The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 1950–1979

ASYMPTOTIC BEHAVIOR OF WILCOXON TYPE CONFIDENCE REGIONS IN MULTIPLE LINEAR REGRESSION[1]

BY HIRA LAL KOUL[2]

University of California, Berkeley

0. Summary. In the multiple linear regression model a class of confidence regions based on rank statistics is constructed. The asymptotic behavior of the center of gravity of a region corresponding to the Wilcoxon type rank statistic is considered.

Section 1 consists of assumptions, introduction and notation. Section 2 consists of a monotonicity lemma. A uniform continuity theorem for Wilcoxon type linear rank statistics is proved in Section 3. This and the monotonicity lemma are used to show that the regions are bounded for large sample size, with large probability. Section 4 gives the asymptotic normality of the above mentioned quantity and defines a consistent estimator of $(\int f^2(x))$. Finally, the Appendix contains some results on relative compactness of some stochastic processes which are used in Section 3.

1. Notation, assumptions and introduction. Suppose we are observing a double sequence $\{Y_{in}, 1 \leqq i \leqq n\}$ $n \geqq 1$ of independent random variables such that

$$\text{Prob } [Y_{in} \leqq y] = F(y - \mathbf{X}_{in}\boldsymbol{\theta}) \qquad i = 1, \cdots, n, \tag{1.0}$$

where

$$\mathbf{X}_{in} = (x_{in}(1), \cdots, x_{in}(p))$$

is the vector of non-random regression scores,

$$\boldsymbol{\theta}' = (\theta_1, \cdots, \theta_p)$$

is the parameter vector of interest.

Our problem here is to use the rank statistics to construct a certain confidence region and study the asymptotic behavior of its center of gravity. For this we will need the following assumptions in the sequel.

In (1.0) cdf F is not known but is assumed to be a member of the class of distributions

$$\begin{aligned} \mathfrak{F}_0 = \{F;\ & \text{(i) } F \text{ is absolutely continuous,} \\ & \text{(ii) } \frac{dF}{dx}(x) = f(x) \text{ is absolutely continuous and bounded,} \\ & \text{(iii) } f(x) = f(-x), \\ & \text{(iv) } \int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)}\right)^2 f(x)dx < \infty\}. \end{aligned} \tag{1.1}$$

Received 2 February 1968.

[1] This research was partially supported by the Office of Naval Research Contract Nønr N000014-A-0114-0004.

[2] Presently at Michigan State University.

In the sequel we shall write $x_{i\nu}$ for $x_{in}(\nu)$ $\nu = 1, \cdots, p$ $i = 1, \cdots, n$. This does not mean that $x_{i\nu}$ are independent of n, but rather dependence on n is suppressed for the sake of convenience.

Let

$$X_n = ((x_{i\nu})) \quad \nu = 1, \cdots, p; \ i = 1, \cdots, n.$$

We need to assume that

$$\lim \left(\max_{1 \leqq i \leqq n} x_{i\nu}^2\right) / \left(\sum_{i=1}^n x_{i\nu}^{\frac{1}{2}}\right) = 0 \tag{1.2}$$

for all $\nu = 1, 2, \cdots, p$.

$$\lim n^{-1} X_n' X_n = \tilde{\Sigma} \tag{1.3}$$

exists and is positive definite matrix, where the limit of a matrix is the matrix formed by the limit of each entry in the matrix. All limits in this paper unless otherwise stated are taken as $n \to \infty$.

For any two sets A, B, $A \triangle B$ will stand for their symmetric difference. By $\mathfrak{L}_\theta(X)$ we will mean the law of a r.v. X when $\boldsymbol{\theta}$ is the true parameter point. For a set A, A^c denotes the complement of A.

For a known distribution function $G \in \mathfrak{F}_0$, define the score function

$$\psi(u) = -g'(G^{-1}(u + 1)/2)/g(G^{-1}(u + 1)/2), \quad 0 \leqq u \leqq 1. \tag{1.4}$$

Also define

$$\psi_n(u) = \psi(i/(n + 1)) \text{ for } i - 1/n < u \leqq i/n, \ 1 \leqq i \leqq n. \tag{1.5}$$

Assume

$$\lim \int_0^1 [\psi_n(u) - \psi(u)]^2 \, du = 0. \tag{1.6}$$

Let R_{in} be the rank of $|Y_{in}|$ among $\{Y_{in}, 1 \leqq i \leqq n\}$.

Let $y_n \equiv (Y_{in}, \cdots, Y_{nn}) \equiv Y_n \equiv y$ denote the sample.

Define

$$S_{n\nu}(y_n) = n^{-1} \sum_{i=1}^n x_{i\nu} \psi_n(R_{in}/(n + 1)) \operatorname{sgn}(Y_{in}) \qquad \nu = 1, \cdots, p, \tag{1.7}$$

where

$$\begin{aligned} \operatorname{sgn}(x) &= 1 \qquad \text{if } x \geqq 0, \\ &= -1 \qquad \text{if } x < 0. \end{aligned}$$

Let

$$\begin{aligned} \mathbf{S}_n' \equiv \mathbf{S}_n'(y_n) &= (S_{n1}(y_n), \cdots, S_{np}(y_n)), \\ M_n(y_n) &= n\mathbf{S}_n'(y_n)\Sigma_n^{-1}\mathbf{S}_n(y_n), \end{aligned} \tag{1.8}$$

where

$$\Sigma_n = n^{-1} X_n' X_n \cdot \int_0^1 \psi^2(u) \, du. \tag{1.9}$$

To test the hypothesis $H_0 : \boldsymbol{\theta} = \mathbf{0}$ in the model (1.0), the test statistic M_n defined above is used. The test is to reject the H_0 if $M_n(y_n) \geqq k_\alpha$ where k_α is so determined as to yield the level α test. For a small sample, k_α can be determined by an ordinary permutation technique, because under H_0 all $\{Y_{in}, 1 \leqq i \leqq n\}$ are i.i.d. and consequently the test based on M_n is distribution free.

Adichie, in [1], considered the class of test statistics defined above by (1.4), (1.5), (1.7) and (1.8) for the case $p = 2$, $x_{i1} = 1$, $1 \leqq i \leqq n$ and arbitrary x_{i2}, $1 \leqq i \leqq n$. For this case he gave the asymptotic distribution of M_n under both the hypothesis H_0 and contiguous alternatives. Generalizing his result to general p in a straightforward fashion, we shall state the following lemma without any proof.

LEMMA 1.1. *Under the conditions* (1.2) *and* (1.3) *on the regression scores and* (1.6) *on the score function* ψ *and for* $F \varepsilon \mathfrak{F}_0$

$$\mathfrak{L}(M_n(Y) \mid H_0) \rightarrow \chi_p^2.$$

where χ_p^2 *is a chi square variable with p degrees of freedom.*

Thus for large n, k_α can be computed from chi-square tables.

Clearly, to test the hypothesis $H(\boldsymbol{\theta}_0): \boldsymbol{\theta} = \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}_0$ is some known vector, one would use the test statistic $M_n(y_n - \mathbf{X}_n\boldsymbol{\theta}_0)$: the statistic M_n based on the ranks of $\{|Y_{in} - \mathbf{X}_{in}\boldsymbol{\theta}_0|, 1 \leqq i \leqq n\}$. From now on we will invariably write $M_n(y_n, \boldsymbol{\theta})$, $M_n(\boldsymbol{\theta}, y_n)$, $M_n(\boldsymbol{\theta})$ or $M_n(y_n - \mathbf{X}_n\boldsymbol{\theta})$ for the test statistic M_n based on the ranks of $|Y_{in} - \mathbf{X}_{in}\boldsymbol{\theta}|$. The same will apply to $\{S_{n\nu}, 1 \leqq \nu \leqq p\}$.

Define the confidence region

$$(1.10) \qquad \mathfrak{R}_n(y_n) = \{\boldsymbol{\theta}: M_n(\boldsymbol{\theta}, y_n) \leqq k_\alpha\}.$$

This confidence region has confidence coefficient $1 - \alpha$. We shall assume that α is fixed and

$$(1.11) \qquad 0 < \alpha < 1,$$

so that for all n, k_α stays bounded away from 0 and ∞. This follows from Lemma 1.1.

Note that, since for each sample point y_n the sequence of sets $\{\mathfrak{R}_n(y_n), n \geqq 1\}$ is Lebesgue measurable because for each y_n the sequence $\{M_n(y_n, \boldsymbol{\theta})\ n \geqq 1\}$ is Lebesgue measurable in $\boldsymbol{\theta}$, it makes sense to talk about the Lebesgue measure of $\mathfrak{R}_n(y_n)$. Let λ denote p-dimensional Lebesgue measure on p-dimensional Euclidean space. Define

$$(1.12) \qquad \hat{\boldsymbol{\theta}}_n(y_n) = (\lambda[\mathfrak{R}_n(y_n)])^{-1} \int_{\mathfrak{R}_n(y_n)} \mathbf{t}\lambda(d\mathbf{t})$$

where the integral is supposed to be the vector valued integral.

(1.10), (1.12) define a class of $\hat{\boldsymbol{\theta}}_n$'s, one corresponding to each ψ. If the score function ψ is monotone, and if $p = 1$, $\alpha = 0$ and $x_{i1} = 1$ for $1 \leqq i \leqq n$, then it is easily seen that the corresponding $\hat{\boldsymbol{\theta}}_n$ is the Hodges and Lehmann [8] estimator of the location parameter in one sample problem. The same applies to the case when $p = 1$, $x_{i1} = 1$ for $1 \leqq i \leqq n_1$, and $x_{i1} = 0$ for $n_1 + 1 \leqq i \leqq n$. Adichie in [2] gave a class of estimators for the case $p = 1$ using Hájek [7] type statistics—the linear rank statistics based on the ranks of the original observations and not of their absolute values. If we take $p = 1$, $\alpha = 0$ and assume ψ to be monotone, as is done by Adichie [2], and use S_{n1} defined by (1.7), then it can easily be seen, in view of the following Lemma 2.2 that $\hat{\boldsymbol{\theta}}_n$ in this case reduces to an Adichie type estimator for θ_1. Therefore, this motivates defining $\hat{\boldsymbol{\theta}}_n$ as a vector estimator of $\boldsymbol{\theta}$. But it is not clear whether for small n $\hat{\boldsymbol{\theta}}_n$ is well defined, i.e., whether it has or has not a coordinate of indeterminate form.

However, we have been able to show here that if $\psi(u) = u$, $0 \leqq u \leqq 1$, then under some regularity conditions $n^{\frac{1}{2}}\mathcal{R}_n$ is bounded for large n with large probability and $\lambda(n^{\frac{1}{2}}\mathcal{R}_n(y_n))$ has finite and positive limit in probability. Consequently, $n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n$ is well defined for large n with large probability. The main idea of the proof is to approximate $M_n(y_n, \boldsymbol{\theta}n^{-\frac{1}{2}})$, uniformly in $\boldsymbol{\theta}$ in a bounded set, in probability by a smoother statistic and to use this, plus Lemma 2.2, in conjunction with a form of Cauchy-Schwartz inequality to show that for large n, the region $n^{\frac{1}{2}}\mathcal{R}_n(y_n)$ covers parameter points too far from the true parameter point with very small probability. This method automatically yields the asymptotic normality of $n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n$ and the convergence of the Lebesgue measure of $n^{\frac{1}{2}}\mathcal{R}_n(y_n)$ in probability to a finite and positive limit.

It is believed that this asymptotic result will remain valid for a class of score functions ψ which are monotone, square integrable and havefirst two integrable derivatives.

For the rest of the paper, therefore, we shall assume $\psi(u) = u$, $0 \leqq u \leqq 1$. Regions corresponding to this will be called Wilcoxon Type regions.

2. Invariance of $\mathcal{R}_n$ and a monotonicity lemma.

LEMMA 2.1. *For any vector* $\mathbf{b}$ *and all* n

$$\mathcal{R}_n(y_n + X_n\mathbf{b}) = \mathcal{R}_n(y_n) + \mathbf{b} \tag{2.1}$$

$$\mathcal{R}_n(-y_n) = -\mathcal{R}_n(y_n) \tag{2.2}$$

with probability 1. *Consequently, if* $\hat{\boldsymbol{\theta}}_n(y_n)$ *is well defined, and if the underlying distribution* F *is symmetric around* 0, *then* $\hat{\boldsymbol{\theta}}_n(y_n)$ *is symmetrically distributed about the true parameter point* $\boldsymbol{\theta}$.

PROOF. The proof of (2.1) is immediate, and that of (2.2) follows by noting that $M_n(-y_n - X_n\boldsymbol{\theta}) = M_n(y_n + X_n\boldsymbol{\theta})$ with probability 1. This terminates the proof.

To conclude this section we shall prove a monotonicity lemma for a linear rank statistic S when $p = 1$. This lemma will be used in the following section to derive the claimed asymptotic theory.

Let θ be a real number. Let $\{x_i, 1 \leqq i \leqq n\}$ be any real numbers. Define

$$S(\theta) = \sum_{i=1}^{n} x_i R_i \operatorname{sgn}(Y_i + \theta x_i) \tag{2.3}$$

where R_i is the rank of $|Y_i + \theta x_i|$ among $\{|Y_i + \theta x_i|, 1 \leqq i \leqq n\}$.

LEMMA 2.2. $S(\theta') \geqq S(\theta'')$ *for all those* θ', θ'' $(\theta' \geqq \theta'')$ *for which* S *is well defined and for all* $\{Y_i, 1 \leqq i \leqq n\}$ *such that* $Y_i \neq Y_j$ *for* $i \neq j$. (Well defined here means 'no ties'.)

PROOF. Let

$$Y_i' = Y_i + \theta' x_i, \qquad Y_i'' = Y_i + \theta'' x_i \qquad i = 1, \cdots, n.$$

Let R_i' and R_i'' denote the ranks of $|Y_i'|$ and $|Y_i''|$ respectively. The proof will be split into several cases.

CASE 1. Suppose there is only one change between the sets of ranks $\{R_i'\}$ and $\{R_i''\}$. Then

$$R_i' = R_j'', \qquad R_j' = R_i'' \quad \text{and} \quad R_h' = R_h'' \quad h \neq i, j. \tag{2.4}$$

Also assume that

$$\operatorname{sgn}(Y_i') = \operatorname{sgn}(Y_i''), \qquad i = 1, \cdots, n. \tag{2.5}$$

Under (2.4) and (2.5)

$$\begin{aligned} S(\theta') - S(\theta'') &= x_i\{R_i' - R_i''\} \operatorname{sgn}(Y_i') + x_j\{R_j' - R_j''\} \operatorname{sgn}(Y_i'') \\ &= (R_j'' - R_j')\{x_i \operatorname{sgn}(Y_i') - x_j \operatorname{sgn} Y_j'\}. \end{aligned} \tag{2.6}$$

Without loss of generality, suppose

$$R_j'' > R_i'', \quad \text{implying} \quad R_i' > R_j' \quad (\text{by } (2.4)). \tag{2.7}$$

Therefore,

$$R_j'' - R_j' > 0. \tag{2.8}$$

It remains to show that the second factor of (2.6) is nonnegative. (2.7) and (2.4) $\Rightarrow |Y_j''| > |Y_i''|$ and $|Y_i'| > |Y_j'| \Leftrightarrow Y_j'' \operatorname{sgn}(Y_j') > Y_i'' \operatorname{sgn}(Y_i'')$ and $Y_i' \operatorname{sgn}(Y_i') > Y_j' \operatorname{sgn}(Y_i') \Leftrightarrow (\theta' - \theta'')(x_i \operatorname{sgn}(Y_i') - x_j \operatorname{sgn}(Y_j')) \geqq 0$, which implies

$$x_i \operatorname{sgn}(Y_i') - x_j \operatorname{sgn}(Y_j') \geqq 0, \tag{2.9}$$

because $\theta' - \theta'' \geqq 0$.
Combining (2.8) and (2.9), we get $S(\theta') - S(\theta'') \geqq 0$ under (2.7), and similarly if $R_j'' < R_i''$. This concludes Case 1.

