) \), with probability \( \geq 1 - O(n^{-s}) \). Also, by (2.9), (3.11) and (3.13), the first term on the right hand side of (3.62) is bounded (for all \( b \in I_n^* \)) above by

\[
\{K_1 n^{-\delta}(\log n)^k [1 + O(n^{-\frac{1}{2}})]\} K \int_{X(n) + Cn^{-\frac{1}{2}}(\log n)^k}^{X(n) + Cn^{-\frac{1}{2}}(\log n)^k} \{F(x)[1 - F(x)]\}^{-1} dF(x)
\]

\[
\leq O(n^{-\delta}(\log n)^k) \left\{-\log F(X(1) - Cn^{-\frac{1}{2}}(\log n)^k) - \log[1 - F(X(n) + Cn^{-\frac{1}{2}}(\log n)^k)]\right\}, \tag{3.63}
\]

with probability \( \geq 1 - O(n^{-s}) \), where \( c > 0 \). In the same way as in after (3.62), it can be shown that \( F(X(1) - Cn^{-\frac{1}{2}}(\log n)^k) \geq O(n^{-k}) \) with probability \( \geq 1 - O(n^{-s}) \), where \( h \) is some fixed positive number, and a similar statement holds for \( 1 - F(X(n) + Cn^{-\frac{1}{2}}(\log n)^k) \). Hence, with probability \( \geq 1 - O(n^{-s}) \), the right hand side of (3.63) is bounded above by

\[
0(n^{-\delta}(\log n)^k) \cdot O(h \log n) = O(n^{-\delta}(\log n)^{k+1}). \tag{3.64}
\]
The proof of the theorem is then completed by (3.55), (3.62), (3.63), (3.64) and the definition of \( W(\bar{X}_n, b) \), given at the beginning of this section. Q.E.D.

4. A MARTINGALE PROPERTY OF \( T(\bar{X}_n) \).

Let \( \mathcal{F}_n \) denote the \( \sigma \)-field generated by \( R_n = (R_{n1}, \ldots, R_{nn}) \); note that \( \mathcal{F}_n \) is \( \uparrow \) in \( n \). Also, let \( T^*_n = c_i \mathcal{T}_n \) and assume that \( H_0 : \beta = 0 \) holds. Then, the following theorem holds; the result is of fundamental use in proving the "uniform continuity in probability" (to be explained in section 5) of \( T(\bar{X}_n) \) with respect to \( C^{-1} \).

**Theorem 4.1.** If \( J_n(u) = EJ(U_{ni}) \), \( (i-1)/n < u < i/n, 1 \leq i \leq n \), then under \( H_0 : \beta = 0 \), \( \{T^*_n, \mathcal{F}_n\} \) forms a martingale sequence.

**Proof.** We have \( T^*_n = \sum_{i=1}^{n} (c_i - \bar{c}_n) J_n(R_{ni}/(n+1)) \), and hence,

\[
E[T^*_{n+1} | \mathcal{F}_n] = (c_{n+1} - \bar{c}_n) J_n(R_{n1}/(n+1)) + \sum_{i=1}^{n} (c_i - \bar{c}_n) J_n(R_{ni}/(n+1)) \tag{4.1}
\]

Since, \( E[J_{n+1}(R_{n1}/(n+2)) | \mathcal{F}_n] = (n+1)^{-1} \sum_{i=1}^{n+1} (1/(i+2))^1 J(u) du = \mu \), and also for any \( 1 \leq i \leq n \), \( E[J_{n+1}(R_{ni}/(n+2)) | \mathcal{F}_n] = (n+1)^{-1} R_{ni} J_{n+1}(R_{ni}/(n+2)) + \sum_{i=1}^{n+1} (1/(i+2))^1 J(u) du = \mu \), and finally, \( [i/(n+1)] J_{n+1}(i/(i+2)) + [(n-1)/(n+1)] J_{n+1}((i-1)/i) = (n-1)^{i-1} \int_{0}^{1} J(u) u^{i-1} du + (n-1)^{i-1} \int_{0}^{1} J(u) u^{i-1} du = J_n(i/(n+1)), \) for \( 1 \leq i \leq n \), we have from (4.1),

\[
E[T^*_{n+1} | \mathcal{F}_n] = (c_{n+1} - \bar{c}_n) J_n(R_{n1}/(n+1)) + \sum_{i=1}^{n} (c_i - \bar{c}_n) J_n(R_{ni}/(n+1)) + (\bar{c}_n - \bar{c}_n) J_n(R_{n1}/(n+1)) = \sum_{i=1}^{n} (c_i - \bar{c}_n) J_n(R_{ni}/(n+1)) = T^*_n, \) for all \( n \geq 1. \)  

\[
\tag{4.2}
\]
Hence, the theorem.

