ABSTRACT

ASYMPTOTIC NORMALITY OF SIMPLE LINEAR RANDOM RANK STATISTICS UNDER THE ALTERNATIVES

By

Janet Tolson Eyster

In this paper we study the problem of the asymptotic normality of random signed rank statistics under the alternatives when the score functions, φ, are bounded. When the random variables are independent and identically distributed, φ is assumed to be square integrable. This extends the work of Koul (1970) and (1972), Sen-Ghosh (1971) and Ghosh-Sen (1972) to φ which may be discontinuous. To relax the assumptions of differentiability on φ, restrictions are placed on the distribution functions of the random variables similar to those used by Dupač-Hájek (1969). This work is useful in generating bounded length confidence intervals for the regression problems.

ASYMPTOTIC NORMALITY OF SIMPLE LINEAR RANDOM RANK STATISTICS UNDER THE ALTERNATIVES

By

Janet Tolson Eyster

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1977

TO

George, Mark and Sandy

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to Professor Hira Koul for his guidance and patience during the preparation of this dissertation. His advice and encouragement given during my study at Michigan State University are greatly appreciated.

I also wish to thank Professors D. Gilliland, R. Phillips and J. Stapleton for reading this thesis and for their helpful comments. Special thanks are due to Mrs. Noralee Burkhardt for her typing of the manuscript and the patience with which she did it.

Finally, I wish to express my gratitude to the National Institutes of Health and to the Department of Statistics and Probability, Michigan State University for their financial support during my stay at Michigan State University.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

We consider the problem of the asymptotic normality of simple linear signed rank statistics $S^+(\varphi)$ under the alternatives based on a random number of observations. They are called random signed rank statistics. Corresponding theorems are also presented for random rank statistics. In particular suppose $\{X_i\}$, $i \geq 1$, is a sequence of independent random variables with continuous distribution functions $\{F_i\}$, $i \geq 1$, and $\{N_r\}$, $r \geq 1$, is a sequence of positive integer valued random variables. All random variables are assumed to be defined on the same probability space. Let $\{c_i\}$, $i \geq 1$, be a sequence of real numbers. We will be investigating the asymptotic distribution of random signed rank and random rank statistics which correspond to a φ and are defined as

$$S_n^+(\varphi) = \Sigma_1^n c_i a_n(R_{in}^+) s(X_i) \quad \text{and}$$

$$S_n(\varphi) = \Sigma_1^n c_i a_n(R_{in})$$

where

(1.1) $$s(x) = I(x \geq 0) - I(x \leq 0), \quad -\infty < x < \infty$$

(1.2) $$R_{in}^+ = \Sigma_{j=1}^n I(|X_j| \leq |X_i|),$$

(1.3) $$R_{in} = \Sigma_{j=1}^{n} I(X_j \leq X_i), \text{ and}$$

(1.4) $$a_n(i) = \varphi(i/(n+1)) \quad \text{or}$$

(1.5) $$a_n(i) = E(\varphi(U_{ni})), \ i = 1,\ldots,n$$

where $U_{n1} \leq \cdots \leq U_{nn}$ is the order statistic of a sample of size n from a uniform distribution on $(0,1)$. When X_i, $i \geq 1$, have continuous distribution functions F_i, $i \geq 1$, φ is assumed to be bounded and and to satisfy some other conditions presented later. These φ may be discontinuous. When X_i, $i \geq 1$, are independent and identically distributed (i.i.d.), φ is assumed to be square integrable and

Also presented are theorems related to the asymptotic normality of $S(\varphi)$ for non-random sample sizes and an application of the theorems for random sample size which are of interest by themselves.

This work is motivated by a desire to use φ which are discontinuous to generate bounded length confidence intervals for the regression problem as presented by Koul (1970) and (1972) and Ghosh-Sen (1972). In particular the median test can be used since $\varphi(t) = I(\frac{1}{2} < t \leq 1)$ is bounded and discontinuous at $t = \frac{1}{2}$.

This work should be compared to the papers by Pyke-Shorack (1968a) and (1968b), Koul (1970) and (1972), Sen-Ghosh (1971) and (1972) and Ghosh-Sen (1972).