CASE 2. Here assume one sign change between sets $\{Y_i'\}$ and $\{Y_i''\}$ and no rank change. Say

$$\operatorname{sgn}(Y_i') = \operatorname{sgn}(Y_i''), \quad i > 1, \quad \text{and} \quad \operatorname{sgn}(Y_i') \neq \operatorname{sgn}(Y_1''). \tag{2.10}$$

Then, since $R_i' = R_i''$ $1 \leqq i \leqq n$,

$$S(\theta') - S(\theta'') = x_1 R_1' \{\operatorname{sgn}(Y_1') - \operatorname{sgn}(Y_1'')\}. \tag{2.11}$$

Without loss of generality suppose $Y_1' \geqq 0$ and $Y_1' < 0$; then $(\theta' - \theta'')x_1 \geqq 0 \Rightarrow x_1 \geqq 0$ for $(\theta' - \theta'') \geqq 0$. Therefore $x_1 R_1' \{\operatorname{sgn}(Y_1') - \operatorname{sgn}(Y_1'')\} = 2x_1 R_1' \geqq 0$. Hence $S(\theta') - S(\theta'') \geqq 0$.

In general the interval from θ'' to θ' may be split into steps $\theta'' = \theta_0 < \theta_1 < \theta_n = \theta'$ such that from θ_i to θ_{i+1} there is either (i) exactly one rank change and no sign change or (ii) exactly one sign change and no rank change.

To see this, suppose on the contrary, no matter how small δ is, there is always an interval $[\theta, \theta + \delta)$ containing at least one rank exchange and at least one sign exchange. Then there must exist $\theta < \theta^* < \theta + \delta$ and i, j, k $(i \neq j)$ such that $|Y_i - \theta^* x_i| = |Y_j - \theta^* x_j|$, $Y_k = \theta^* x_n \Rightarrow |Y_i - Y_k(x_i/x_k)| = |Y_j - Y_k(x_j/x_k)|$ which can happen with probability zero. This concludes the proof.

It might be remarked that the above lemma remains true for a class of statistics $S = \sum x_i \psi(R_i/(n + 1)) \operatorname{sgn}(Y_i + \theta x_i)$ where ψ is a nondecreasing score function.

3. A uniform continuity theorem for linear rank statistic. In this section the statistic $M_n(y_n, \boldsymbol{\theta})$ will be approximated in probability by another quadratic form, uniformly in $\boldsymbol{\theta}$ in a bounded set. Also it will be shown that the parameter points too far from the true parameter value do not give us much trouble. These results will be used in proving the claimed result in the following section.

Because of (2.1) of Lemma 2.1, there is no harm in assuming that the true parameter point is $\mathbf{0}$. Let P_n denote the probability distribution generated by $\{Y_{in}, 1 \leqq i \leqq n\}$, which are i.i.d. F, that is, $\mathbf{0}$ is the true parameter point.

Since M_n is a quadratic combination of $S_{n\nu}$, $\nu = 1, \cdots, p$, we will first approximate $S_{n\nu}$, $\nu = 1, \cdots, p$. Define for a p-vector $\mathbf{t}$, and real x, the functions

$$(3.1) \quad \mu_{n\nu}(\mathbf{t}, x) = (n^{-1}) \sum_{i=1}^{n} x_{i\nu} I(Y_{in} - \mathbf{X}_{in}\mathbf{t} \leqq x) \operatorname{sgn}(Y_{in} - \mathbf{X}_{in}\mathbf{t}), \qquad \nu = 1, \cdots, p.$$

$$(3.2) \qquad \bar{\mu}_{n\nu}(\mathbf{t}, x) = \mathcal{E}\mu_n(\mathbf{t}, x).$$

Here $I(A)$ is the indicator of the set A, and sgn (x) is defined by (1.7).

Also define

$$(3.3) \qquad H_n(\mathbf{t}, |x|) = (n + 1)^{-1} \sum_{i=1}^{n} I(|Y_{in} - \mathbf{X}_{in}\mathbf{t}| \leqq |x|),$$

$$(3.4) \qquad \bar{H}_n(\mathbf{t}, |x|) = \mathcal{E}H_n(\mathbf{t}, |x|).$$

$\mathcal{E}$ in the above definitions denotes the expectation under P_n.

By $X_{in}t$ we shall mean $\mathbf{X}_{in}\mathbf{t} = \sum_{\nu=1}^{p} x_{i\nu} t_\nu$. Define for $\mathbf{t}$,

$$(3.5) \qquad \|\mathbf{t}\| = \sum_{\nu=1}^{p} |t_\nu|.$$

In view of (3.1), (3.3), we can write

$$(3.6) \qquad S_{n\nu}(\mathbf{t}) = \int_{-\infty}^{\infty} H_n(\mathbf{t}, |x|)\, d\mu_{n\nu}(\mathbf{t}, x), \qquad \nu = 1, \cdots, p,$$

where x is the integrating variable. This representation is of the Chernoff-Savage type where the function $\mu_{n\nu}(\mathbf{t}, \cdot)$ gives mass $n^{-1} x_{i\nu} \operatorname{sgn}(Y_{in} - X_{in}t)$ to the point $Y_{in} - X_{in}t$ and $H_n(\mathbf{t}, \cdot)$ gives mass $(n + 1)^{-1}$ to the point $|Y_{in} - X_{in}t|$.

Define

$$(3.7) \qquad A_{n\nu}(\mathbf{t}) = \int_{-\infty}^{\infty} \bar{H}_n(\mathbf{t}, |x|)\, d\bar{\mu}_{n\nu}(\mathbf{t}, x), \qquad \nu = 1, \cdots, p.$$

For an a, $0 < a < \infty$, define $V_n(a) = \{\mathbf{t};\ \|\mathbf{t}\| \leqq an^{-\frac{1}{2}}\}$ and $V(a) = \{\mathbf{t}, \|\mathbf{t}\| \leqq a\}$. Also define $\mathcal{F}_1 = \{F; F$ is absolutely continuous and $F' = dF = f$ is absolutely continuous and bounded, vanishing at most at a finite number of intervals.$\}$

THEOREM 3.1. *Let* $\{Y_{in}, 1 \leqq i \leqq n\}$ $n \geqq 1$ *be the sequences of independent* rv *as in* 1.1. *Let* $F \varepsilon \mathfrak{F}_1$. *Assume that* $\{x_{i\nu}, 1 \leqq i \leqq n, 1 \leqq \nu \leqq p\}, n \geqq 1$, *satisfy* (1.2) *and* (1.3). *Then for every* $\epsilon > 0$ *there exists* n_ϵ, *which may depend on* F *and* $\{x_{i\nu}\}$, *such that* $n > n_\epsilon$ *yields*

$$(3.8)\quad P_n[\sup_{t \varepsilon V_n(a)} n^{\frac{1}{2}} |\{S_{n\nu}(\mathbf{t}) - A_{n\nu}(\mathbf{t})\} - (S_{n\nu}(\mathbf{0}) - A_{n\nu}(\mathbf{0})\}| \geqq \epsilon] \leqq \epsilon$$

for any $0 < a < \infty$ *and for any* $\nu = 1, \cdots, p$ *fixed.*

PROOF. Since the proof for each $\nu = 1, \cdots, p$ is the same, we fix ν and will drop it from the suffixes of statistics. Thus by μ_n we will mean $\mu_{n\nu}$ for some ν and so on.

We have the following decomposition of $S_n(\mathbf{t}) - A_n(\mathbf{t})$.

$$\begin{aligned} S_n(\mathbf{t}) - A_n(\mathbf{t}) &= \int_{-\infty}^{\infty} H_n(\mathbf{t}, |x|)\, d\mu_n(\mathbf{t}, x) - \int_{-\infty}^{\infty} \bar{H}_n(\mathbf{t}, |x|)\, d\bar{\mu}_n(t, x) \\ &= \int_{-\infty}^{\infty} \bar{H}_n(\mathbf{t}, |x|)\, d\{\mu_n(\mathbf{t}, x) - \bar{\mu}_n(\mathbf{t}, x)\} + \int_{-\infty}^{\infty} \{H_n(\mathbf{t}, |x|) \\ &\qquad - \bar{H}_n(\mathbf{t}, |x|)\}\, d\bar{\mu}_n(\mathbf{t}, x) + \int_{-\infty}^{\infty} \{H_n(\mathbf{t}, |x|) \\ &\qquad - \bar{H}_n(\mathbf{t}, |x|)\}\, d\{\mu_n(\mathbf{t}, x) - \bar{\mu}_n(\mathbf{t}, x)\}. \\ &= B_{n1}(\mathbf{t}) + B_{n2}(\mathbf{t}) + R_n(\mathbf{t}) \quad \text{(say)}. \end{aligned} \tag{3.9}$$

We will show that for every $\epsilon > 0$

$$(3.10)\quad P_n[\sup_{t \varepsilon V_n(a)} n^{\frac{1}{2}} |B_{ni}(\mathbf{t}) - B_{ni}(\mathbf{0})| \geqq \epsilon] \leqq \epsilon \quad \text{for} \quad n \geqq n_\epsilon, \quad i = 1, 2,$$

and that

$$(3.11)\quad P_n[\sup_{t \varepsilon V_n(a)} n^{\frac{1}{2}} |R_n(\mathbf{t})| \geqq \epsilon] \leqq \epsilon.$$

Consider the difference

$$\begin{aligned} n^{\frac{1}{2}}\{B_{n2}(\mathbf{t}) - B_{n2}(\mathbf{0})\} &= n^{\frac{1}{2}}[\int_{-\infty}^{\infty} \{H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)\}\, d\bar{\mu}_n(\mathbf{t}, x) \\ &\qquad - \int_{-\infty}^{\infty} \{H_n(\mathbf{0}, |x|) - \bar{H}_n(\mathbf{0}, |x|)\}\, d\bar{\mu}_n(\mathbf{0}, x)]. \end{aligned}$$

Let

$$(3.12)\quad Z_n(\mathbf{t}, |x|) = n^{\frac{1}{2}}\{H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)\}.$$

It is easy to see, with $\operatorname{sgn}(x) = I(x \geqq 0) - I(x \leqq 0)$, that

$$(3.13)\quad \bar{\mu}_n(\mathbf{t}, x) = n^{-1} \sum_{i=1}^n x_{i\nu}\{F(x + X_{in}t) \operatorname{sgn}(x) - 2F(X_{in}t) I(x \geqq 0)\}.$$

(3.13) and (3.12) yield

$$\begin{aligned} n^{\frac{1}{2}}\{B_{n2}(\mathbf{t}) - B_{n2}(\mathbf{0})\} &= n^{-1} \sum_{i=1}^n x_{i\nu}[\int_{-\infty}^{\infty} z_n(\mathbf{t}, |x|) \operatorname{sgn}(x)\, dF(x + X_{in}t) \\ &\qquad - \int_{-\infty}^{\infty} Z_n(\mathbf{0}, |x|) \operatorname{sgn}(x)\, dF(x)] \\ &= n^{-1} \sum_{i=1}^n x_{i\nu}[\int_{-\infty}^{\infty} \{Z_n(\mathbf{t}, |x - X_{in}t|) \operatorname{sgn}(x - X_{in}t) \\ &\qquad - Z_n(\mathbf{0}, |x|) \operatorname{sgn}(x)\}\, dF(x)]. \end{aligned}$$

And therefore

$$\sup_{t \in V_n(a)} n^{\frac{1}{2}} |B_{n2}(\mathbf{t}) - B_{n2}(\mathbf{0})|$$

$$\leqq n^{-1} \textstyle\sum_{i=1}^{n} |x_{i\nu}| [\sup_x \sup_{t \in V_n(a)} \{ |Z_n(\mathbf{t}, |x - X_{in}t|) - Z_n(\mathbf{0}, |x - X_{in}t|)| + |Z_n(\mathbf{0}, |x - X_{in}t|) - Z_n(\mathbf{0}, |x|)| + |Z_n(\mathbf{0}, |x - X_{in}t|)| \\ 2P_n[0 \leqq Y \leqq X_{in}t \quad \text{or} \quad X_{in}t \leqq Y \leqq 0] \}]. \tag{3.14}$$

Observe that

$$\sup_{-\infty \leqq x \leqq +\infty} \sup_{t \in V_n(a)} \max_{1 \leqq i \leqq n} |Z_n(\mathbf{t}, |x - X_{in}t|) - Z_n(\mathbf{0}, |x - X_{in}t|)| \\ \leqq \sup_{-\infty \leqq x \leqq +\infty} \sup_{t \in V_n(a)} |Z_n(\mathbf{t}, |x|) - Z_n(\mathbf{0}, |x|)| \tag{3.15}$$

which tends to zero in P_n probability by (52) of Theorem A5.

Also, if $\delta_n = a \max_{1 \leqq i \leqq n} \|X_{in}\| n^{-\frac{1}{2}}$,

$$\max_{1 \leqq i \leqq n} \sup_{-\infty \leqq x \leqq +\infty} \sup_{t \in V_n(a)} |Z_n(\mathbf{0}, |x - X_{in}t|) - Z_n(0, |x|)| \\ \leqq \sup_{\| |y| - |x| \| \leqq \delta_n} |Z_n(\mathbf{0}, |y|) - Z_n(\mathbf{0}, |x|)| \tag{3.16}$$

which tends to zero in P_n probability because by assumption $\delta_n \to 0$ and because $Z_n(0, |x|)$ is a standard empirical cumulative process with continuous Gaussian limit (see [3], e.g.).

Furthermore,

$$\max_{1 \leqq i \leqq n} \sup_{-\infty \leqq x \leqq +\infty} \sup_{t \in V_n(u)} |Z_n(\mathbf{0}, |x - X_{in}t|)| \\ \leqq \sup_{-\infty \leqq x \leqq +\infty} |Z_n(\mathbf{0}, |x|)|, \tag{3.17}$$

and by (53) of the Appendix the right hand side is bounded in P_n probability for large n. Finally since F has bounded density, and since

$$\sup_{t \in V_n(a)} \max_{1 \leqq i \leqq n} |X_{in}t| \to 0,$$

we have

$$\sup_{t \in V_n(a)} \max_{1 \leqq i \leqq n} P_n[0 \leqq Y_{in} \leqq X_{in}t \quad \text{or} \quad X_{in}t \leqq Y_{in} \leqq 0] \to 0. \tag{3.18}$$

Combining (3.18), (3.17), (3.16) and (3.15) with the condition (1.3), we have (3.10) for B_{n2} term.