Remark. Note that the above martingale property does not necessarily hold when we let \( J_n(i/(n+1)) = J(i/(n+1)), 1 \leq i \leq n \) [unless \( J(u) = u: 0 < u < 1 \)]. Also, when \( \beta \neq 0 \), so that given the ranks of \( X_1, \ldots, X_n \), \( X_{n+1} \) does not have a constant probability \( 1/(n+1) \) of having the rank \( i(1 \leq i \leq n+1) \), the above martingale property does not hold, in general.

5. THE PROOF OF THE MAIN THEOREM

The proof is accomplished in several steps. First, we prove the following lemma.

Lemma 5.1. For \( \forall s > 0 \), \( \exists \) positive constants \( K_s^{(1)}, K_s^{(2)} \) and a sample size \( n_s \), such that for \( n \geq n_s \),

\[
P\left( [C_n(\hat{\beta}_{U,n} - \beta) - A \alpha/2/B(F)] > K_s^{(1)} (\log n)^2 \right) < K_s^{(2)} n^{-s},
\]

\[
P\left( [C_n(\hat{\beta}_{L,n} - \beta) + A \alpha/2/B(F)] < -K_s^{(1)} (\log n)^2 \right) < K_s^{(2)} n^{-s}.
\]

Proof. We shall only consider the proof of (5.1) as the proof of (5.2) follows on the same line. By virtue of the translation-invariance of the estimates \( \hat{\beta}_{u,n} \) and \( \hat{\beta}_{l,n} \), we may, without any loss of generality, assume that \( \beta = 0 \). Then, by (2.7), we have

\[
P\left( [C_n(\hat{\beta}_{U,n} - A \alpha/2/B(F)] > K_s^{(1)} (\log n)^2 \right) = P\left( T_n - C_n^{-1}[A \alpha/2/B(F) + K_s^{(1)} (\log n)^2]c_n \right) > T_n^{(1)} | \beta = 0\}
\]

\[
= P_{\beta = 0} \{ T_n - T_n^{(1)} > T_n^{(1)} - T_n \},
\]

\[
(5.3)
\]
where we write $\tilde{T}_n = T(x_n - \beta o_n c_n^*)$, $b_o = A_T \alpha / 2 / B(F)$ and $\hat{T}_n = T(x_n - [b_o + K_s(1) (\log n)^2] c_n^*)$.

By theorem 3.2, with probability $\geq 1 - O(n^{-8})$,

$$\hat{T}_n - \tilde{T}_n = -K_s(1) (\log n)^2 B(F) + O(n^{-\delta} (\log n)^3), \delta > 0.$$  \hfill (5.4)

Also, using the fact that $\lim_{n \to \infty} T_n^{(1)} = -A_T \alpha / 2$ and writing $S_n(x; b_o) = \sum_{i=1}^{n} c_n^* f(x - b_o c_n^*)$,

we obtain by some standard computations that

$$\int_{-\infty}^{\infty} J[F(x)] dS_n(x; b_o) = b_o B(F) + O(n^{-\frac{1}{2}}) = T_n^{(1)} + O(n^{-\frac{1}{2}}).$$  \hfill (5.5)

Hence, it suffices to show that with probability $\geq 1 - O(n^{-8})$,

$$| \int_{-\infty}^{\infty} J[F(x)] dS_n(x; b_o) - \hat{T}_n | < K_s(1) (\log n)^{3/2}.$$  \hfill (5.6)

The left hand side of (5.6) can be written as

$$| \int_{-\infty}^{\infty} \{ J[F(x)] - J[H_n(x; b_o)] \} \, dS_n(x; b_o) + \int_{-\infty}^{\infty} J[H_n(x; b_o)] \, d[S_n(x; b_o) - S_n(x; b_o)] |$$

$$\leq | \int_{-\infty}^{\infty} \{ J[F(x)] - J[H_n(x; b_o)] \} \, dS_n(x; b_o) |$$

$$+ | \int_{-\infty}^{\infty} \{ S_n(x; b_o) - S_n(x; b_o) \} J'[H_n(x; b_o)] \, dH_n(x; b_o) |$$  \hfill (5.7)