Pyke-Shorack (1968a) and (1968b) presented a Chernoff-Savage theorem on the asymptotic normality of the two sample linear rank statistic for non-random and random sample sizes with a broad class of φ which include φ with jumps. With slightly more restrictive score functions, Braun (1976) shows that the two sample linear rank statistic with random sample sizes converges weakly to a Wiener process. Koul (1970) proved the asymptotic normality of the random signed rank statistics when $\{a_n(i)\}$ satisfy (1.4), $\{c_i\}$ satisfy Noether's condition, $\{X_i\}$ are i.i.d. and φ is square integrable and absolutely continuous with derivative φ' which exists a.e. and has bounded variation on $[0,1]$. Koul (1972) proved the asymptotic normality of the random rank and signed rank statistics under the alternatives when a_n satisfies (1.4) and φ satisfies the conditions in Koul (1970) and has a continuous derivative φ'. These φ are bounded and uniformly continuous.

Sen-Ghosh (1971) presented the asymptotic normality of functions of random signed rank statistics when $c_i \equiv 1$, $\{a_n(i)\}$ satisfy either (1.4) or (1.5), φ satisfy various smoothness conditions and $\{X_i\}$, $i \geq 1$, is a sequence of i.i.d. random variables having an absolutely continuous distribution function F symmetric about zero for which both the density function f and its first derivative f' exist and are bounded for almost all x. These φ include the normal scores and Wilcoxon, but not the median tests.

With essentially the same assumptions on φ, $\{a_n(i)\}$ and F, except symmetry which is not assumed, Ghosh-Sen (1972) generalized to the linear regression model with $\{c_i\}$ satisfies the boundedness condition (2.2).

The results in Sen-Ghosh (1971) are used in obtaining bounded length confidence intervals for the median while the results in Koul (1972) and Ghosh-Sen (1972) are useful in obtaining sequential bounded length confidence intervals for the regression coefficient β.

In two of our main theorems we establish the asymptotic normality of random signed rank and rank statistics under the alternatives for a class of bounded φ which may be discontinuous.

In particular, in Theorems 2.1 and 2.2 we assume that φ is the difference of two non-decreasing bounded functions. To relax the usual assumptions of differentiability on φ, we need to place restrictions on F_i similar to those used by Dupač-Hájek (1969). In view of the example in Hájek (1968), these conditions are close to being necessary. As an application of these theorems, we present the regression problem.

To complete the proof of the main theorems, the appendix contains the fixed sample size theorems. They are presented in a form which relaxes the boundedness condition on φ by assuming φ satisfies Hoeffding's (1973) condition. By imposing Hoeffding's condition on φ we are able to eliminate condition D-H(2.12) and relax condition D-H(2.13) on F_i in Dupač-Hájek (1969).

Theorem 2.3 presents the weak convergence of random signed rank statistics when X_i, $i \geq 1$, are i.i.d., a_n is defined by (1.5) and φ is square integrable. The proof uses techniques similar to Sen-Ghosh (1972). The essential step in proving the uniform continuity in probability needed to apply the Anscombe theorem (1952) is the proof that if $\mathcal{F}_n$ is the σ-field generated by $\{(s(X_i), R_i), 1 \leq i \leq n\}$ and the distribution F of $\{X_i\}$ is symmetric about zero, then $\{S_n^+, \mathcal{F}_n\}$ is a martingale sequence and the Kolmogorov inequality for martingales holds. Sen-Ghosh (1971) and Ghosh-Sen (1972) used this technique to prove their results.

In general this technique cannot be used. If F is not symmetric about 0 or if a_n is defined by (1.4) instead of (1.5), the martingale property need not hold for S_n^+. Similarly if a_n is defined by (1.4), the martingale property need not hold for S_n. Since S_n^+ and S_n are sums of dependent random variables, the uniform continuity in probability needed to apply the Anscombe theorem is in general difficult to prove.

To prove the main theorems for a_n defined by (1.4), we utilize the so called weak convergence technique as used by earlier authors, Pyke-Shorack (1968a) and (1968b) and Koul (1970) and (1972). We decompose the normalized random signed rank or rank statistic into two parts and approximate each part by a fixed sample size sum of independent random variables.