Next, consider the remainder term

$$n^{\frac{1}{2}} R_n(\mathbf{t}) = n^{\frac{1}{2}} \textstyle\int_{-\infty}^{\infty} \{H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)\} \, d\{\mu_n(\mathbf{t}, x) - \bar{\mu}_n(\mathbf{t}, x)\},$$

$$= \textstyle\int_{-\infty}^{\infty} \{H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)\} \, dL_n(\mathbf{t}, x),$$

where

$$L_n(\mathbf{t}, x) = n^{\frac{1}{2}} \{\mu_n(\mathbf{t}, x) - \bar{\mu}_n(\mathbf{t}, x)\}. \tag{3.19}$$

We shall integrate the $R_n(\mathbf{t})$ term by parts. To see that the integration by parts is justified here, see Apostol [0].

Note that by definition of $H_n(\mathbf{t}, |x|)$ and $L_n(\mathbf{t}, x)$

$$(3.20)\quad P_n[\sup_{t\varepsilon V_n(a)} \lim_{x\to\pm\infty} [H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)] = 0] = 1 \quad \text{for all} \quad n,$$
$$P_n[\sup_{t\varepsilon V_n(a)} |L_n(\mathbf{t}, -\infty)| = 0] = 1 \quad \text{for all} \quad n,$$

and $\sup_{t\varepsilon V_n(a)} |L_n(\mathbf{t}, +\infty)|$ has a limiting distribution. Therefore, integrating by parts the $n^{\frac{1}{2}}R_n(\mathbf{t})$, for every $\epsilon > 0$, one gets

$$(3.21)\quad \sup_{t\varepsilon V_n(a)} n^{\frac{1}{2}}|R_n(\mathbf{t})| = \sup_{t\varepsilon V_n(a)} |\textstyle\int_{-\infty}^{\infty} L_n(\mathbf{t}, x)\, d\{H_n(\mathbf{t}, |x|) - \bar{H}_n(\mathbf{t}, |x|)\}|$$

with P_n—probability at least $1 - \epsilon$ for $n \geqq n_\epsilon$. Let

$$(3.22)\quad R_{n1}(\mathbf{t}) = \textstyle\int_0^\infty L_n(\mathbf{t}, x)\, d\{H_n(\mathbf{t}, x) - \bar{H}_n(\mathbf{t}, x)\},$$

$$(3.23)\quad R_{n2}(\mathbf{t}) = \textstyle\int_{-\infty}^0 L_n(\mathbf{t}, x)\, d\{H_n(\mathbf{t}, -x) - \bar{H}_n(\mathbf{t}, -x)\}.$$

Define, for $0 \leqq y \leqq 1$,

$$(3.24)\quad H_n^{-1}(\mathbf{t}, y) = \inf\{x > 0; H_n(\mathbf{t}, x) \geqq y\},$$
$$\bar{H}_n^{-1}(\mathbf{t}, y) = \inf\{x > 0; \bar{H}_n(\mathbf{t}, x) \geqq y\}.$$

By an ordinary substitution process one sees that

$$(3.25)\quad R_{n1}(\mathbf{t}) = \textstyle\int_0^1 \{L_n(\mathbf{t}, H_n^{-1}(\mathbf{t}, y)) - L_n(\mathbf{t}, \bar{H}_n^{-1}(\mathbf{t}, y))\}\, dy,$$

$$(3.26)\quad \sup_{t\varepsilon V_n(a)} |R_{n1}(\mathbf{t})|$$
$$\leqq \sup_{0\leqq y\leqq 1} \sup_{t\varepsilon V_n(a)} |L_n(\mathbf{t}, H_n^{-1}(\mathbf{t}, y)) - L_n(\mathbf{t}, \bar{H}_n^{-1}(\mathbf{t}, y))|.$$

However, by Theorem A6 of the Appendix, the right hand side of (3.26) tends to zero in P_n—probability as $n \to \infty$. Therefore for every $\epsilon > 0$

$$(3.27)\quad P_n[\sup_{t\varepsilon V_n(a)} |R_{n1}(\mathbf{t})| \geqq \epsilon] < \epsilon \quad \text{for} \quad n \geqq n_\epsilon.$$

Similarly, it can be shown that for every $\epsilon > 0$

$$(3.28)\quad P_n[\sup_{t\varepsilon V_n(a)} |R_{n2}(\mathbf{t})| \geqq \epsilon] \leqq \epsilon \quad \text{for} \quad n \geqq n_\epsilon.$$

Moreover, note that

$$(3.29)\quad \sup_{t\varepsilon V_n(a)} n^{\frac{1}{2}}|R_n(\mathbf{t})| \leqq \sup_{t\varepsilon V_n(a)} |R_{n1}(\mathbf{t})| + \sup_{t\varepsilon V_n(a)} |R_{n2}(\mathbf{t})|.$$

Combining (3.28), (3.27) and the above remarks, (3.11) follows. Finally, the proof for the B_{n1} term uses integration by parts, Theorem A4 and a slight variation of Theorem A6. We terminate the proof.

Our next objective is to approximate the function $A_{n\nu}(\mathbf{t})$ by a linear function of $\mathbf{t}$.

LEMMA 3.1. *Define*

$$(3.30)\quad \mathfrak{F} = \mathfrak{F}_0 \cap \{F; \textstyle\int_\infty^\infty \sup_{\infty<s<\infty} |f'(x - s)|\, dF(x) < \infty\}.$$

Assume $F \varepsilon \mathfrak{F}$. Also assume that the regression scores $\{x_{i\nu}\}$ satisfy conditions (1.2), (1.3). *Then*

$$(3.31)\quad \lim \sup_{t\varepsilon V_n(a)} n^{\frac{1}{2}}|A_{n\nu}(\mathbf{t}) - A_{n\nu}(\mathbf{0}) + 2\mathbf{t}'\dot{\mathbf{A}}_{n\nu}(\mathbf{0})| = 0$$

for any $0 < a < \infty$ *and all* $\nu = 1, \cdots, p$.
Here

$$\dot{\mathbf{A}}_{n\nu}(\mathbf{0}) = n^{-1} \sum_{i=1}^{n} x_{i\nu}\mathbf{X}'_{in} \cdot \int_{-\infty}^{\infty} f^2(x)\, dx. \tag{3.32}$$

PROOF. Recall that

$$A_{n\nu}(\mathbf{t}) = \int_{-\infty}^{\infty} \bar{H}_n(\mathbf{t}, |x|)\, d\bar{\mu}_{n\nu}(\mathbf{t}, x).$$

It is easy to see from the definition that

$$\bar{H}_n(\mathbf{t}, |x|) = (n+1)^{-1} \sum_{i=1}^{n} \{F(|x| + X_{in}t) - F(-|x| + X_{in}t)\} \tag{3.33}$$

After some simple algebraic manipulations, using the definitions of $\bar{H}_n$ given by (3.33) and $\bar{\mu}_{n\nu}$ given by (3.13), one can see that

$$\begin{aligned} A_{n\nu}(\mathbf{t}) = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{j\nu} \int_{-\infty}^{\infty} \{F(x + (X_{in} - X_{jn})t) \\ - F(-x + (X_{in} + X_{jn})t)\}\, dF(x). \end{aligned} \tag{3.34}$$

Taylor's expansion around $\mathbf{t} = \mathbf{0}$ of the integrand above yields

$$\begin{aligned} &\{F(x + (\mathbf{X}_{in} - \mathbf{X}_{jn})\mathbf{t}) - F(-x + (\mathbf{X}_{in} + \mathbf{X}_{jn})t)\} \\ &\quad = \{F(x) - F(-x)\} - 2\mathbf{X}_{jn}\mathbf{t}f(x) + R_{nij}(\mathbf{t}, x) - R'_{nij}(\mathbf{t}, x) \end{aligned} \tag{3.35}$$

where

$$\begin{aligned} R_{nij}(\mathbf{t}, x) &= (X_{in} - X_{jn})tt'(X_{in} - X_{jn})' \cdot f'(x + \xi(X_{in} - X_{jn})'), \\ R'_{nij}(\mathbf{t}, x) &= (X_{in} + X_{jn})tt'(X_{in} + X_{jn})'f'(x + \xi(X_{in} + X_{jn})'), \end{aligned} \tag{3.36}$$

where $\|\boldsymbol{\xi}\| \leqq \|\mathbf{t}\|$ and $\boldsymbol{\xi}$ may depend on $\mathbf{t}$, n, x. Let

$$e_n(\mathbf{t}) = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{j\nu} \int_{-\infty}^{\infty} R_{nij}(\mathbf{t}, x)\, dF(x). \tag{3.37}$$

Then

$$\begin{aligned} \sup_{t \varepsilon V_n(a)} n^{\frac{1}{2}}|e_n(\mathbf{t})| \leqq a^2 n^{\frac{1}{2}} \cdot n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_{j\nu}| \cdot \|X_{in} - X_{jn}\|^2 \\ \cdot n^{-1} \times \sup_{\|\xi\| \leqq an^{-\frac{1}{2}}} \int_{-\infty}^{\infty} |f'(x + (X_{in} - X_{jn})\xi)|\, dF(x). \end{aligned} \tag{3.38}$$

By (3.30),

$$\begin{aligned} &\lim \sup_{\|\xi\| \leqq an^{-\frac{1}{2}}} \int_{-\infty}^{\infty} |f'(x + (X_{in} - X_{jn})\boldsymbol{\xi})|\, dF(x) \\ &\quad \leqq \lim \int_{-\infty}^{\infty} \sup_{\|\xi\| \leqq an^{-\frac{1}{2}}} |f'(x + (X_{in} - X_{jn})\boldsymbol{\xi})|\, dF(x) < \infty \end{aligned} \tag{3.39}$$

uniformly in $1 \leqq i, j \leqq n$.
Also, by (1.2), (1.3)

$$\begin{aligned} &\lim \sum_{i=1}^{n} \sum_{j=1}^{n} |x_{j\nu}| \cdot \|X_{in} - X_{jn}\|^2 \cdot n^{-3} \cdot n^{\frac{1}{2}} \\ &\quad \leqq \lim\, (\max_{1 \leqq i,j \leqq n} \|X_{in} - X_{jn}\| \cdot n^{-\frac{1}{2}}) \\ &\qquad \cdot \lim\, (n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} |x_{j\nu}| \cdot \|X_{in} - X_{jn}\|) = 0. \end{aligned} \tag{3.40}$$

Combining (3.40), (3.39) and (3.38), we have

$$\lim \sup_{t \varepsilon V_n(a)} n^{\frac{1}{2}}|e_n(\mathbf{t})| = 0. \tag{3.41}$$

Similarly, if

$$e_n'(\mathbf{t}) = (n^{-2}) \sum_{i=1}^n \sum_{j=1}^n x_{j\nu} \int_{-\infty}^{\infty} R'_{nij}(\mathbf{t}, x)\, dF(x),$$

then

$$\sup_{\mathbf{t}\in V_n(a)} n^{\frac{1}{2}}|e_n'(\mathbf{t})| \to 0 \quad \text{as} \quad n \to \infty. \tag{3.42}$$

Recalling the definition of $A_{n\nu}(\mathbf{0})$, and combining this with (3.41), (3.42), (3.35) and (3.34), it follows that

$$\begin{aligned}\lim \sup_{\mathbf{t}\in V_n(a)} n^{\frac{1}{2}}|A_{n\nu}(\mathbf{t}) - A_{n\nu}(\mathbf{0}) + 2\mathbf{t}'\dot{\mathbf{A}}_{n\nu}(\mathbf{0})| \\ = \lim \sup_{\mathbf{t}\in V_n(a)} n^{\frac{1}{2}}|e_n(\mathbf{t}) - e_n'(\mathbf{t})| = 0.\end{aligned}$$

This concludes the proof.

Next define

$$\hat{S}_{n\nu}(\mathbf{0}) = n^{-1} \sum_{i=1}^n x_{i\nu}\{2F(|Y_{in}|) - 1\} \operatorname{sgn}(Y_{in}) \tag{3.43}$$

$$T_{n\nu}(\mathbf{t}) = \hat{S}_{n\nu}(\mathbf{0}) - 2\dot{\mathbf{A}}'_{n\nu}(\mathbf{0})\mathbf{t}. \qquad \nu = 1, \cdots, p. \tag{3.44}$$

Adichie [1] under the assumptions (1.2) and (1.3), has shown that for every $\epsilon > 0$

$$\lim P_n[n^{\frac{1}{2}}|\hat{S}_{n\nu}(\mathbf{0}) - S_{n\nu}(\mathbf{0})| \geqq \epsilon] = 0. \tag{3.45}$$

Combining (3.4) with Lemma 3.1 and Theorem 3.1 we have

COROLLARY 3.1. *Under the assumptions of Theorem* 3.1 *and Lemma* 3.1, *for every* $\epsilon > 0$ *there exists* n_ϵ *such that* $n \geqq n_\epsilon$ *yields*

$$P_n[\sup_{\mathbf{t}\in V_n(a)} n^{\frac{1}{2}}|S_{n\nu}(\mathbf{t}) - T_{n\nu}(\mathbf{t})| \geqq \epsilon] \leqq \epsilon$$

for all $\nu = 1, \cdots, n$ *and any* $0 < a < \infty$.

THEOREM 3.2. *Under the conditions of Theorem* 3.1 *and Lemma* 3.1, *for every* $\epsilon > 0$ *there exists an* n_ϵ *such that* $n \geqq n_\epsilon \Rightarrow$

$$P_n[\sup_{\mathbf{t}\in V_n(a)} |M_n(Y_n, \mathbf{t}) - T_n(Y_n, \mathbf{t})| \geqq \epsilon] \leqq \epsilon \tag{3.46}$$

for any $0 < a < \infty$, *with*

$$\begin{aligned} T_n(Y_n, \mathbf{t}) &= n\mathbf{V}_n'(Y_n, \mathbf{t})\Sigma_n^{-1}\mathbf{V}_n(Y_n, \mathbf{t}), \\ \mathbf{V}_n'(Y_n, \mathbf{t}) &= (T_{n1}(Y_n, \mathbf{t}), \cdots, T_{np}(Y_n, \mathbf{t})). \end{aligned} \tag{3.47}$$

$T_{n\nu}(Y_n, \mathbf{t})$ *being defined by* (3.46).

PROOF. Note that

$$\begin{aligned}|M_n(\mathbf{t}) - T_n(\mathbf{t})| \leqq n^{\frac{1}{2}}\|\mathbf{S}_n(\mathbf{t}) - \mathbf{V}_n(\mathbf{t})\|\{\|\Sigma_n^{-1}n^{\frac{1}{2}}\mathbf{S}_n(Y_n, \mathbf{t})\| \\ + \|n^{\frac{1}{2}}\Sigma_n^{-1}\mathbf{V}_n(Y_n, \mathbf{t})\|\}\end{aligned} \tag{3.48}$$

where $\|\cdot\|$ is defined by (3.5).