It can be shown by using theorem 2 of Hoeffding (1963) that $\sup_{x} |S_n(x; b_o) - S_n(x; b_o)| = O((\log n)^{1/2})$, with probability $\geq 1 - O(n^{-8})$. Also, $\int_{-\infty}^{\infty} |J'[H_n(x; b_o)]| \, dH_n(x; b_o) \leq (n+1)^{-1} \sum_{i=1}^{n} (i(n+1-i)/(n+1)^2)^{-1} = O(\log n)$ [by (2.9)]. Hence, the second term on the right hand side of (5.7) is $O((\log n)^{3/2})$ with probability $\geq 1 - O(n^{-8})$.

Finally, precisely on the same line as in (3.42)-(3.55), it follows that the first term on the right hand side of (5.7) is also $O((\log n)^{3/2})$, with probability $\geq 1 - O(n^{-8})$. Hence the proof is complete.
A direct consequence of theorem 3.2 and lemma 5.1 is the following.

**Lemma 5.2.** For every \( s > 0 \), there exists positive constants \( (K_s^{(1)}, K_s^{(2)}) \) and a sample size \( n_s \), such that for \( n > n_s \),

\[
P\{|C_n(\hat{\beta}_{U,n} - \hat{\beta}_{L,n}) - 2\Delta \alpha^2 / B(F)| > K_s^{(1)} n^{-\delta} (\log n)^{\frac{1}{2}} \} \leq K_s^{(2)} n^{-s}.
\]  

(5.8)

Also, we have the following lemma whose proof follows along the lines of Sen (1969, Section 3).

**Lemma 5.3.** For every real \( x \) \((-\infty, x) \),

\[
\lim_{n \to \infty} P\{C_n(\hat{\beta}_{U,n} - \beta)B(F)/A - \alpha^2 / 2 < x\} = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-\frac{1}{2}t^2)dt.
\]  

(5.9)

Finally, for the "uniform continuity in probability" (for definition, see Anscombe (1952)) of \( \{\hat{\beta}_{U,n}\} \) with respect to \( \{C^{-1}_n\} \), we have the following:

**Lemma 5.4.** For every positive \( \varepsilon \) and \( \eta \), there exists a \( \delta(>0) \), such that as \( n \to \infty \),

\[
P\{\sup_{n' - n} |C_n(\hat{\beta}_{U,n} - \hat{\beta}_{U,n'})| > \eta \} < \varepsilon,
\]  

(5.10)

and a similar statement holds for \( \{\hat{\beta}_{L,n}\} \).

**Proof.** By theorem 3.2, lemma 5.1 (where we let \( s = 1 + h \), \( h > 0 \)), and (2.17), we have, with probability \( \geq 1 - O(n^{-h}) \), for all \( |n' - n| < \delta n \),

\[
C_n(\hat{\beta}_{U,n} - \hat{\beta}_{U,n'}) = C_n(\hat{\beta}_{U,n} - \beta)[C_n/C_n] - C_n(\hat{\beta}_{U,n} - \beta)
\]

\[
= -[C_n/C_n][1/B(F)][T(X_n - \hat{\beta}_{U,n}C_n) - T(X_n - \beta C_n) + o(1)]
\]

\[
+ [1/B(F)][T(X_n - \hat{\beta}_{U,n}C_n) - T(X_n - \beta C_n) + o(1)].
\]
\[-[A/B(F)]\{\alpha/2[1-C_n/C_n]_n+(C_n/C_n')_n[T(X_n,-\beta_{e_n})_n-T(X_n,-\beta_{c_n})_n]/B(F)\]
\[-T(X_n,-\beta_{e_n})_n[1-C_n/C_n]_n]/B(F)+o(1). \tag{5.11}\]