Chapter 2 contains the main theorems. In Chapter 3 Theorems 2.1 and 2.2 are specialized to the regression problem. Also some sufficient conditions for Theorems 2.1 and 2.2 are presented. Chapter 4 contains preliminary results for the empirical process which are necessary for the proofs that follow. Chapters 5 and 6 contain the proofs of the main theorems. The asymptotic normality of the non-random signed rank and rank statistics is presented in the Appendix.

CHAPTER 2

MAIN THEOREMS

Suppose $\{X_i\}$, $\{F_i\}$, $\{N_r\}$ and $\{c_i\}$ are defined as before. About $\{N_r\}$ we assume the following:

(2.1) there exist sequences of positive integers $\{a_r\}$ and $\{b_r\}$ such that $a_r \to \infty$, $b_r \to \infty$, $a_r/b_r \to 1$ and if $A_r = [a_r \le N_r \le b_r]$, then $P[A_r] \to 1$; or equivalently there exists an increasing sequence of positive integers $\{n_r\}$ such that $N_r/n_r \to 1$ in probability as r tends to infinity.

For Theorems 2.1 and 2.2 we will assume that $\{c_i\}$ satisfy the boundedness condition for all n

(2.2) $$k_n \le k_c < \infty \quad \text{where} \quad \sigma_n^2 = \Sigma_1^n c_i^2 \quad \text{and} \quad \sigma_n^{-1} n^{\frac{1}{2}} \max_{1 \le i \le n} |c_i| = k_n .$$

Since $1 \ge \sigma_{a_r}^2 \sigma_{b_r}^{-2} \ge 1 - (b_r - a_r) \max\limits_{1 \le i \le b_r} |c_i^2| \sigma_{b_r}^{-2}$, (2.1) and (2.2) imply

(2.3) $$\sigma_{a_r}/\sigma_{b_r} \to 1 .$$

(2.1) and (2.3) imply

(2.4) $$\sigma_{N_r}/\sigma_{b_r} \to 1 \quad \text{w.p. } 1 \quad \text{and}$$

$$\sigma_{a_r}/\sigma_{N_r} \to 1 \quad \text{w.p. } 1 .$$

For Theorem 2.3 we will only assume that $\{c_i\}$ satisfy (2.3) and Noether's condition

(2.5) $$\max_{1 \le i \le n} |c_i| \sigma_n^{-1} \to 0 \quad \text{as} \quad n \to \infty.$$

Define for any n

$$S_n^+(t) = \Sigma_1^n \, c_i I(R_i^+ \le (n+1)t) s(X_i), \quad 0 \le t \le 1,$$

and

$$S_n(t) = \Sigma_1^n \, c_i I(R_i \le (n+1)t), \quad 0 \le t \le 1,$$

where $s(X_i)$, R_i^+ and R_i are defined by (1.1), (1.2) and (1.3) respectively. One observes that

$$S_n^+(\varphi) = \int_0^1 \varphi(t)\, dS_n^+(t)$$

and

$$S_n(\varphi) = \int_0^1 \psi(t)\, dS_n(t) \ .$$

This representation makes it clear that one needs to study the standardized weighted empirical processes $\{S_{N_r}^+(t), 0 \le t \le 1\}$ and $\{S_{N_r}(t), 0 \le t \le 1\}$ in the cases in which φ is bounded. This is done in Chapter 4.

In the remainder of this chapter we will state our main theorems for random signed rank and random rank statistics.

In Theorems 2.1 and 2.2 we assume

(2.6) $$\varphi = \varphi_1 - \varphi_2$$

where φ_m is $\searrow$ bounded on $(0,1)$, $m = 1,2$.