Since the vector $n^{\frac{1}{2}}\hat{\mathbf{S}}_n'(\mathbf{0}) = n^{\frac{1}{2}}(S_{n1}(\mathbf{0}), \cdots, S_{np}(\mathbf{0}))$ is known to have limiting normal distribution (see [1]) with covariance matrix Σ_n, therefore, by

Theorem 3.1, Lemma 3.1 and Corollary 3.1 it follows that both the quantities $\sup_{t\varepsilon V_n(a)} \|n^{\frac{1}{2}}\Sigma_n^{-1}\mathbf{S}_n(Y_n, \mathbf{t})\|$ and $\sup_{t\varepsilon V_n(a)} \|n^{\frac{1}{2}}\Sigma_n^{-1}\mathbf{V}_n(Y_n, \mathbf{t})\|$ have a limiting distribution. Moreover, again, by Theorem 3.1

$$P_n[\sup_{t\varepsilon V_n(a)} n^{\frac{1}{2}}\|\mathbf{S}_n(\mathbf{t}) - V_n(\mathbf{t})\| \geqq \epsilon] \leqq \epsilon \quad \text{for every} \quad \epsilon > 0 \quad \text{and} \quad n \geqq n_\epsilon .$$

The proof is concluded.

After some algebraic manipulations one can see that if

$$\hat{\mathbf{S}}_n'(\mathbf{0}) = (\hat{S}_{n1}(\mathbf{0}), \cdots, \hat{S}_{np}(\mathbf{0})) \quad \text{and} \tag{3.49}$$

$$b(f) = 6(\textstyle\int_{-\infty}^{\infty} f^2(x)\, dx), \tag{3.50}$$

then

$$T_n(\mathbf{t}) = n(\hat{\mathbf{S}}_n(\mathbf{0}) - b(f)\Sigma_n\mathbf{t})\Sigma_n^{-1}(\hat{\mathbf{S}}_n(\mathbf{0}) - b(f)\Sigma_n\mathbf{t}). \tag{3.51}$$

Define

$$D_n(y_n) = \{\boldsymbol{\theta} : T_n(y_n, \boldsymbol{\theta}) \leqq k_\alpha\}. \tag{3.52}$$

One can easily see that (see Cramér [4])

$$\lambda(n^{\frac{1}{2}}D_n(y_n)) = (\pi k_\alpha)^{p/2}/\{\Gamma((p/2) + 1)[b(f)]^p \cdot |\Sigma_n|^{\frac{1}{2}}\} \tag{3.53}$$

where $\Gamma(x)$ is standard Gamma function, $|(\Sigma_n)|$ = determinant of Σ_n. Since for $F \varepsilon \mathfrak{F}_0$, $0 < b(f) < \infty$ and since $k_\alpha > 0$ and $|\Sigma_n|$ has a limit and is positive by (1.3), it follows that

$$0 < \lim \lambda(n^{\frac{1}{2}}D_n(y_n)) < \infty \tag{3.54}$$

with probability 1. Furthermore observe that the center of gravity of D_n is given by

$$\tilde{\boldsymbol{\theta}}_n(y_n) = [b(f)]^{-1}\hat{\mathbf{S}}_n'(\mathbf{0})\Sigma_n^{-1}. \tag{3.55}$$

This will be used later to derive asymptotic normality of $\hat{\boldsymbol{\theta}}_n$.

We are now ready to turn to the problem of boundedness of the regions $n^{\frac{1}{2}}\mathfrak{R}_n(y_n)$. We begin with the following definitions. Define for $a - \infty \leqq \gamma \leqq +\infty$ and a vector $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta}\| = 1$, the functions

$$g_n(\gamma, \boldsymbol{\theta}) = n^{\frac{1}{2}}[\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0}) - b\gamma\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta}]/(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{\frac{1}{2}} \tag{3.56}$$

$$h_n(\gamma, \boldsymbol{\theta}) = n^{\frac{1}{2}}[\boldsymbol{\theta}'\mathbf{S}_n(\gamma\boldsymbol{\theta})]/(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{\frac{1}{2}} \tag{3.57}$$

where $b \equiv b(f)$ defined in (3.50).

Assume that the conditions of Theorem 3.1, Lemma 3.1 hold.

LEMMA 3.2. *For every $\epsilon > 0$ and a given $d > k_\alpha$ there exists an a, $0 < a < \infty$, large enough such that $n \geqq n_\epsilon$ implies*

$$P_n[\inf_{\|\theta\|=1} \inf_{|\gamma|=an^{-\frac{1}{2}}} |g_n(\gamma, \boldsymbol{\theta})| \geqq d] \geqq 1 - \epsilon, \tag{3.58}$$

$$P_n[\inf_{\|\theta\|=1} \inf_{|\gamma|=an^{-\frac{1}{2}}} |h_n(\gamma, \boldsymbol{\theta})| \geqq d] \geqq 1 - \epsilon, \tag{3.59}$$

$$P_n[\inf_{|\gamma|<an^{-\frac{1}{2}}} |g_n(\gamma, \boldsymbol{\theta})| = 0 \quad \textit{for every} \quad \|\boldsymbol{\theta}\| = 1] \geqq 1 - \epsilon \tag{3.60}$$

and

$$(3.61) \qquad P_n[\inf_{|\gamma|<an^{-\frac{1}{2}}} |h_n(\gamma, \boldsymbol{\theta})| \leqq \epsilon \quad \textit{for every} \quad \|\boldsymbol{\theta}\| = 1] \geqq 1 - \epsilon.$$

PROOF. We begin by proving (3.58) and (3.60). By definition $\hat{\mathbf{S}}_n'(0) = (\hat{S}_{n1}(\mathbf{0}) \cdots \hat{S}_{np}(\mathbf{0}))$, where $\hat{S}_{n\nu}(\mathbf{0})$, $\nu = 1, \cdots, p$ are defined by (3.43) which are the sums of independent random variables, and by the Lindeberg central limit theorem, $\lim \mathcal{L}(n^{\frac{1}{2}}\Sigma_n^{-1}\hat{\mathbf{S}}_n(0)) = N(0, 1)$. Therefore $n^{\frac{1}{2}}\|\hat{\mathbf{S}}_n(\mathbf{0})\|$ has a limiting distribution.

Also note that since $\lim \Sigma_n$ is a positive definitive matrix, there are two constant $\eta > 0$ and $K < \infty$ such that

$$(3.62) \qquad 0 < \eta \leqq (\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{\frac{1}{2}} \leqq K < \infty$$

uniformly for all $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta}\| = 1$. Therefore, for any $\|\boldsymbol{\theta}\| = 1$ we have

$$(3.63) \qquad b^{-1}n^{\frac{1}{2}}[\boldsymbol{\theta}'\hat{\mathbf{S}}(\mathbf{0})][\boldsymbol{\theta}'\Sigma_n^{-1}\boldsymbol{\theta}]^{-\frac{1}{2}} \leqq b^{-1}\eta^{-1}n^{\frac{1}{2}}\|\hat{\mathbf{S}}_n(\mathbf{0})\|.$$

Hence, combining this with the above remark, we have that for every $\epsilon > 0$ there is a constant $c > 0$ large enough such that for $n \geqq n_\epsilon$

$$(3.64) \qquad P_n[\inf_{\|\theta\|=1} b^{-1}(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{-\frac{1}{2}}n^{\frac{1}{2}}|\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0})| \leqq c] \geqq 1 - \epsilon.$$

Choose

$$(3.65) \qquad a \geqq d + c.$$

Then using this choice of a and the definition of g_n and (3.63) one can easily see that

$$\begin{aligned} P_n[\inf_{\|\theta\|=1} \inf_{|\gamma|=an^{-\frac{1}{2}}} |g_n(\gamma, \boldsymbol{\theta})| \geqq b] & \\ &\geqq P_n[\inf_{\|\theta\|=1} |\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{-\frac{1}{2}}n^{\frac{1}{2}}|\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0})|b^{-1} - a(\boldsymbol{\theta}\Sigma_n\boldsymbol{\theta})| \geqq d] \\ &\geqq P_n[\inf_{\|\theta\|=1} b'(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{-\frac{1}{2}}n^{\frac{1}{2}}|\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0})| \leqq cK/\eta] \\ &\geqq 1 - \epsilon. \end{aligned}$$

We shall next establish (3.60). By the definition of $g_n(\gamma, \boldsymbol{\theta})$ it is obvious that over-all infimum of g_n is zero and is attained at the point

$$(3.66) \qquad \gamma_i(\boldsymbol{\theta}) = [\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0})][\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta}]^{-1} d^{-1}.$$

(3.60) will, therefore, follow if we show that there exists an a, $0 < a < \infty$ large enough such that

$$(3.67) \qquad P_n[|\gamma_1(\boldsymbol{\theta})| < an^{-\frac{1}{2}} \quad \text{for every} \quad \|\boldsymbol{\theta}\| = 1] \geqq 1 - \epsilon.$$

However, in view of (3.63) and the fact that $b^{-1}n^{\frac{1}{2}}\|\hat{\mathbf{S}}_n(\mathbf{0})\|$ has a limiting distribution, it follows that for every $\epsilon > 0$ and $n \geqq n_\epsilon$ there exists a c such that

$$\begin{aligned} \epsilon &\geqq P_n[\sup_{\|\theta\|=1} [n^{\frac{1}{2}}|\boldsymbol{\theta}'\hat{\mathbf{S}}_n(\mathbf{0})|][\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta}]^{-1}b^{-1} > c] \\ (3.68) \qquad &= P_n[n^{\frac{1}{2}}|\gamma_1(\boldsymbol{\theta})| > c\eta \quad \text{for some} \quad \|\boldsymbol{\theta}\| = 1] \\ &= 1 - P_n[n^{\frac{1}{2}}|\gamma_1(\boldsymbol{\theta})| \leqq c\eta \quad \text{for every} \quad \|\boldsymbol{\theta}\| = 1]. \end{aligned}$$

Therefore for a given by (3.65), (3.67) holds in view of (3.68). This completes the proof for (3.60).

To see that the other two claims are true we make the following observations. From the definitions of g_n and h_n and Theorem 3.1, Lemma 3.1, one can see that for every $\epsilon > 0$ there exists n_ϵ such that $n \geqq n_\epsilon$ yields

$$(3.69)\quad P_n[|\inf_{|\gamma|<an^{-\frac{1}{2}}} |g_n(\gamma, \boldsymbol{\theta})| - \inf_{|\gamma|<an^{-\frac{1}{2}}} |h_n(\gamma, \boldsymbol{\theta})|\,| \leqq \epsilon \quad \text{for every} \quad \|\boldsymbol{\theta}\| = 1] \geqq 1 - \epsilon$$

and

$$(3.70)\quad P_n[|\inf_{|\gamma|=an^{-\frac{1}{2}}, \|\theta\|=1} |g_n(\gamma, \boldsymbol{\theta})| - \inf_{|\gamma|=an^{-\frac{1}{2}}, \|\theta\|=1} |h_n(\gamma, \boldsymbol{\theta})|\,| \leqq \epsilon] \geqq 1 - \epsilon$$

for any $0 < a < \infty$.

Now if we choose $0 < a < \infty$ so as to satisfy (3.58) and (3.60) then by (3.69), (3.61) is satisfied and by (3.70), (3.59) is satisfied for the same a. The proof is terminated.

LEMMA 3.3. *For every $\epsilon > 0$ and a given $d > k_\alpha$ there exists an a such that*

$$(3.71)\quad P_n[\inf_{|\gamma| \geqq an^{-\frac{1}{2}}, \|\theta\|=1} |g_n(\gamma, \boldsymbol{\theta})| \geqq d] \geqq 1 - \epsilon,$$

$$(3.72)\quad P_n[\inf_{|\gamma| \geqq an^{-\frac{1}{2}}, \|\theta\|=1} |h_n(\gamma, \boldsymbol{\theta})| \geqq d] \geqq 1 - \epsilon$$

for $n \geqq n_\epsilon$.

PROOF. By definition $g_n(\gamma, \boldsymbol{\theta})$ is a monotone function of γ for every $\|\boldsymbol{\theta}\| = 1$. Therefore (3.72) holds, in view of (3.58) and (3.60). On the other hand, h_n can be rewritten as $(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})^{\frac{1}{2}}h_n(\gamma, \boldsymbol{\theta}) = n^{-\frac{3}{2}}(n + 1)^{-1} \sum_{i=1}^n d_{in}R_{in} \operatorname{sgn}(Y_{in} - \gamma d_{in})$ where $d_{in} = \boldsymbol{\theta}\mathbf{X}_{in}'$ and R_{in} is the rank of $|Y_{in} - \gamma\, d_{in}|$.

To this statistic we can apply monotonicity Lemma 2.2 so that $h_n(\gamma, \boldsymbol{\theta})$ is also monotone in γ for every $\|\boldsymbol{\theta}\| = 1$. This plus (3.60) and (3.61) imply (3.72). This proves the lemma.

THEOREM 3.3. *Assume the conditions of Lemma 3.1, Theorem 3.1. For every $\epsilon > 0$ and a given $d > k_\alpha$ there exists n_ϵ and an a, $0 < a < \infty$ large enough such that*

$$(3.73)\quad P_n[\inf_{\|\theta\| \geqq an^{-\frac{1}{2}}} M_n(Y_n, \boldsymbol{\theta}) \geqq d] \geqq 1 - \epsilon,$$

$$(3.74)\quad P_n[\inf_{\|\theta\| \geqq an^{-\frac{1}{2}}} T_n(Y_n, \boldsymbol{\theta}) \geqq d] \geqq 1 - \epsilon$$

or all $n \geqq n_\epsilon$.

PROOF. Assume $n \geqq n_\epsilon$ so that Theorem 3.1, Lemmas 3.2 and 3.3 hold. Now, for any vector $\boldsymbol{\theta}$ in a direction there is another vector $\boldsymbol{\theta}^*$ in the same direction and a real number γ such that $\|\boldsymbol{\theta}^*\| = 1$, and

$$(3.75)\quad \boldsymbol{\theta} = \gamma\boldsymbol{\theta}^*.$$

Also from Rao [10, page 48], we have, because $\lim \Sigma_n$ is positive definite, for any vector $\boldsymbol{\theta}$, for $n \geqq n_\epsilon$

$$(3.76)\quad M_n(\boldsymbol{\theta}) \geqq |n^{\frac{1}{2}}\boldsymbol{\theta}'\mathbf{S}_n(\boldsymbol{\theta})|^2/(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta})$$

and

$$T_n(\boldsymbol{\theta}) \geqq |n^{\frac{1}{2}}\boldsymbol{\theta}'(\mathbf{S}_n(\mathbf{0}) - b(f)\Sigma_n\boldsymbol{\theta})|^2/(\boldsymbol{\theta}'\Sigma_n\boldsymbol{\theta}). \tag{3.77}$$

Therefore in view of (3.75), (3.77), (3.76) and the definitions (3.56) and (3.57) it is enough to prove (3.72) and (3.73), after noting that

$$[y; \inf_{\|\theta\| \geqq an^{-\frac{1}{2}}} M_n(y, \boldsymbol{\theta}) \geqq d]$$
$$= [y; \inf_{|\gamma| \geqq an^{-\frac{1}{2}}} M_n(y, \gamma\boldsymbol{\theta}^*) \geqq d \quad \text{for every} \quad \|\boldsymbol{\theta}^*\| = 1]$$

and the similar relation for the T_n statistic. The proof is terminated.