By (2.12), \(\sup_{n'\to n}<\delta_n|1-C_n/C_n',<\delta', where \(\delta'(>0) depends on \(\delta\), and can be made arbitrarily small when \(\delta\) is made so. Also, the asymptotic normality of 
\(T(X_n,-\beta_{e_n})_n/A\) (with zero mean and unit variance) implies that \(|T(X_n,-\beta_{e_n})_n| is bounded, in probability, as \(n\to\infty\). Hence, it suffices to show that as \(n\to\infty\)

\[P\{\sup_{n'\to n}<\delta_n|T(X_n,-\beta_{e_n})_n-T(X_n,-\beta_{c_n})_n|>\eta\} < \varepsilon. \tag{5.12}\]

Since \(|T(X_n,-\beta_{e_n})_n-T(X_n,-\beta_{c_n})_n| \leq |(C_n,-C_n')_n/C_n,|T_n| + C_n^{-1}|T^*_n,-T_n^*|\), (where \(T^*_n\) is defined in section 4), it seems sufficient to show that

\[P_{\beta=0}\{\sup_{n'\to n}<\delta_n|T^*_n,-T^*_n|>\eta C_n\} < \varepsilon, \tag{5.13}\]

and in view of the martingale property of \(T^*_n\), see theorem 4.1, from the Kolmogorov inequality (cf. Loève (1963, p. 386)) for martingales, we get,

\[P_{\beta=0}\{\sup_{n'\to n}<\delta_n|T^*_n,-T^*_n|>\eta C_n\} \leq (\eta C_n)^{-2}E(T^2_n+[\delta_n]-T^2_n-[\delta_n])\]

\[= (\eta C_n)^{-2}[C^2_n+[\delta_n]A^2_n+[\delta_n]-c^2_n-[\delta_n]A^2_n-[\delta_n]]\]

The rest of the proof follows from (2.12) and the fact that \(A^2_n = A^2+o(1)\), for large \(n\).

We now return to the proof of theorem 2.1. By (2.17) lemma 5.2 and the definition of \(N(d)\), it follows that for all \(d>0\), \(N(d)\) is finite a.s. and is \(ـd\) in \(d\). We also note that
\[ \text{E}[N(d)] = \sum_{n=0}^{\infty} n \text{P}(N(d)=n) = \sum_{n>0} \text{P}(N(d)>n). \]  

(5.14)

Hence, in order to show that \( \text{E}[N(d)]<\infty \), it suffices to show that for large \( n \),

\[ \text{P}(N(d)>n) = O(n^{-1-\eta}), \text{ where } \eta>0. \]  

(5.15)

By definition, the left hand side of (5.15) is equal to

\[ \text{P}(\hat{\beta}_U,n^{\frac{k}{2}}-\hat{\beta}_L,n^{\frac{k}{2}}>2d \text{ for all } k<n) \leq \text{P}(\hat{\beta}_U,n^{\frac{k}{2}}-\hat{\beta}_L,n^{\frac{k}{2}}>2d) \]

\[ = \text{P}(C_n(\hat{\beta}_U,n^{\frac{k}{2}}-\hat{\beta}_L,n^{\frac{k}{2}})>2dC_n) = O(n^{-s}), \]  

(5.16)

(where we can let \( s>1 \)), as by (5.8),

\[ C_n(\hat{\beta}_U,n^{\frac{k}{2}}-\hat{\beta}_L,n^{\frac{k}{2}}) = 2A_{n/2}/B(F) + O(n^{-s}(\log n)^2), \]

where as by (2.11),

\[ 2dC_n \geq 2dC_n^k. \]  

Hence, for every \( d>0 \), \( \text{E}[N(d)]<\infty \). Using the fact that \( \hat{\beta}_U,n^{\frac{k}{2}}-\hat{\beta}_L,n^{\frac{k}{2}}>0 \) a.s. for each \( n \), we have \( \lim_{d \to 0} N(d) = \infty \) and finally by the monotone convergence theorem \( \lim_{d \to 0} \text{E}[N(d)] = \infty \).

Now (2.19) follows directly from (5.8), (2.12) and (2.22); (2.20) also follows from theorem 1 of Anscombe (1952) after using our lemmas 5.3 and 5.4.