For signed rank statistics we define for $0 \le x < +\infty$ and $0 \le t \le 1$,

$$K_n(x) = n^{-1}\Sigma_1^n I\{|X_i| \le x\},$$

$$\underline{K}_n(x) = n^{-1}\Sigma_1^n[F_i(x) - F_i(-x)],$$

(2.7) $$L_{in1}^+(t) = F_i(\underline{K}_n^{-1}(t)) - F_i(0),$$

$$L_{in2}^+(t) = F_i(0) - F_i(-\underline{K}_n^{-1}(t)),$$

$$L_{in}^+(t) = L_{in1}^+(t) + L_{in2}^+(t),$$

$$\mu_{in}^+(t) = L_{in1}^+(t) - L_{in2}^+(t),$$

$$\bar{\mu}_n^+(t) = \Sigma_1^n c_i\mu_{in}^+(t), \text{ and}$$

$$\bar{\mu}_n^+(\varphi) = \Sigma_1^n c_i\int\varphi(t)d\mu_{in}^+(t).$$

Obviously for any n

(2.8) $$n^{-1} \Sigma_1^n L_{in}^+(t) \equiv t .$$

<u>Theorem 2.1</u>. <u>Let</u> $\{a_n(i)\}$, $\{N_r\}$, $\{c_i\}$, φ <u>and</u> φ_m, $m = 1,2$, <u>satisfy</u> (1.4), (2.1), (2.2) <u>and</u> (2.6). <u>Assume there exist functions</u> $\ell_{ijk}^+(s)$, $0 < s < 1$, $1 \le i \le j$, $a_r \le j \le b_r$, $k = 1,2$, <u>and sets</u> E_r <u>such that</u> $\forall\, d$, $0 < d < \infty$, <u>and</u> $\forall\, \eta > 0$

(2.9) $$\lim_{r\to\infty} \varphi_m(E_r) = 0, \ m = 1,2,$$

<u>where</u> $E_r = E_{r1} \cup E_{r2}$

$$(2.10)\qquad E_{r1} = \{s;\ \max_{k=1,2}\ \max_{a_r \leq j \leq b_r}\ \max_{1 \leq i \leq j}\ \sup_{|t-s| \leq dj^{-\frac{1}{2}}} j^{\frac{1}{2}} |L^{+}_{ijk}(t) - L^{+}_{ijk}(s) - (t-s)\ell^{+}_{ijk}(s)| > \eta\},$$

$$(2.11)\qquad E_{r2} = \{s;\ \sigma^{-1}_{a_r} a_r^{\frac{1}{2}} \max_{a_r \leq j \leq b_r} |\hat{c}^{+}_{j}(s) - \hat{c}^{+}_{a_r}(s)| > \eta\},$$

and for any integer n

$$(2.12)\qquad \hat{c}^{+}_{n}(t) = n^{-1} \Sigma_{1}^{n} c_i(\ell^{+}_{in1}(t) - \ell^{+}_{in2}(t)),\ 0 < t < 1.$$

Finally for any n define

$$(2.13)\qquad \begin{aligned} \tau^{2}_{n+}(\varphi) = \Sigma_{i=1}^{n} \mathrm{Var}\{\textstyle\int[&(c_i - \hat{c}^{+}_{n})[I(0 \leq X_i \leq K_n^{-1}(t)) - L_{in1}(t)] \\ &- (c_i + \hat{c}^{+}_{n})(I(-K_n^{-1}(t) \leq X_i < 0) - L_{in2}(t))]d\varphi(t) \\ &- c_i[s(X_i) - \mu^{+}_{i}(1)]\varphi(1)\}\ . \end{aligned}$$

Then

$$(2.14)\qquad \liminf_{r}\ \tau^{2}_{a_r+}(\varphi)\sigma^{-2}_{a_r} > 0$$

implies

$$(2.15)\qquad \tau^{-1}_{a_r+}(\varphi)[S^{+}_{N_r}(\varphi) - \bar{\mu}^{+}_{N_r}(\varphi)] \xrightarrow[D]{} N(0,1) \quad \text{r.v.}$$

Proof: See Chapter 5.