The immediate consequence of Theorem 3.3 is that for every $\epsilon > 0$ there exists an n_ϵ and $0 < a < \infty$, such that

$$P_n[y; n^{\frac{1}{2}}\mathcal{R}_n(y) \subset V(a)] \geqq 1 - \epsilon, \tag{3.78}$$

$$P_n[y; n^{\frac{1}{2}}D_n(y) \subset V(a)] \geqq 1 - \epsilon \tag{3.79}$$

for all $n \geqq n_\epsilon$, and hence there exists a constant $k < \infty$, depending on a, such that

$$P_n[y; \lambda(n^{\frac{1}{2}}\mathcal{R}_n(y)) \leqq k] \geqq 1 - \epsilon \qquad \text{for} \quad n \geqq n_\epsilon. \tag{3.80}$$

4. Asymptotic normality, and efficiency. The asymptotic normality of $n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n$ will be derived by approximating it by $n^{\frac{1}{2}}\tilde{\boldsymbol{\theta}}_n$ in probability. For this we need to show that the Lebesgue measures of the two regions $n^{\frac{1}{2}}\mathcal{R}_n(y)$ and $n^{\frac{1}{2}}D_n(y)$ are close to each other in probability. The nature of $n^{\frac{1}{2}}D_n(y)$, together with Theorem 3.2, enables us to conclude this fact. We begin with the following definitions. Define for random sample y and a constant a, the sets

$$W_n(y) = n^{\frac{1}{2}}\mathcal{R}_n(y)\Delta n^{\frac{1}{2}}D_n(y), \tag{4.1}$$

$$K_n(a, \mathbf{t}) = \{y; \sup_{t \in V_n(a)} |M_n(y, \mathbf{t}) - T_n(y, \mathbf{t})| \geqq \epsilon\} \tag{4.2}$$

$$Q_n(a) = \{y; W_n(y) \subset V(a)\}, \tag{4.3}$$

and

$$U_n(y, \epsilon) = \{\mathbf{t}, k_\alpha - \epsilon \leqq T_n(y, \mathbf{t}n^{-\frac{1}{2}}) \leqq k_\alpha + 2\epsilon\}. \tag{4.4}$$

LEMMA 4.1. *If the conditions of Theorem 3.1 and Lemma 3.1 are satisfied, then for every $\epsilon > 0$ there exists n_ϵ such that $n \geqq n_\epsilon$ yields*

$$P_n[n^{\frac{1}{2}}\|\hat{\boldsymbol{\theta}}_n(Y_n) - \tilde{\boldsymbol{\theta}}_n(Y_n)\| \geqq \epsilon] \leqq \epsilon$$

where $\tilde{\boldsymbol{\theta}}_n$ is defined by (3.55).

PROOF. We will first show that for every $\epsilon > 0$

$$P_n[y; \lambda(W_n(y)) \leqq \epsilon] \geqq 1 - \epsilon \qquad \text{for} \quad n \geqq n_\epsilon. \tag{4.5}$$

From Theorem 3.2 it follows that for every $\epsilon > 0$ there exists n_ϵ and an a, $0 < a < \infty$ such that

$$P_n[K_n{}^c(a, \epsilon) \cap Q_n(a)] \geqq 1 - \epsilon \tag{4.6}$$

and

$$P_n[Q_n(a)] \geqq 1 - \epsilon \tag{4.7}$$

for all $n \geqq n_\epsilon$.

But for a $y \varepsilon K_n^c(a, \epsilon) \cap Q_n(a)$, a $\mathbf{t}$ belonging to $W_n(y)$ is such that for every $\epsilon > 0$ it belongs also to $U_n(y, \epsilon)$. Therefore, for every $\epsilon > 0$ there exists n_ϵ such that for $n \geqq n_\epsilon$

$$P_n[y; W_n(y) \subset U_n(y, \epsilon)] \geqq 1 - \epsilon. \tag{4.8}$$

But because of the definition of T_n it is easy to see that

$$\lambda(U_n(y, \epsilon)) = d_n\{(k_\alpha + 2\epsilon)^{p/2} - (k_\alpha - \epsilon)^{p/2}\}, \tag{4.9}$$

where d_n is a constant depending on n only through $|\Sigma_n|$ and hence by assumption $\lim d_n < \infty$. Obviously (4.9) can be made small for an arbitrarily small ϵ. This, plus (4.7), yields (4.5).

Therefore, in view of (3.8), we can conclude that

$$P_n[y: |\lambda(n^{\frac{1}{2}}\mathcal{R}_n(y)) - \lambda\{n^{\frac{1}{2}}D_n(y)\}| \geqq \epsilon] \leqq \epsilon \tag{4.10}$$

for every $\epsilon > 0$ and all $n \geqq n_\epsilon$.

And, since by (1.11) $k_\alpha > 0$, there exists a $\delta > 0$ such that

$$P_n[y; \lambda(n^{\frac{1}{2}}\mathcal{R}_n(y)) \geqq \delta > 0] \geqq 1 - \epsilon \tag{4.11}$$

for every $\epsilon > 0$ and $n \geqq n_\epsilon$.

Also note that

$$\lim \lambda(n^{\frac{1}{2}}D_n(y)) \geqq \delta > 0 \qquad \text{with probability 1.} \tag{4.12}$$

Now consider the difference

$$\begin{aligned} &n^{\frac{1}{2}}\|\hat{\boldsymbol{\theta}}_n(y) - \tilde{\boldsymbol{\theta}}_n(y)\| \\ &\leqq [\lambda(n^{\frac{1}{2}}\mathcal{R}_n(y))\lambda(n^{\frac{1}{2}}D_n(y))]^{-1}|\lambda(n^{\frac{1}{2}}\mathcal{R}_n(y)) - \lambda(n^{\frac{1}{2}}D_n(y))| \\ &\quad \cdot \int_{n^{\frac{1}{2}}\mathcal{R}_n(y)} \|\mathbf{t}\| \, d\lambda(\mathbf{t}) + [\lambda(n^{\frac{1}{2}}D_n(y))]^{-1} \int \|\mathbf{t}\| I(W_n(y)) \, d\lambda(\mathbf{t}). \end{aligned} \tag{4.13}$$

By (3.78), for every $\epsilon > 0$

$$P_n[y; \int_{n^{\frac{1}{2}}\mathcal{R}_n(y)} \|\mathbf{t}\| \cdot d\lambda(\mathbf{t}) \leqq a \cdot \lambda(n^{\frac{1}{2}}\mathcal{R}_n(y)) \leqq ak_1 < \infty] \geqq 1 - \epsilon \tag{4.14}$$

where $k_1 = \lambda(V(a))$.

By (4.10), (4.12) and (4.9) the multiplier of the integral on the right hand side of (4.13) is arbitrarily small with high probability. Combining this with (4.14), the first term on the right hand side of (4.13) can be at most equal to $\epsilon/2$ for $n \geqq n_\epsilon$ with P_n probability at least $1 - \epsilon/2$, for every $\epsilon > 0$.

Similarly, by (4.6) and (4.7),

$$P_n[y; \int \|\mathbf{t}\| I(W_n(y)) \, d\lambda(\mathbf{t}) \leqq \epsilon] \geqq 1 - \epsilon \tag{4.15}$$

for $n \geqq n_\epsilon$ and for every $\epsilon > 0$. Therefore, since $\lambda(n^{\frac{1}{2}}D_n(y))$ has a finite limit with probability 1, the second term of (4.12) can also be made to be at most equal to $\epsilon/2$ with P_n probability at least $1 - \epsilon/2$ for $n \geqq n_\epsilon$. This completes the proof.

We are now ready to state the main theorem of this section.

THEOREM 4.1. *Let* $\{Y_{in}\ 1 \leqq i \leqq n\}\ n \geqq 1$ *be the sequence of random variables as given by model* 1.0. *Let* $\{x_{i\nu}\}$ *satisfy the conditions of* (1.2) *and* (1.3). *Let* $F \varepsilon \mathfrak{F}$, *where* $\mathfrak{F}$ *is defined in Lemma* 3.1. *Then*

$$\lim \mathcal{L}_{\theta_0}(n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)) = N(\mathbf{0}, \Sigma^{-1}(b(f))^{-2}). \tag{4.16}$$

(N stands for the multivariate normal distribution and $\Sigma^{-1} = \lim \Sigma_n^{-1}$.)

PROOF. First of all, in view of Lemma 2.1, it is enough to prove (4.15) if $\boldsymbol{\theta}_0 = \mathbf{0}$. Under the assumptions made here, the conclusions of Theorem 3.1, Lemma 3.2, and hence of Theorem 3.2 and Lemma 4.1, hold. Therefore it is enough to show that

$$\lim \mathcal{L}_0(n^{\frac{1}{2}}\tilde{\boldsymbol{\theta}}_n) = N(\mathbf{0}, \Sigma^{-1}(b(f))^{-2}). \tag{4.17}$$

However, by the definition $n^{\frac{1}{2}}\tilde{\boldsymbol{\theta}}_n = n^{\frac{1}{2}}(b(f)^{-1}\Sigma_n^{-1}\hat{\mathbf{S}}_n(\mathbf{0})$, and the fact that we know

$$\mathcal{L}_\theta(n^{\frac{1}{2}}\hat{\mathbf{S}}_n(\mathbf{0})) \rightarrow N(\mathbf{0}, \Sigma),$$

we conclude that

$$\mathcal{L}_0(n^{\frac{1}{2}}\tilde{\boldsymbol{\theta}}_n) \rightarrow N(\mathbf{0}, \Sigma^{-1}(b(f))^{-2}).$$

This proves the theorem.

Asymptotic efficiency. Essentially, under the conditions (1.2), (1.3) on the regression scores and $F \varepsilon \mathfrak{F}$, Eicker [6] has proved that the asymptotic distribution of the least squares estimate $n^{\frac{1}{2}}\boldsymbol{\theta}_n^*$ is $N(\mathbf{0}, \tilde{\Sigma}^{-1})$. Actually, our conditions imply his conditions. $\tilde{\Sigma}$ is defined by (1.3). By (1.9), we have $\Sigma_n = \frac{1}{3}\tilde{\Sigma}_n$, because $\psi(u) = u$.

If we define the asymptotic efficiency of $n^{\frac{1}{2}}\hat{\boldsymbol{\theta}}_n$ relative to the least squares estimator on the inverse ratio of their generalized limiting variances, and denote it by $e(F)$, then we have

$$\begin{aligned} e(F) &= \lim b^2(f)(1/|\Sigma_n^{-1}|)/[3\cdot(1/|\Sigma_n^{-1}|)] = \tfrac{1}{3}b^2(f) \\ &= 12(\textstyle\int_{-\infty}^{\infty} f^2(x)\,dx). \end{aligned} \tag{4.18}$$

It might be noted that this is nothing but the asymptotic efficiency of the M_n statistic relative to the corresponding F statistic based on least squares estimates. See [1].

A consistent estimator of $(\int_{-\infty}^{\infty} f^2(x)\,dx)$. Let

$$b_1(f) = \textstyle\int_{-\infty}^{\infty} f^2(x)\,dx, \tag{4.19}$$

$$K_{\alpha,p} = (6)^p|\Sigma|^{\frac{1}{2}}\Gamma((p/2) + 1)(\chi^2_{p,\alpha}\pi)^{-p/2}. \tag{4.20}$$

Define

$$[\hat{b}_1(f)]^{-p} = \lambda(n^{\frac{1}{2}}\mathcal{R}_n(y))K_{\alpha,p}. \tag{4.21}$$

$\hat{b}_1(f)$ defined by (4.21) is the proposed estimator of $b_1(f)$. That it is consistent follows from (3.54), (4.10), and Lemma 1.1, which enables us to replace k_α in (3.53) by $\chi^2_{p,\alpha}$.

APPENDIX

Notation here is the same as in Section 3.

Here it will be shown that the stochastic processes $\{L_n(\mathbf{t}, x), -\infty \leqq x \leqq \infty\}$ and $\{Z_n(\mathbf{t}, |x|) -\infty \leqq x \leqq +\infty\}$ converge weakly to some processes, uniformly in $\mathbf{t} \varepsilon V_n(a)$. For this we shall make use of the following known results, which can be found in Billingsley [3].

Let $\{V_n(x) - 1 \leqq x \leqq +1\}$ be a sequence of stochastic process. Define for some $\delta > 0$

$$W''(V_n, \delta) = \sup_{x_1 \leqq x \leqq x_2,\, x_2 - x_1 \leqq \delta} \min \{|V_n(x) - V_n(x_1)|, |V_n(x_2) - V_n(x)|\}.$$

THEOREM A1. *Let $\{V_n(x) - 1 \leqq x \leqq +1\}$ be a sequence of stochastic processes in $D[-1, +1]$. Suppose that*

$$\mathfrak{L}(V_n(x_1) \cdots V_n(x_k)) \to \mathfrak{L}(V(x_1) \cdots V(x_k)) \tag{1}$$

holds for all continuity points of V. Suppose further that

$$\Pr[V(1) \neq V(1-)] = 0. \tag{2}$$

Suppose finally that for each $\epsilon > 0$

$$\lim_{\delta \to 0} \lim_{n \to \infty} P[W''(V_n, \delta) \geqq \epsilon] = 0; \tag{3}$$

then $V_n \to_D V$. (For the definitions of D see [3].)

From the statement and the proof of Theorem 15.6 of [3] one can deduce the following

THEOREM A2. *Suppose there are two constants $\gamma \geqq 0$ and $\alpha > \frac{1}{2}$ and a nondecreasing continuous function G_n on $[-1, +1]$, such that for each n, if $t_1 \leqq t \leqq t_2$,*

$$E[|V_n(t) - V_n(t_1)|^\gamma [V_n(t_2) - V_n(t)|^\gamma] \leqq [G_n(t) - G_n(t_1)]^\alpha [G_n(t_2) - G_n(t)]^\alpha \tag{3'}$$

holds for all t_1, t, t_2.

Then for any $\epsilon > 0$

$$P(W''(V_n, \delta) \geqq \epsilon) \leqq \epsilon^{-2\gamma} K(\Sigma_n' + \Sigma_n'') \tag{4}$$

where K is a constant independent of n but may depend on γ and α, and where Σ_n' and Σ_n'' are the sums of the form $\sum_{k=1}^r [G_n(t_k) - G_n(t_{k-1})]^{2\alpha}$ with $-1 \leqq t_1 \leqq t_2 \leqq \cdots \leqq t_r \leqq 1$ and $t_k - t_{k-1} \leqq 2\delta$, $k = 1, \cdots r$.

REMARK. Expression (4) corresponds to (15.29) of [3].

We now consider the following stochastic processes. Define, for $-\infty \leqq x \leqq +\infty$, $\mathbf{t}$ a p-vector,

$$W_n(t, x) = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in} [I(Y_{in} - \delta_{in}(t) \leqq x) - F(x + \delta_{in}(t))] \tag{5}$$

with $\delta_{in}(t) = n^{-\frac{1}{2}}\mathbf{t}'\mathbf{X}_{in}$, $\{c_{in}\}$ some constants. In the following, whenever convenient, we will write t for $\mathbf{t}$ and 0 for $\mathbf{0}' = (0, \cdots, 0)$.