To prove (2.21), we let for each \( d(>0) \),

\[ n_1(d) = [Q^{-1}(\upsilon(d))(1+(-1)^{d-1}e) + 1+(-1)^{d-1}] \]

(5.17)

\([x]\) being the largest integer contained in \( x \), and write,

\[ \text{E}[N(d)]/Q^{-1}(\upsilon(d)) = [Q^{-1}(\upsilon(d))]^{-1}\{\Sigma_1 + \Sigma_2 + \Sigma_3 \text{ nP}(N(d)=n)\}, \]

where the summations \( \Sigma_1, \Sigma_2 \) and \( \Sigma_3 \) extend over \( n<n_1(d), n_1(d)<n<n_2(d) \) and \( n>n_2(d) \), respectively. Since \( Q(\upsilon) \) is \( \uparrow \), \( \lim_{d \to 0} \upsilon(d) = \infty \) and (2.19) holds, for every \( \varepsilon>0 \), there exists a \( d_\varepsilon(>0) \), such that for all \( 0<d<d_\varepsilon \),

\[ \text{P}(n_1(d)<N(d)<n_2(d)) \geq \text{P}(\frac{N(d)}{Q^{-1}(\upsilon(d))}-1|<\varepsilon) \geq 1-\eta, \]

where \( \eta(>0) \) is a (pre-assigned) small number. Hence, for \( d<d_\varepsilon \),
\[
\{Q^{-1}(v(d))\}^{-1}\Sigma_1 nP\{N(d)=n\} \leq (1-\varepsilon)P\{N(d)<n_1(d)\} \leq \eta(1-\varepsilon).
\] (5.18)

Also, by the same technique as in (5.16), we have

\[
\{Q^{-1}(v(d))\}^{-1}\Sigma_3 nP\{N(d)=n\} = \{Q^{-1}(v(d))\}^{-1}n_2(d)P\{N(d) = n_2(d)\}
\]
\[+ \{Q^{-1}(v(d))\}^{-1}\Sigma_3 P\{N(d)>n\}\]
\[= \{Q^{-1}(v(d))\}^{-1}\{O(n_2(d))O[(n_2(d))^{-s}] + O[(n_2(d))^{-s+1}]\}\]
\[= O((n_2(d))^{-s+1}) + 0 \text{ as } n\rightarrow\infty,\] (5.19)

as in (5.8) we can let \(s>1\) and \(Q^{-1}(v(d))\rightarrow\infty\) as \(d\rightarrow0\). Finally,

\[
|\{Q^{-1}(v(d))\}^{-1}\Sigma_2 nP\{N(d)=n\}-1| \leq \varepsilon\Sigma_2 P\{N(d)=n\} + \eta \leq \varepsilon + \eta.
\] (5.20)

Thus, the proof of (2.21) follows from (5.18)-(5.20).

Remark: Using (5.12), the classical Wald-Wolfowitz-Noether-Hájek theorem [cf. Hájek and Šidák (1967, p. 160)] on the asymptotic normality of \(T_n(\beta)\) for fixed but large \(n\), and (2.19), it readily follows from theorem 1 (p. 601) of Anscombe (1952) that \(T_{N(d)}(\beta)\) has asymptotically (as \(d\rightarrow0\)) a normal distribution with zero mean and variance \(A^2\) defined in (2.23). Also, on using theorem 3.2, it follows that for every \(b\in I_{N(d)}\), \(T_{N(d)}(\beta+b) + bB(F)\) has asymptotically the same normal distribution. These results give simple proofs of the asymptotic normality of regression rank statistics based on random sample sizes both in the null (\(b=0\)) and the non-null (\(b\neq0\)) situations. The results trivially extend to any stopping variable \(N_r\), indexed by a sequence \(\{r\}\) and a sequence of positive integers \(\{n_r\}\), such that \(n_r\rightarrow\infty\) and \(n_r/n_r \rightarrow 1\) as \(r\rightarrow\infty\).
6. ASYMPTOTIC RELATIVE EFFICIENCY

For any two procedures A and B for determining (sequentially) bounded
length confidence intervals for \( \beta \) (with the same bound 2d), let \( P_A(d) \) and \( P_B(d) \)
be the coverage probabilities and \( N_A(d) \) and \( N_B(d) \) the stopping variables
corresponding to the respective procedures. Then, the ARE of the procedure A
with respect to the procedure B is given by

\[
e_{A,B} = \lim_{d \to 0} \left[ \frac{E_B(d)}{E_A(d)} \right],
\]

(6.1)

provided \( \lim_{d \to 0} P_A(d) = \lim_{d \to 0} P_B(d) \) and either of the limit exists.