For linear rank statistics we define for $-\infty < x < \infty$, $0 \leq t \leq 1$

$$H_n(x) = n^{-1} \Sigma_1^n I\{X_i \le x\},$$

$$\underline{H}_n(x) = n^{-1} \Sigma_1^n F_i(x),$$

(2.16) $$L_{in}(t) = F_i(\underline{H}_n^{-1}(t)),\ 1 \le i \le n \quad \text{and}$$

$$\mu_n(\varphi) = \Sigma_1^n c_i \int \varphi(t)\, dL_{in}(t).$$

Theorem 2.2. *Assume* $\{a_n(i)\}$, $\{N_r\}$, $\{c_i\}$, φ *and* φ_m, $m = 1,2$ *satisfy* (1.4), (2.1), (2.2) *and* (2.6) *respectively. Assume there exist functions* $\ell_{ij}(s)$, $0 < s < 1$, $1 \le i \le j$, $a_r \le j \le b_r$, *and sets* D_r *such that* $\forall\, d$, $0 < d < \infty$, *and* $\forall\, \eta > 0$

(2.17) $$\lim_{r\to\infty} \varphi_m(D_r) = 0,\ m = 1,2,$$

where $D_r = D_{r1} \cup D_{r2}$,

$$D_{r1} = \{s;\ \max_{a_r\le j\le b_r}\ \max_{1\le i\le j}\ \sup_{|t-s|\le dj^{-\frac{1}{2}}} j^{\frac{1}{2}} |L_{ij}(t) - L_{ij}(s) - (t-s)\ell_{ij}(s)| > \eta\},$$

(2.18) $$D_{r2} = \{s;\ \sigma_{a_r}^{-1} a_r^{\frac{1}{2}} \max_{a_r\le j\le b_r} |\hat{c}_j(s) - \hat{c}_{a_r}(s)| > \eta\}$$

and for any n

$$\hat{c}_n(t) = n^{-1} \Sigma_1^n c_i \ell_{in}(t).$$

Define for any n

$$(2.19)\quad \tau_n^2(\varphi) = \Sigma_{i=1}^n \mathrm{Var}[\int (c_i - \hat{c}_n)(I(X_i \le H_n^{-1}(t)) - L_{in}(t))d\varphi(t)].$$

<u>Then</u> $\liminf_r \tau_{a_r}^2(\varphi)\sigma_{a_r}^{-2} > 0$ <u>implies</u>

$$\tau_{a_r}^{-1}(\varphi)(S_{N_r}(\varphi) - \mu_{N_r}(\varphi)) \xrightarrow[D]{} N(0,1) \quad \text{r.v.}$$

<u>Proof:</u> See Chapter 5.

<u>Remark 1</u>. Without loss of generality we may assume that

$$(2.20)\qquad n^{-1}\Sigma_{k=1}^2 \Sigma_{i=1}^n \ell_{ink}^+(t) = 1,\ 0 < t < 1$$

and that $\ell_{ink}^+(t)$ are measurable functions on $(0,1)$. These follow from the fact that if there exist ℓ_{ink}^+ satisfying (2.9) then it is satisfied by $\ell_{ink}^+(t) = n^{\frac{1}{2}}[L_{ink}^+(t + n^{-\frac{1}{2}}) - L_{ink}^+(t)]$.

<u>Remark 2</u>. When $\{X_i\}$ are i.i.d. and F symmetric about zero, $\ell_{ijk}^+(t) \equiv \frac{1}{2}$ (in i,j,k and $0 \le t \le 1$) and for any n, $\hat{c}_n(t) \equiv 0$. Therefore $E_r = \phi$ and (2.9) is automatically satisfied.

<u>Remark 3</u>. The theorems above allow a broader class of φ than Koul (1972) by removing the absolute continuity and uniform continuity conditions on φ, but we impose the conditions (2.9) through (2.11) on the F_i.

<u>Remark 4</u>. In the fixed sample size problem when φ is an indicator function (2.17) and (2.9) reduce to D-H (2.13) in Dupač-Hájek (1969) and a similar condition in Koul-Staudte (1972a).

The following corollary allows the application of Theorem 2.2 to the two sample Chernoff-Savage problem.

Define λ_0 such that $0 < \Lambda \le \lambda_0 \le 1 - \Lambda < 1$.