THEOREM A3. *If* $\{x_{i\nu}\}$ *satisfy* (1.2) *and* (1.3), $\{c_{in}\}$ *satisfy* (1.2), *and* $n^{-1}\Sigma c_{in}^2 \to C^2 < K' < \infty$, *and if* F *is absolutely continuous with uniformly bounded and continuous density, then for each fixed* t

$$(6) \qquad \{W_n(t, x), -\infty \leqq x \leqq +\infty\} \to_D \{L(x), -\infty \leqq x \leqq +\infty\}$$

where L *is a Gaussian process with continuous sample functions. Consequently for every* $\epsilon > 0$ *and each fixed* $\mathbf{t}$,

$$(7) \qquad \lim_{h\to 0} \lim_{n\to\infty} \Pr [\sup_{|x-y|\leqq h} |W_n(t, x) - W_n(t, y)| \geqq \epsilon] = 0.$$

PROOF. It is easy to see that $W_n(t, x)$ is a stochastic process in $D[-\infty, +\infty]$ and $W_n(t, -\infty) = 0$, $W_n(t, +\infty) = 0$ hold, each with probability 1 for all n.

Since $\mathbf{t}$ is fixed, we will suppress $\mathbf{t}$ in our notation, for convenience, and define

$$(8) \qquad Q_n(x) = W_n(\mathbf{t}, \tan(\tfrac{1}{2}\pi x)), \qquad -1 \geqq x \leqq +1.$$

Then $Q_n \,\varepsilon\, D[-1, +1]$ for all n.

Under this transformation, since L is continuous at $+\infty$, it is enough to prove

$$(9) \qquad \{Q_n(x), -1 \leqq x \leqq +1\} \to_D \{Q(x), -1 \leqq x \leqq +1\}$$

with $Q(x) = L(\tan(\frac{1}{2}\pi x))$. This is so because $\tan(\frac{1}{2}\pi)$ is a homeomorphism between $D[-1, 1]$ and $D[-\infty, +\infty]$, when $D[-1, +1]$ is metrized by Skorhods metric. For the definition of Skorhods metric see [3].

Thus it is enough to verify the conditions of the Theorem A1. We begin by verifying (3). This we will do via (3′) and (4). First we shall exhibit a $\gamma > 0$, and $\alpha > \frac{1}{2}$ and G_n such that (3′) holds.

Let

$$(10) \qquad g(x) = \tan(\tfrac{1}{2}\pi x), \qquad -1 \leqq x \leqq +1,$$

$$(11) \qquad \alpha_{in} = I[g(x_1) \leqq Y_{in} - \delta_{in} \leqq g(x)] - [p_{in}(x) - p_{in}(x_1)],$$

$$(12) \qquad \beta_{in} = I[g(x) \leqq Y_{in} - \delta_{in} \leqq g(x_2)] - [p_{in}(x_2) - p_{in}(x)],$$

with $\delta_{in} = \delta_{in}(t)$ and

$$(13) \qquad p_{in}(x) = F(g(x) + \delta_{in}).$$

Recalling the definition of W_n from (5) we can write, in view of (13), (12), (11), and (10),

$$(14) \qquad [Q_n(x) - Q_n(x_1)] = n^{-\frac{1}{2}}\Sigma c_{in}\alpha_{in},$$

$$(15) \qquad [Q_n(x_2) - Q_n(x)] = n^{-\frac{1}{2}}\Sigma c_{in}\beta_{in}.$$

Consequently, using the independence of $\{(\alpha_{in}, \beta_{in}), 1 \leqq i \leqq n\}$, $\{\alpha_{in}, 1 \leqq i \leqq n\}$, $\{\beta_{in}, 1 \leqq i \leqq n\}$ and that α_{in} are independent of β_{jn} $i \neq j = 1, \cdots, n$ and that $E\alpha_{in} = E\beta_{in} = 0$, $1 \leqq i \leqq n$, one can see that

$$(16)\quad E[Q_n(x) - Q_n(x_1)]^2[Q_n(x_2) - Q_n(x)]^2 = n^{-2}[\textstyle\sum_{i=1}^n c_{iv}^4 E\alpha_{in}^2\beta_{in}^2 + 2\sum\sum_{i\neq j=1}^n c_{in}^2 c_{jn}^2 E\alpha_{in}^2 E\beta_{jn}^2 + \sum\sum_{i\neq j=1}^n c_{in}^2 c_{jn}^2 E\alpha_{in}\beta_{in}E\alpha_{jn}\beta_{jn}].$$

Using the definitions (12), (11) and

$$(17)\qquad a_{in} = p_{in}(x) - p_{in}(x_1), \qquad b_{in} = p_{in}(x_2) - p_{in}(x),$$

one can see that $E\alpha_{in}^2\beta_{in}^2 \leqq 3a_{in}b_{in}$, $E\alpha_{in}^2 E\beta_{jn}^2 \leqq a_{in}b_{jn}$, and $E\alpha_{in}\beta_{in}E\alpha_{jn}\beta_{jn} \leqq a_{in}b_{jn}$.

Consequently the right side of (16) is bounded by

$$n^{-2}[3\textstyle\sum_{i=1}^n c_{in}^4 a_{in}b_{in} + 3\sum\sum_{i\neq j=1}^n c_{in}^2 c_{jn}^2 a_{in}b_{jn}] = 3n^{-2}(\Sigma c_{in}^2 a_{in})(\sum_i c_{in}^2 b_{in}).$$

But note that

$$n^{-1}\textstyle\sum c_{in}^2 a_{in} = n^{-1}\sum c_{in}^2[F(g(x) + \delta_{in}) - F(g(x_1) + \delta_{in})]$$

and $$n^{-1}\textstyle\sum c_{in}^2 b_{in} = n^{-1}\sum c_{in}^2[F(g(x_2) + \delta_{in}) - F(g(x) + \delta_{in})].$$

Consequently $(3')$ is satisfied for $\gamma = 2$, $\alpha = 1$ and $G_n(x) = n^{-1}\sum c_{in}^2 F(g(x) + \delta_{in})$.

Therefore (4) of Theorem A2 is satisfied.

However,

$$(18)\quad \textstyle\sum_{k=1}^r [G_n(x_k) - G_n(x_{k-1})]^2 \leqq \max_{1\leqq k\leqq r} [G_n(x_k) - G_n(x_{k-1})]\sum_{k=1}^r G_n[(x_k) - G_n(x_{k-1})].$$

Now note that by (1.2) and (1.3) $\max_{1\leqq i\leqq n}\delta_{in} \to 0$ as $n \to \infty$, and that by assumption, $\lim_{n\to\infty} n^{-1}\sum c_{iv}^2 = K' < \infty$. Define $G(x) = F(g(x))$ for $-1 \leqq x \leqq +1$.

Observe that F continuous implies that G_n and G are uniformly continuous for all n and that $\lim_{n\to\infty} G_n(x) = K'[G(x)]$ where $K' = \lim_{n\to\infty} n^{-1}\sum c_{in}^2$.

Also note that for $-1 \leqq x_1 \leqq \cdots \leqq x_r \leqq 1$,

$$\lim_{n\to\infty} \textstyle\sum_{k=1}^r [G_n(x_k) - G_n(x_{k-1})] \leqq \lim_{n\to\infty} [G_n(1) - G_n(-1)] = K' < \infty.$$

Therefore

$$(19)\quad \lim_{n\to\infty} \max_{1\leqq k\leqq r} |G_n(x_k) - G_n(x_{k-1})| \textstyle\sum_{k=1}^r [G_n(x_k) - G_n(x_{k-1})] \leqq K' \max_{1\leqq k\leqq r} |G(x_k) - G(x_{k-1})|,$$

and now it is obvious that for δ sufficiently small the right hand side of the above can be made arbitrarily small for G being uniformly continuous. Hence by (4), (18) and (19)

$$(20)\qquad \lim_{\delta\to 0}\lim_{n\to\infty} \Pr[W''(Q_n, \delta) \geqq \epsilon] = 0 \quad \text{for every} \quad \epsilon > 0.$$

Next we show that

$$(21)\qquad \lim_{n\to\infty} \mathfrak{L}(Q_n(x_j),\ j = 1, \cdots, k) = \text{MVN}.$$

MVN = Multivariate Normal law.

Now let $\boldsymbol{\lambda}' = (\lambda_1, \cdots, \lambda_k)$ be a k-dimensional real vector and $\mathbf{Q}_n = (Q_n(x_1), \cdots, Q_n(x_k))$. Then to prove (21), it is enough to prove (see [10])

$$(22) \qquad \lim_{n\to\infty} \mathcal{L}(\boldsymbol{\lambda}'\mathbf{Q}_n) = \text{Normal Law for every } \boldsymbol{\lambda}.$$

However, $\lambda' Q_n = \sum_{s=1}^k \lambda_s W_n(t, g(x_s)) = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in} Z_{in}$ with

$$(23) \qquad Z_{in} = \sum_{s=1}^k \lambda_s \{I[Y_{in} - \delta_{in} \leqq g(x_s)] - F(g(x_s) + \delta_{in})\}.$$

Note that $\{Z_{in}, 1 \leqq i \leqq n\}$ are independent rv's.

Under the condition of the theorem it can be shown that the Lindeberg-Feller Theorem is applicable to the rv's $\boldsymbol{\lambda}'\mathbf{Q}_n$ for every $\boldsymbol{\lambda}$. Consequently (21) follows. It is clear that (21) holds at all continuity points $(x_1, \cdots, x_k)$ of $Q = L \tan(\frac{1}{2}\pi)$ because $L \tan(\frac{1}{2}\pi)$ has continuous sample paths almost surely. Also the condition (2) of Theorem A1 is then trivially satisfied. Thus, in view of (21) and (20) and the above remarks, conditions of Theorem A1 are satisfied and hence the consequences (6). (7) is a well known consequence in view of the Gaussian limit with continuous sample paths. The proof is terminated.

Remark. The above theorem remains true also for the processes defined by

$$(24) \quad W_n^*(t, x) = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}[I(Y_{in} - \delta_{in} + \lambda_{in} \leqq x] - F(x + \delta_{in} - \lambda_{in})]$$

where $\lambda_{in} = \epsilon' n^{-\frac{1}{2}} \|\mathbf{X}_{in}\|$, $\epsilon' > 0$.

Lemma A1. *If F is absolutely continuous and has absolutely continuous and uniformly bounded density f, and if the constants $\{x_{i\nu}\}$ and $\{c_{in}\}$ satisfy the conditions of Theorem* A3, *then for any fixed t_0,*

$$(25) \quad \sup_{|x-y| \leqq \delta} \sup_{\| t-t_0 \| \leqq a} |\bar{J}_n(t, x) - \bar{J}(t, y) - \bar{J}_n(t_0, x) + \bar{J}_n(t_0, y)| \to 0$$

as $n \to \infty$ and then $\delta \to 0$, for any $0 < a < \infty$ fixed.

$$\bar{J}_n(t, x) = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in} F(x + \delta_{in}(t)).$$

Proof. Note that

$$\begin{aligned}
&\lim_{n\to\infty} \sup |\bar{J}_N(t, x) - \bar{J}_n(t, y) - \bar{J}_n(t_0, x) + \bar{J}_n(t_0, y)| \\
&\quad \leqq \lim_{n\to\infty} \sup |n^{-\frac{1}{2}} \sum c_{in}[F(x + \delta_{in}(t)) - F(y + \delta_{in}(t)) - F(x + \delta_{in}(t_0)) \\
&\qquad\quad + F(y + \delta_{in}(t_0))]| \\
&\quad \leqq \lim_{n\to\infty} \sup [n^{-1} \sum_{i=1}^n |c_{in}| \, \|\mathbf{X}_{in}\| \, \|\mathbf{t} - \mathbf{t}_0\| \, |f(x) - f(y)| \\
&\quad \leqq k_1 a \sup_{|y-x| \leqq \delta} |f(y) - f(x)|
\end{aligned}$$

where $k_1 = \lim n^{-1} \sum_{i=1}^n |c_{in}| \, \|\mathbf{X}_{in}\| < \infty$, and which now can be made arbitrarily small for small δ. Sup in the above inequalities up to the penultimate one is taken over all $x, y, |y - x| \leqq \delta$ and all t; $\|r - t_0\| \leqq a$. The proof is completed.

Lemma A2. *Let the conditions of Theorem* A3 *be satisfied. Then for each fixed* $\mathbf{t}$,

$$(26) \qquad \lim_{n\to\infty} P_n[\sup_{-\infty \leqq x \leqq +\infty} |W_n(\mathbf{t}, x) - W_n(\mathbf{0}, x)| \geqq \epsilon] = 0$$

for every $\epsilon < 0$.

PROOF. (26) will follow if we show

(i) $\lim_{n\to\infty} P_n[|W_n(t, x) - W_n(0, x)| \geqq \epsilon] = 0$ for every $\epsilon > 0$, and that

(ii) for every $\epsilon \geqq 0$

$$\lim_{\delta\to 0} \lim_{n\to\infty} P_n[\sup_{|x-y| \leqq \delta} |W_n(t, x) - W_n(0, x) - W_n(t, y) + W_n(0, y)| \geqq \epsilon] = 0.$$

For, if (i) and (ii) are satisfied, then the stochastic process $\{|W_n(t, x) - W_n(0, x)| \ -\infty \leqq x \leqq +\infty\}$ is relatively compact, with degenerate process, degenerate at zero, as its limit. However, from the inequality

$$\sup_{|x-y| \leqq h} |W_n(t, x) - W_n(0, x) - W_n(t, y) + W_n(0, y)|$$
$$\leqq \sup_{|x-y| \leqq h} |W_n(t, x) - W_n(t, y)| + \sup_{|x-y| \leqq h} |W_n(0, x) - W_n(0, y)|,$$

(ii) follows immediately by applying (7) to $W_n(t, x)$ and $W_n(0, x)$ separately. It remains to prove (i).

By definition

$$W_n(t, x) - W_n(0, x) = n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n c_{in}\{I(Y_{in} \leqq x + \delta_{in}(t)) - I(Y_{in} \leqq x) - p_{in}(t, x) + p_{in}(0, x)\} \tag{27}$$

where $p_{in}(\mathbf{t}, x)$ is exactly the $p_{in}(x)$ defined by (13). Notice in (13) $p_{in}(x)$ depends on $\mathbf{t}$ through $\delta_{in} = \delta_{in}(\mathbf{t})$. Let

$$U_{in}(t, x) = \{I(Y_{in} - \delta_{in}(t) \leqq x) - I(Y_{in} \leqq x) - p_{in}(t, x) + p_{in}(0, x)\}. \tag{28}$$

(28) and (27) yield

$$W_n(t, x) - W_n(0, x) = n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n c_{in} U_{in}(t, x). \tag{29}$$

Now for each t, x fixed, $U_{in}(t, x)$ are independent random variables and have means zero. If we show that the variance of the above sum goes to zero, we will have completed the proof in view of the Chebychev inequality. But

$$\begin{aligned} &\operatorname{Var}(\{W_n(t, x) - W_n(0, x)\}) \\ &\quad = n^{-1}\textstyle\sum c_{in}^2 \operatorname{Var}(U_{in}(t, x)) \\ &\quad = n^{-1}\textstyle\sum^+ x_{i\nu}^2 \operatorname{Var}(U_{in}(t, x)) + n^{-1}\textstyle\sum^- x_{i\nu}^2 \operatorname{Var}(U_{in}(t, x)) \end{aligned} \tag{30}$$

where $\sum^+$ and $\sum^-$ are the summations over those i for which $\delta_{in}(t) \geqq 0$ and $\delta_{in}(t) \geqq 0$ respectively.