Gleser (1965) considered the case when \( n^{-1} C \to C > 0 \) as \( n \to \infty \). However, one can
easily extend his results when (2.10) and (2.11) hold. Thus, if \( G \) denotes his
procedure, theorem 2.1 also holds with the change that \( \nu(d) \) has to be replaced
by \( \nu_G(d) = \sigma^2 \tau^{2 \nu} / d^2 \), \( \sigma^2 \) being the variance of the distribution of \( F(x) \) in (2.1).
Writing \( R \) for the proposed procedure and using (2.21), we have now,

\[
e_{R,G} = \lim_{d \to 0} \left\{ Q^{-1}(\nu_G(d))/Q^{-1}(\nu(d)) \right\}.
\]

(6.2)

By definition, \( \nu_G(d)/\nu(d) = \sigma^2 B^2(F)/A^2 \) is independent of \( d \). Write \( e = s(e^*) \).
Then, \( e^* = s^{-1}(e) \) is monotonic in \( e \) with \( e^* = 1 \) when \( e = 1 \). Write \( \phi_d = \psi_G(d)/\psi(d), \)
where \( \psi_G(d) = Q^{-1}(\nu_G(d)), \psi(d) = Q^{-1}(\nu(d)) \). Both \( \psi_G(d) \) and \( \psi(d) \) tend to \( \infty \)
as \( d \to 0 \). Thus,

\[
s(e^*) = e = \nu_G(d)/\nu(d) = \lim_{d \to 0} \left\{ \nu_G(d)/\nu(d) \right\}
= \lim_{d \to 0} \left\{ Q(\phi_d \psi^*(d))/Q(\psi^*(d)) \right\}.
\]

(6.3)

Using (2.12) and proving by contradiction, we have,
\[ \lim_{d \to 0} \phi(d) = e^* = s^{-1}(e) = s^{-1}(\sigma^2B^2(F)/A^2) \] (6.4)

The expression \( \sigma^2B^2(F)/A^2 \) is the Pitman-efficiency of a general rank order test with respect to Student's t test. In the particular case, when \( J(u) = \phi^{-1}(u) \), \( \phi \) being the distribution function (d.f.) of a standard normal variable, (normal score) it is well-known (cf. Chernoff and Savage (1958)) that \( \sigma^2B^2(F)/A^2 \geq 1 \) for all d.f. \( F \) with a density \( f \) and a finite second moment, equality being attained when and only when \( F \) is normal \((0, \sigma^2)\) d.f. From monotonicity, it follows now from (4.4) that in this case \( e^* \geq 1 \), equality being attained if and only if \( F \) is normal. Also, when \( J(u)=u \) (Wilcoxon score), \( e^* = s^{-1}(12\sigma^2(\int_{-\infty}^{\infty} f^2(x)dx)^2) \geq s^{-1} \) (0.864) (cf. Hájek and Šidák (1967, p. 280)). In the case of equispaced regression line, \( t_i = a+(i-1), \ i=1,2,\ldots, \ C_n^2 = n(n^2-1)/12, \ e^* = \{12\sigma^2(\int_{-\infty}^{\infty} f^2(x)dx)^2\}^{1/3}. \) For normal \( F \), this reduces to \((0.955)^{1/3} \approx .985, \) while the infimum is given by \((0.864)^{1/3} \approx .953. \)

In the special case when \( c_i \) is either 0 or 1, \( \beta \) is the difference in the location parameters of the two distributions \( F(x-\alpha') \) and \( F(x-\alpha'-\beta) \). This is the classical two-sample problem. If at the \( n^{th} \) stage, \( m_n \) of the \( c_i \) are 1 and rest zero, \( C_n^2 = m_n (n-m_n)/4 \leq n/4, \) for all \( n \geq 1. \) Looking at the definition of \( C_n^2 \), (2.19) and (2.21), we may observe that an optimum choice of \( m_n \) is \([\frac{n}{2}]\), the integral part of \( \frac{n}{2} \). Thus, among all designs for obtaining a bounded length confidence interval for \( \beta \), in this problem, and optimum design (which minimizes \( EN(d) \) for small \( d \)) consists in taking every alternate observation for the two distributions. Here, \( n^{-1}C_n^2 \) and the ARE reduces to \( \sigma^2B^2(F)/A^2 \) various bounds for which have been discussed earlier.
REFERENCES