<u>Corollary 2.1</u>. <u>Theorems 2.1 and 2.2 hold when (2.2) is replaced by</u>

(2.21) $$c_i = I(i \le n_r)$$

<u>where</u> n_r <u>is a random variable and there exist positive integers</u> $\{a'_r\}$ <u>and</u> $\{b'_r\}$ <u>such that</u> $a'_r \to \infty$, $b'_r \to \infty$, $a'_r/b'_r \to 1$ <u>and if</u> $A'_r = [a'_r \le n_r \le b'_r]$, <u>then</u> $P[A'_r] \to 1$ <u>and</u> $\lim_r a'_r/a_r = \lambda_0$.

<u>Proof</u>: See Chapter 5.

We now present a theorem for random signed rank statistics with a different centering constant when φ is square integrable and $\{X_i\}$ are independent and identically distributed.

<u>Theorem 2.3</u>. <u>Assume</u> $X_1,\ldots,X_n$ <u>have the continuous distribution function</u> F <u>satisfying</u> $F(x) = 1 - F(-x)$. <u>Assume</u> φ <u>is square integrable in</u> $(0,1)$. <u>Assume</u> $\{a_n(i)\}$, $\{N_r\}$ <u>and</u> $\{c_i\}$ <u>satisfy</u> (1.5), (2.1), (2.3) <u>and</u> (2.5). <u>Then</u> $(S^+_{N_r} - ES^+_{N_r})(A\sigma_{N_r})^{-1} \to N(0,1)$ r.v. <u>where</u> $A^2 = \int_0^1 \varphi^2(t)dt$.

<u>Proof</u>: See Chapter 6.

CHAPTER 3

APPLICATIONS

The following lemma replaces (2.11) in Theorem 2.1 by a stronger condition on F_i but perhaps an easily verifiable one. Define

$$E_{r3} = \{0 < s < 1;\ \max_{k=1,2}\ \max_{a_r \le j \le b_r}\ \max_{1 \le i \le a_r} |\ell^+_{ijk}(s) - \ell^+_{ia_rk}(s)| > \eta\} .$$

<u>Lemma 3.1</u>. <u>Define</u> $E_{r4} = E_{r1} \cup E_{r3}$. <u>Condition</u> (2.9) <u>is implied by</u>

$$\lim_r \varphi_m(E_{r4}) = 0,\ m = 1,2, \tag{3.1}$$

<u>where the</u> ℓ^+_{ijk} <u>are defined in Theorem</u> 2.1.

<u>Proof</u>: Unless otherwise specified a_r, N_r and b_r will be denoted by r, N and b. To prove the lemma it is sufficient to prove $E_{r3} \supset E_{r2}$.

In the following we use the fact that for all j and r

$$j^{-1} = r^{-1} + (j - r)(rj)^{-1}. \tag{3.2}$$

Since s will be fixed in $(0,1)$ we will simplify notation by not revealing the s. For fixed $s \in (0,1)$ and $r < j \le b$ we have

$$\sigma_r^{-1} r^{\frac{1}{2}} |\hat{c}_j^+ - \hat{c}_r^+|$$

$$\leq \sigma_r^{-1} r^{\frac{1}{2}} |\Sigma_1^r c_i r^{-1}([\ell_{ij1}^+ - \ell_{ij2}^+] - [\ell_{ir1}^+ - \ell_{ir2}^+])$$

$$+ \Sigma_1^j c_i (j - r)(rj)^{-1}(\ell_{ij1}^+ - \ell_{ij2}^+)$$

$$+ \Sigma_{r+1}^j c_i r^{-1}(\ell_{ij1}^+ - \ell_{ij2}^+)| \qquad \text{by (3.2)}$$

$$\leq \sigma_r^{-1} r^{\frac{1}{2}} \max_{1 \leq i \leq b} |c_i| [2 \max_{k=1,2} \max_{r \leq j \leq b} \max_{1 \leq i \leq r} |\ell_{ijk}^+ - \ell_{irk}^+|$$

$$+ r^{-1}(j - r) j^{-1} \Sigma_1^j (\ell_{ij1}^+ + \ell_{ij2}^+)$$

$$+ r^{-1}[j - \Sigma_1^r (\ell_{ij1}^+ + \ell_{ij2}^+)]]$$

$$\leq 8k_c \max_{k=1,2} \max_{r \leq j \leq b} \max_{1 \leq i \leq r} |\ell_{ijk}^+ - \ell_{irk}^+| + o(1) \qquad \text{by (2.2) and (2.20).}$$

This completes the proof that $E_{r3} \supset E_{r2}$.