Let $\delta_{in}(t) \geqq 0$. Then,

$$U_{in}(t, x) = [I(z < Y_{in} \leqq x + \delta_{in}(t)) - \{p_{in}(t, x) - p_{in}(0, x)\}]$$

where $z = \max(x, \delta_{in}(t))$. Consequently, in this case,

$$\begin{aligned} \operatorname{Var}(U_{in}(\mathbf{t}, x)) = {} & [F(x + \delta_{in}(t)) - F(z)][1 - p_{in}(t, x) + p_{in}(0, x)]^2 \\ & + [1 - F(x + \delta_{in}(t)) + F(z)][p_{in}(\mathbf{t}, x) - p_{in}(0, x)]^2. \end{aligned}$$

Now, since F is uniformly continuous and $\max_{1 \leqq i \leqq n} \delta_{in}(\mathbf{t}) \to 0$ as $n \to \infty$, it can be easily seen that

$$(31) \qquad n^{-1} \sum^{+} c_{in}^2 \operatorname{Var}(U_{in}(t, x)) \to 0 \quad \text{as} \quad n \to \infty .$$

The case when $\delta_{in}(\mathbf{t}) < 0$ can be handled similarly. Therefore, combining the above remarks, (30) and (31), we can conclude (i). This terminates the proof.

LEMMA A3. *Let the conditions of Theorem* A3 *hold. Then given an* $\epsilon > 0$, *there is an* $\epsilon' > 0$ *such that for each fixed* $\mathbf{t}_0$

$$(32) \quad \lim_{\delta \to 0} \lim_{n \to \infty} P_n[\sup_{|x-y| \leqq \delta} \sup_{\| t - t_0 \| \leqq \epsilon'} |W_n(t, x) - W_n(t, y) - W_n(t_0, x) + W_n(t_0, y)| \geqq \epsilon] = 0.$$

PROOF. Let

$$J_n(t, y, x) = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}[I(Y_{in} - \delta_{in}(t) \leqq y) - I(Y_{in} - \delta_{in}(t) \leqq x)]$$

and $\bar{J}_n(t, y, x) = EJ_n(t, y, x)$, so that

$$W_n(t, y) - W_n(t, x) - W_n(t_0, y) + W_n(t_0, y) = [J_n(t, y, x) - J_n(t_0, y, x)] - [\bar{J}_n(t, y, x) - \bar{J}_n(t_0, y, x)].$$

Therefore, in view of Lemma A1, it will be enough to prove that given an $\epsilon > 0$, there is an $\epsilon' > 0$ such that for each $\mathbf{t}_0$

$$(33) \quad \lim_{\delta \to 0} \lim_{n \to \infty} P_n[\sup_{|x-y| \leqq \delta} \sup_{\| t - t_0 \| \leqq \epsilon'} |J_n(t, y, x) - J_n(t_0, y, x)| \geqq \epsilon] = 0.$$

Without loss of generality let $x < y$. Assume $c_{in} \geqq 0$, $1 \leqq i \leqq n$. Note that, with

$$(34) \qquad \lambda_{in}(\epsilon') = \sup_{\| t \| \leqq \epsilon'} |\delta_{in}(t)| = n^{-\frac{1}{2}} \epsilon' \|\mathbf{X}_{in}\|$$

we have

$$(35) \qquad \delta_{in}(\mathbf{t}) \leqq \delta_{in}(\mathbf{t}_0) + \lambda_{in}(\epsilon'), \qquad 1 \leqq i \leqq n.$$

Combining (35), (36) and the fact $c_{in} \geqq 0$, with $x < y$, one can see that

$$\begin{aligned} &[J_n(t, y, x) - J_n(t_0, y, x)] \\ &\quad \leqq n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}\{I(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'), y + \delta_{in}(t_0) + \lambda_{in}(\epsilon')] \\ &\qquad - I(x + \delta_{in}(t_0), y + \delta_{in}(t_0)]\} \\ (36) \quad &\quad = n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}\{I(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'), x + \delta_{in}(t_0)]\} + n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}\{I(y \\ &\qquad + \delta_{in}(t_0), y + \delta_{in}(t_0) + \lambda_{in}(\epsilon')]\} \end{aligned}$$

where $I(a, b] = I(a < Y_{in} \leqq b)$ for any $a < b$. Similarly one sees

$$\begin{aligned} (37) \quad [J_n(t, y, x) - J_n(t, y, x)] \geqq &-n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}\{I(x + \delta_{in}(t_0), x + \delta_{in}(t_0) \\ &+ \lambda_{in}(\epsilon')]\} - n^{-\frac{1}{2}} \sum_{i=1}^n c_{in}\{I(y + \delta_{in}(t_0) - \lambda_{in}(\epsilon'), y + \delta_{in}(t_0)]\}. \end{aligned}$$

Now define stochastic processes

$$M_n(x) = n^{-\frac{1}{2}}\sum_{i=1}^{n} c_{in}\{I(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'), x + \delta_{in}(t_0)] - F(x + \delta_{in}(t_0)) + F(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'))\}.$$

Observe that

$$\text{(38)} \quad \sup_x |\text{first term on the right hand side of 36}| \leqq \sup_x M_n(x) + \sup_x |n^{-\frac{1}{2}}\sum_{i=1}^{n} c_{in}\{F(x + \delta_{in}(t_0)) - F(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'))\}|.$$

Now observe that since F has bounded and absolutely continuous density f, we have

$$\sup_x |n^{-\frac{1}{2}}\sum c_{in}\{F(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon')) - F(x + \delta_{in}(t_0)\}| \leqq \epsilon' n^{-1}\sum |c_{in}|\, \|\mathbf{X}_{in}\| \sup_x f(x + \delta_{in}(t_0)) \leqq \epsilon' k_2$$

where k_2 is such that

$$\lim_{n\to\infty} n^{-1}\sum |c_{in}|\, \|\mathbf{X}_{in}\| \sup_x f(x + \delta_{in}(t_0) \leqq k_2\,,$$

the existence of which is guaranteed by assumption.

So we choose ϵ' such that $k_2\epsilon' \leqq \epsilon/2$. Next we will show that $\sup_x |M_n(x)| \to 0$ in P_n-probability.

First observe that for each x,

$$\lim_{n\to\infty} \operatorname{Var}(M_n(x)) \leqq \lim_{n\to\infty} n^{-1}\sum c_{in}^2[F(x + \delta_{in}(t_0)) - F(x + \delta_{in}(t_0)) - \lambda_{in}(\epsilon')] = 0,$$

so that finite dimensional distribution n of M_n tends to degenerate distribution, degenerate at 0. Next observe that, for $x < z$, ($x > z$ is similarly handled)

$$\begin{aligned}
|M_n(z) - M_n(x)| &= |n^{-\frac{1}{2}}\sum c_{in}\{I[z - \lambda_{in}(\epsilon') + \delta_{in}(t_0), z + \delta_{in}(t_0)] - I[x + \delta_{in}(t_0) \\
&\quad - \lambda_{in}(\epsilon'), x + \delta_{in}(t_0)] - F(z + \delta_{in}(t_0)) + F(z - \lambda_{in}(\epsilon') + \delta_{in}(t_0) \\
&\quad + F(x + \delta_{in}(t_0)) - F(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'))\}| \\
\text{(39)} \qquad &= |n^{-\frac{1}{2}}\sum_{i=1}^{n} c_{in}\{I[x + \delta_{in}(t_0), z + \delta_{in}(t_0)] - F(z + \delta_{in}(t_0) \\
&\quad + F(x + \delta_{in}(t_0))\} - n^{-\frac{1}{2}}\sum_{i=1}^{n} c_{in}\{I[x + \delta_{in}(t_0) \\
&\quad - \lambda_{in}(\epsilon'), z + \delta_{in}(t_0) - \lambda_{in}(\epsilon')] - F(z + \delta_{in}(t_0) - \lambda_{in}(\epsilon')) \\
&\quad + F(x + \delta_{in}(t_0) - \lambda_{in}(\epsilon'))\}| \\
&= |\{W_n(t_0\,, z) - W_n(t_0\,, x)\} - \{W_n^*(t_0\,, z) - W_n^*(t_0\,, x\}|
\end{aligned}$$

where W_n^* is defined in the remark after Theorem A3. This relation, consequently, along with (7) and the remark after Theorem A3, gives us

$$\sup_{|x-y|\leqq\delta} |M_n(x) - M_n(y)| \to 0$$

in P_n-probability as $n \to \infty$, then $\delta \to 0$. Therefore $\{M_n(x); -\infty \leqq x \leqq +\infty\} \to_D 0$ and hence $\sup_x |M_n(x)| \to 0$ in P_n-probability.

Therefore, if we choose an ϵ' such that $k_2\epsilon' \leqq \epsilon/2$, we would have then made the first term of the right hand side of (36) arbitrarily small for arbitrarily small ϵ, and large n. It may be shown similarly that for the same choice of ϵ', the second term on the right side of (36) can be made very small for large n. Again a similar argument as given above will show that sup |right hand side of (37)| $\to 0$ in probability. Thus we have proved the theorem when $c_{in} \geqq 0$, $1 \leqq i \leqq n$. For $\{c_{in}\}$ $1 \leqq i \leqq n$, with variable sign, we use decomposition of $c_{in} = c_{in}^+ - c_{in}^-$ into positive and negative parts and the result then follows because of linearity of J_n function in $\{c_{in}\}$ and the above argument for nonnegative $\{c_{in}\}$'s. The proof is terminated.

LEMMA A4. *Under the conditions of Theorem* A3, *for every* $\epsilon > 0$ *and any fixed* $0 < a < \infty$,

$$\lim_{\delta\to 0} \lim_{n\to\infty} P_n[\sup_{|x-y| \leqq \delta} \sup_{t \varepsilon V(a)} |W_n(\mathbf{t}, x) - W_n(\mathbf{t}, y) - W_n(\mathbf{0}, x) + W_n(\mathbf{0}, y)| \geqq \epsilon] = 0, \tag{40}$$

where $V(a) = \{\mathbf{t}; \|\mathbf{t}\| \leqq a\}$.

PROOF. Given an $\epsilon > 0$ choose an $\epsilon' > 0$ such that (32) is satisfied, the existence of which is guaranteed by Lemma A3. Also choose vectors $\mathbf{t}_j \varepsilon V(a)$, $j = 1, \cdots, r$ such that for every $\mathbf{t} \varepsilon V(a)$ there is a $\mathbf{t}_j$ such that

$$\|\mathbf{t} - \mathbf{t}_j\| \leqq \epsilon'. \tag{41}$$

Then

$$\begin{aligned} &P_n[\sup_{|x-y| \leqq \delta} \sup_{t\varepsilon V(a)} |W_n(\mathbf{t}, x) - W_n(\mathbf{0}, x) - W_n(\mathbf{t}, y) + W_n(\mathbf{0}, y)| \geqq \epsilon] \\ &\quad \leqq \textstyle\sum_{j=1}^r P_n\{\sup_{x,y} \sup_{\|t-t_j\| \leqq \epsilon'} |W_n(t, x) - W_n(t, y) - W_n(t_j, x) \\ &\qquad + W_n(t_j, y)| \geqq \epsilon/2] + \textstyle\sum_{j=1}^r P_n[\sup_{|x-y| \leqq \delta} |W_n(t_j, x) \\ &\qquad - W_n(t_j, y) - W_n(0, x) + W_n(0, y)| \geqq \epsilon/2]. \end{aligned} \tag{42}$$

The first term of (42) goes to zero as $n \to \infty$, by Lemma A3, and the second term goes to zero as $n \to \infty$; then $\delta \to 0$, by (ii), which appears in the proof of Lemma A2. This terminates the proof.

LEMMA A5. *Under the conditions of Theorem* A3, *given an* $\epsilon > 0$, *there is an* $\epsilon' > 0$ *such that for each fixed* x *and* $\mathbf{t}_0$,

$$\lim_{n\to\infty} P_n[\sup_{\|t-t_0\| \leqq \epsilon'} |W_n(t, x) - W_n(t_0, x)| \geqq \epsilon] = 0. \tag{43}$$

PROOF. Note that

$$\begin{aligned} &W_n(t, x) - W_n(t_0, x) \\ &\quad = n^{-\frac{1}{2}} \textstyle\sum_{i=1}^n c_{in}[I(Y_{in} \leqq x + \delta_{in}(t)) - I(Y_{in} \leqq x + \delta_{in}(t_0)) - F(x \\ &\qquad + \delta_{in}(t) + F(x + \delta_{in}(t_0))] \\ &\quad = [J_n(t, x) - J_n(t_0, x)] - [\bar{J}_n(t, x) - \bar{J}_n(t_0, x)]. \end{aligned} \tag{44}$$

However,

$$\sup_{\|t-t_0\|\leqq\epsilon'} |\bar{J}_n(t, x) - \bar{J}_n(t_0, x)|$$
$$\leqq \sup_{\|t-t_0\|\leqq\epsilon'} n^{-\frac{1}{2}}\sum |c_{in}|\, |F(x + \delta_{in}(t) - F(x + \delta_{in}(t_0))|$$
$$\leqq \epsilon' n^{-1}\sum_{i=1}^n |c_{in}|\, \|\mathbf{X}_{in}\| \sup_x f(x + \delta_{in}(t_0))$$
$$\leqq k_2\epsilon' \quad \text{for large} \quad n.$$

Thus, if we choose ϵ' such that $k_2\epsilon' \leqq \epsilon/2$ and such that

$$\lim_{n\to\infty} P_n[\sup_{\|t-t_0\|\leqq\epsilon'} |J_n(t, x) - J_n(t_0, x)| \geqq \epsilon] = 0,$$

then (43) will be proved.

Again assume $c_{in} \geqq 0$; then

$$\sup_{\|t-t_0\|\leqq\epsilon'} [J_n(t, x) - J_n(t_0, x)]$$
$$\leqq n^{-\frac{1}{2}}\sum c_{in}[I(Y_{in} - \delta_{in}(t_0) \leqq x + \lambda_{in}(\epsilon')) - I(Y_{in} - \delta_{in}(t_0) \leqq x)]$$
$$= n^{-\frac{1}{2}}\sum c_{in}I(x < Y_{in} - \delta_{in}(t_0) \leqq x + \lambda_{in}(\epsilon'))$$
$$= M_n^*(x) + n^{-\frac{1}{2}}\sum c_{in}[F(x + \lambda_{in}(\epsilon') + \delta_{in}(t_0)) - F(x + \delta_{in}(t_0))].$$

where

$$M_n^*(x) = n^{-\frac{1}{2}}\sum c_{in}\{I(x + \delta_{in}(t_0), x + \delta_{in}(t_0) + \lambda_{in}(\epsilon')]$$
$$- F(x + \delta_{in}(t_0) + \lambda_{in}(\epsilon')) + F(x + \delta_{in}(t_0)\}.$$

Now it is easy to verify that $\lim_{n\to\infty} \operatorname{Var}(M_n^*(x)) = 0$ for each x. Moreover, for ϵ' given above, the second term is at most $\epsilon/2$. The lower bound on $[J_n(t, x) - J_n(t_0, x)]$ may be handled similarly. Again, for $\{c_{in}\}$ with variable sign, the decomposition into negative and positive parts is used. The proof is terminated.

LEMMA A6. *Under the conditions of Theorem* A3 *for every* $\epsilon > 0$ *and each* x *fixed,*

$$\lim_{n\to\infty} P_n[\sup_{t\varepsilon V(a)} |W_n(t, x) - W_n(0, x)| \geqq \epsilon] = 0 \tag{45}$$

for any $0 < a < \infty$.

PROOF. For any $\epsilon > 0$ choose an $\epsilon' > 0$ such that (43) is satisfied. Choose $\mathbf{t}_1, \cdots, \mathbf{t}_r \varepsilon V(a)$ such that for every $\mathbf{t} \varepsilon V(a)$ there is a $\mathbf{t}_j$ such that $\|\mathbf{t} - \mathbf{t}_j\| \leqq \epsilon'$. Then,

$$P_n[\sup_{t\varepsilon V(a)} \sup|W_n(t, x) - W_n(0, x)| \geqq \epsilon]$$
$$\leqq \sum_{j=1}^r P_n[\sup_{\|t-t_j\|\leqq\epsilon'} |W_n(t, x) - W_n(t_j, x)| \geqq \epsilon/2] \tag{46}$$
$$+ \sum_{j=1}^r P_n[|W_n(t_j, x) - W_n(0, x)| \geqq \epsilon/2].$$

The first term of the right hand side tends to zero because of Lemma A5 and the second term tends to zero as $n \to \infty$ because of (26). Hence the lemma.

Finally, we are in a position to state one of our main theorems.

THEOREM A4. *Let* F *be absolutely continuous with absolutely continuous and uni-*

formly bounded density f. Also let $\{x_{i\nu}\}$ *satisfy* (1.2) *and* (1.3). *Then for every* $\epsilon > 0$

$$\lim_{n\to\infty} P_n[\sup_{-\infty \leqq x \leqq +\infty} \sup_{t\varepsilon V(a)} |L_n(\mathbf{t}, x) - L_n(0, x)| \geqq \epsilon] = 0. \tag{47}$$

As a consequence

$$\lim_{n\to\infty} \mathfrak{L}(\sup_{t\varepsilon V_n(a)} L_n(\mathbf{t}, x), -\infty \leqq x \leqq +\infty) = \mathfrak{L}(L'(x), -\infty \leqq x \leqq +\infty); \tag{48}$$

$$\lim \mathfrak{L}(\sup_x \sup_{t\varepsilon V_n(a)} L_n(\mathbf{t}, x)) = \mathfrak{L}' \tag{49}$$

for any $0 < a < \infty$.
Here $\mathfrak{L}'$ *is law essentially determined by a Gaussian process concentrated on continuous functions and* $\mathfrak{L}'$ *is a process determined by a Gaussian process with almost all sample paths continuous.*

PROOF. In the W_n process we now take

$$c_{in} = x_{i\nu}, \qquad 1 \leqq i \leqq n, \tag{50}$$

and then the relationship

$$\begin{aligned} L_n(\mathbf{t}n^{-\frac{1}{2}}, x) &= W_n(\mathbf{t}, x) - 2W_n(\mathbf{t}, 0) \quad &\text{if } x \geqq 0, \\ &= -W_n(\mathbf{t}, x) &\text{if } x < 0 \end{aligned} \tag{51}$$

holds with probability 1.

The claimed result (47) will follow if we show that

$$\sup_{\|t\| \leqq a} \sup_{|x-y| \leqq \delta} |L_n(tn^{-\frac{1}{2}}, x) - L_n(tn^{-\frac{1}{2}}, y) - L_n(0, x) + L_n(0, y)| \to 0$$

in P_n probability and that the finite dimensional distribution of $\{|L_n(\mathbf{t}n^{-\frac{1}{2}}, x) - L_n(\mathbf{0}, x)|, -\infty \leqq x \leqq +\infty\}$ converges to degenerate distribution—degenerate at zero, uniformly in all $t \varepsilon V(a)$. But the latter fact follows from Lemmas A5 and A6 and the relations (50) and (51), while the former one follows from (50) and lemmas A1–A4. (48) and (49) are straight consequences of (47), and $\{L_n(\mathbf{0}, x), -\infty \leqq x \leqq +\infty\}$, $n \geqq 1$, is relatively compact with a continuous Gaussian limit, which follows again from (50) and (51) and Theorem A3 with $\mathbf{t} = \mathbf{0}$. The proof is completed.

Our next problem is to obtain a version of the above theorem for $Z_n(t, |x|)$ processes.

THEOREM A5. *Under the conditions of Theorem* A3 *for any* $\epsilon > 0$

$$\lim_{n\to\infty} P_n[\sup_{-\infty \leqq x \leqq +\infty} \sup_{t\varepsilon V_n(a)} |Z_n(t, |x|) - Z_n(0, |x|)| \geqq \epsilon] = 0. \tag{52}$$

Consequently

$$\lim_{n\to\infty} \mathfrak{L}(\sup_x \sup_{t\varepsilon V_n(a)} (Z_n(t, |x|)) = \mathfrak{L}_1, \tag{53}$$

$$\lim_{n\to\infty} \mathfrak{L}(\sup_{t\varepsilon V_n(a)} Z_n(t, |x|), -\infty \leqq x \leqq +\infty) = \mathfrak{L}(Z(x); -\infty \leqq x \leqq +\infty) \tag{54}$$

where $\mathfrak{L}_1$ is determined by a Gaussian process with continuous sample functions and Z is essentially a Guassian process with continuous sample paths.

PROOF. Again we will use the results about $W_n(\mathbf{t}, x)$ processes. Note that if

$$c_{in} = 1, \qquad 1 \leqq i \leqq n, \tag{55}$$

then

$$\begin{aligned} Z_n(\mathbf{t}n^{-\frac{1}{2}}, |x|) &= n^{-\frac{1}{2}}\sum \{I[|Y_{in} - \delta_{in}(\mathbf{t})| \leqq |x|] - EI[|Y_{in} - \delta_{in}(t)| \leqq |x|]\}, \\ &= W_n(\mathbf{t}, x) - W_n(\mathbf{t}, -x) \quad \text{if} \quad x \geqq 0, \\ &= W_n(\mathbf{t}, -x) - W_n(\mathbf{t}, x) \quad \text{if} \quad x < 0, \\ &= W_n(\mathbf{t}, |x|) - W_n(\mathbf{t}, -|x|) \qquad -\infty < x < +\infty \end{aligned} \tag{56}$$

holds with probability 1.

Therefore

$$\begin{aligned} &Z_n(\mathbf{t}n^{-\frac{1}{2}}, |x|) - Z_n(\mathbf{0}, |x|) \\ &\quad = [W_n(\mathbf{t}, |x|) - W_n(\mathbf{t}, -|x|) - W_n(\mathbf{0}, |x|) + W_n(\mathbf{0}, -|x|)]. \end{aligned} \tag{57}$$

And it follows that

$$\begin{aligned} &\sup_{t \varepsilon V(a)} |Z_n(\mathbf{t}n^{-\frac{1}{2}}, |x|) - Z_n(\mathbf{0}, |x|)| \\ &\quad \leqq \sup_{t \varepsilon V(a)} |W_n(\mathbf{t}, |x|) - W_n(\mathbf{0}, |x|)| + \sup_{t \varepsilon V(a)} |W_n(\mathbf{t}, -|x|) \\ &\qquad - W_n(\mathbf{0}, -|x|)| \\ &\quad \rightarrow 0 \quad \text{in} \quad P_n\text{-probability} \end{aligned} \tag{58}$$

by Lemma A6. This tells us that finite dimensional distribution of the processes $\{|Z_n(tn^{-\frac{1}{2}}, |x|) - Z_n(0, |x|)|\}$, $n \geqq 1$ is degenerate at 0 in the limit, uniformly in $t \varepsilon V(a)$.

Next we similarly observe that

$$\begin{aligned} &\sup |Z_n(\mathbf{t}n^{-\frac{1}{2}}, |x|) - Z_n(\mathbf{0}, |x|) - Z_n(\mathbf{t}n^{-\frac{1}{2}}, |y|) + Z_n(\mathbf{0}, |y|)| \\ &\quad \leqq \sup |W_n(\mathbf{t}, |x|) - W_n(\mathbf{t}, +|y|) - W_n(\mathbf{0}, |x|) + W_n(\mathbf{0}, |y|)| \\ &\qquad + \sup |W_n(\mathbf{t}, -|x|) - W_n(\mathbf{t}, -|y|) - W_n(\mathbf{0}, -|x|) + W_n(\mathbf{0}, -|y|)| \end{aligned} \tag{59}$$

where sup on either side is taken over all x, y; $|x - y| \leqq \delta$ and all $\mathbf{t} \varepsilon V(a)$. But this tends to zero in P_n-probability as $n \rightarrow \infty$ and then $\delta \rightarrow 0$ by Lemma A4. Thus, fluctuations of our processes are small, uniformly in $t \varepsilon V(a)$, in probability for large n. Hence (52) (53) and (54) follow from this fact and the fact that $\{Z_n(0, |x|) - \infty \leqq x \leqq +\infty\}$ $n \geqq 1$ is a relatively compact sequence with a continuous Gaussian limit.

From the above theorem and the well known Glivenko-Cantelli theorem for empirical cumulatives, the following corollary may be concluded.

COROLLARY A5.1. *For every* $\epsilon > 0$ *and any* $0 < a < \infty$

$$\lim_{n\to\infty} P_n[\sup_x \sup_{t\varepsilon V_n(a)} n^{-\frac{1}{2}}|Z_n(t, |x|)| \geqq \epsilon] = 0. \tag{60}$$

But $n^{-\frac{1}{2}}Z_n(t, |x|) = H_n(t, |x|) - \bar{H}_n(t, |x|)$.

Define for $0 \leqq y \leqq 1$,

$$H_n^{-1}(t, y) = \inf\{x \geqq 0;\ H_n(t, x) \geqq y\}, \tag{61}$$

$$\bar{H}_n^{-1}(t, y) = \inf\{x \geqq 0;\ \bar{H}_n(t, x) \geqq y\}.$$

Because of Corollary A5.1 we have

COROLLARY A5.2. *For every* $\epsilon > 0$, *and an* a, $0 < a < \infty$,

$$\lim P_n[\sup_{0\leqq y\leqq 1} \sup_{t\varepsilon V_n(a)} |H_n^{-1}(t, y) - \bar{H}_n^{-1}(t, y)| \geqq \epsilon] = 0. \tag{62}$$

We now state and prove our last

THEOREM A6. *Let the conditions of Theorem* A3 *and Lemma* A1 *be satisfied. Then for every* $\epsilon > 0$ *and an* a, $0 < a < \infty$,

$$\lim_{n\to\infty} P_n[\sup_{0\leqq x\leqq 1} \sup_{t\varepsilon V_n(a)} |L_n(t, H_n^{-1}(t, x)) - L_n(t, \bar{H}_n^{-1}(t, x))| \geqq \epsilon] = 0.$$

PROOF. From Theorem A4 it follows that for any $\eta > 0$ and $\epsilon > 0$ there is an $n_{\epsilon,\eta} = n_0$ and a $\delta_0 = \delta_{\epsilon,\eta}$ such that $n \geqq n_0$ implies

$$P_n[\sup_t \sup_{|x-y|\leqq\delta_0} |L_n(t, x) - L_n(t, y)| \geqq \epsilon] \leqq \eta/2.$$

Similarly from (62) it follows that there exists $n_1 = n'_{\epsilon,\eta}$ such that $n \geqq n_1$ implies

$$P_n[\sup_t \sup_{0\leqq y\leqq 1} |H_n^{-1}(t, y) - \bar{H}_n^{-1}(t, y)| \geqq \delta_0] \leqq \eta/2.$$

Let

$$A_n = [\sup_t \sup_{0\leqq y\leqq 1} |H_n^{-1}(t, y) - \bar{H}_n^{-1}(t, y)| \leqq \delta_0 \quad \text{and} \quad \sup_t \sup_{|x-z|\leqq\delta} |L_n(t, x) - L_n(t, z)| \geqq \epsilon].$$

Then for $n \geqq \max(n_1, n_2)$

$$P_n[\sup_t \sup_{0\leqq y\leqq 1} |L_n(t, H_n^{-1}(t, y)) - L_n(t, \bar{H}_n^{-1}(t, y))| \leqq \epsilon] \geqq P_n[A_n] \geqq 1 - \eta.$$

Note that $\sup_t$ is over $t \varepsilon V_n(a)$. The proof is terminated.

REMARK. All through the above discussion we assumed that our random variable $\{Y_{in}\}$ are defined on $[-\infty, +\infty]$. This does not change anything in our basic problem for $P[Y_{in} = \pm\infty] = 0$.

Acknowledgments. I wish to express my deeply felt gratitude to Professor Peter J. Bickel for suggesting the problem and for his continued guidance and encouragement thereafter. To Professor Lucien M. LeCam I am grateful for various discussions and suggestions. I would also like to thank the referee for his very valuable criticism and helpful comments.

REFERENCES

[0] APOSTOL. (1957) *Mathematical Analysis*. Addison-Wesley, Reading, Mass.

[1] ADICHIE, J. N. (1967). Asymptotic efficiency of a class of nonparametric tests for regression scores. *Ann. Math. Statist.* **38** 884–893.

[2] ADICHIE, J. N. (1967). Estimation of regression parameters based on rank tests. *Ann. Math. Statist.* **38** 894–904.

[3] BILLINGSLEY, P. (1968) *Weak Convergence of Stochastic Processes*. Wiley, New York.

[4] CRAMÉR, H. (1945). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[5] CRAMÉR, H. (1963). *Random Variables and Probabilities Distributions*. (2nd ed.). Cambridge Univ. Press.

[6] EICKER, F. (1965). Limit theorems for regressions with unequal and dependent errors. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** Univ. of California Press.

[7] HÁJEK, J. (1962). Asymptotically most powerful rank order tests. *Ann. Math. Statist.* **33** 1124–1147.

[8] HODGES, J. L. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 599–611.

[9] LECAM, L., GOVINDARAJULU, Z., and RAGHAVACHARI, M. (1965). Generalizations of theorems of Chernoff-Savage on the asymptotic normality of test statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** Univ. of California Press.

[10] RAO, C. R. (1966). *Linear Statistical Inference and its Applications*. Wiley, New York.