<u>Remark 5.</u> Results similar to Lemma 3.1 hold for rank statistics. In particular (2.18) may be replaced by

$$D_{r3} = \{0 < s < 1 \quad \max_{a_r \leq j \leq b_r} \max_{1 \leq j \leq a_r} |\ell_{ij}(s) - \ell_{ia_r}(s)| > \eta\} \quad \text{and (2.17)}$$

by $\lim_{r\to\infty} \varphi_m(D_{r1} \cup D_{r3}) = 0$, $m = 1,2$.

Theorems 2.1 and 2.2 can be applied to the regression problem. The proof will be carried out for Theorem 2.1 although similar results hold for Theorem 2.2.

Assume for all n, $X_{in} = \beta_0 + c_{in}\beta + \varepsilon_i$, $c_{in}^2 = c_i^2/\Sigma_1^n c_i^2$, c_i satisfy (2.5) and ε_i have a distribution function F which has a uniformly continuous derivative f which is positive on $(-\infty,\infty)$. Also assume N_r satisfies (2.1). For any n define $\bar{\mu}_n^+(\varphi)$, L_{ink}^+,

$k = 1,2$, $\underline{K}_n$, $\tau^2_{n+}(\varphi)$ and $k_n(t)$ by (2.7), (2.13) and

$$k_n(t) = n^{-1}\Sigma_1^n[f(\underline{K}_n^{-1}(t) - c_{in}\beta) + f(-\underline{K}_n^{-1}(t) - c_{in}\beta)],\ t \in (0,1).$$

Also define $F_*(x) = F(x) - F(-x)$, $0 \le x < \infty$.

Then $\tau^2_{r+}(\varphi)\sigma_r^{-2} > 0$ implies $\tau^{-1}_{r+}(\varphi)(S_N^+(\varphi) - \bar{\mu}_N^+(\varphi)) \to N(0,1)$.

The proof follows from Theorem 2.1 and Lemma 3.1 by showing that (3.1) holds.

Fix $t_0 \in (0,1)$. For every h such that $(t_0 - h, t_0 + h) \subset (0,1)$ and n there exist e_n, g_n (dependent on h, t_0 and n), $-\infty < e_n < g_n < \infty$ such that $e_n = \underline{K}_n^{-1}(t_0 - h) - \max_i c_{in}\beta < \underline{K}_n^{-1}(t_0 + h) - \min_i c_{in}\beta = g_n$. Since $\underline{K}_n^{-1}(t) \to F_*^{-1}(t)$ and $\max_{1\le i\le n} |c_{in}| \to 0$ as $n \to \infty$, there exists e and g, $-\infty < e < g < \infty$ such that for all n, $e \le e_n < g_n \le g$.

Since f is positive on $(-\infty,\infty)$ and e and g are finite there exists $\varepsilon > 0$ dependent on t_0 and h such that for all $n \ge 1$

$$\inf_{|t-t_0|<h} k_n(t) > \varepsilon.$$

Noting that $[k_n(t)]^{-1}$ represents the derivative of $\underline{K}_n^{-1}(t)$, we obtain for all $n \ge 1$

$$\sup_{|t-t_0|<h} |d\underline{K}_n^{-1}(t)/dt| < \varepsilon^{-1}.$$

Therefore $\underline{K}_n^{-1}(t)$, $n \ge 1$, is uniformly continuous in a neighborhood of t_0. The same is true of

$$dL^{+}_{in1}(t)/dt = f(\underline{K}_n^{-1}(t) - c_{in}\beta)/k_n(t)$$

and

$$dL^{+}_{in2}(t)/dt = f(-\underline{K}_n^{-1}(t) - c_{in}\beta)/k_n(t)$$

since f, k_n and $\underline{K}_n^{-1}(t)$ are uniformly continuous in a neighborhood of t_0 for all $n \geq 1$.

Now for $\ell^{+}_{ijk} = dL^{+}_{ijk}/dt$, $k = 1,2$, $i \leq j$, $a_r \leq j \leq b_r$, we have

$$\max_{k=1,2}\ \max_{a_r \leq j \leq b_r}\ \max_{1 \leq i \leq j}\ \sup_{|t-t_0| \leq dj^{-\frac{1}{2}}} j^{\frac{1}{2}} |L^{+}_{ijk}(t) - L^{+}_{ijk}(t_0) - (t - t_0)\ell^{+}_{ijk}(t_0)| \to 0. \tag{3.3}$$

Since (3.3) holds for each $t_0 \in (0,1)$,

$$\varphi(E_{r1}) \to 0.$$

Next we will show $\varphi(E_{r3}) \to 0$ which completes the proof of (3.1). For $t_0 \in (0,1)$ and $\eta > 0$ choose $\varepsilon > 0 \ni 2\varepsilon[f(F_*^{-1}(t_0)) - \varepsilon]^{-1} < \frac{1}{2}\eta$. Then there exists $r_1 \ni r > r_1$ implies

$$\begin{aligned} f(F_*^{-1}(t_0)) - \varepsilon &\leq \min_{r \leq j \leq b}\ \min_{1 \leq i \leq r} f(\underline{K}_j^{-1}(t_0) - c_{ij}\beta) \\ &\leq \max_{r \leq j \leq b}\ \max_{1 \leq i \leq r} f(\underline{K}_j^{-1}(t_0) - c_{ij}\beta) \\ &\leq f(F_*^{-1}(t_0)) + \varepsilon. \end{aligned}$$

Then for $r > r_1$

$$\max_{1\le i\le r}\ \max_{r\le j\le b}\ |\ell^{+}_{ij1}(t_0) - \ell^{+}_{ir1}(t_0)|$$

$$\le |[f(F_*^{-1}(t_0)) + \varepsilon][f(F_*^{-1}(t_0)) - \varepsilon]^{-1} \tag{3.4}$$

$$- [f(F_*^{-1}(t_0)) - \varepsilon][f(F_*^{-1}(t_0)) + \varepsilon]^{-1}| < \eta \ .$$

Since similar results hold for $\ell^{+}_{ir2}(t_0)$, $i \le r$, $r > r_1$ and (3.4) holds for each $t_0 \in (0,1)$, $\varphi(E_{r3}) \to 0$ and (3.1) holds

Remark 6. By Corollary 2.1 the regression problem can be specialized to the two sample Chernoff-Savage problem with random sample size. Pyke-Shorack (1968b) obtain similar results for a broader class of φ.

CHAPTER 4

RANDOM WEIGHTED EMPIRICAL CUMULATIVES

In this chapter we prove the relative compactness of the weighted empirical cumulatives based on independent random variables with improper but finite distribution functions. This extends the results of Shorack (1973) and Koul (1974).

This result is used to establish the asymptotic normality of random weighted empirical cumulatives based on signed ranks or ranks by establishing that they have the same asymptotic distribution as non-random weighted empirical cumulatives. The asymptotic normality of the latter is established through approximation by their projection on independent random variables.

Define $\{Y_i, 1 \leq i \leq n\}$ to be a sequence of random variables on $[0,1]$ with improper but finite and continuous distribution functions $\{G_{in}, 1 \leq i \leq n\}$. Also define $\sigma_d^2 = \Sigma_1^n d_i^2$ and

$$(4.1) \qquad W_d(t) = \sigma_d^{-1}\Sigma_1^n d_i\{I(Y_i \leq t) - G_{in}(t)\}, \; 0 \leq t \leq 1 .$$

In Proposition 4.1 and the remainder of this work, $\|f\|$ will be the sup norm for any function f on $[0,1]$.

Proposition 4.1. Assume $\{d_i\}$ satisfy a condition similar to (2.2). Assume for the G_{in} defined above that

$$(4.2) \qquad \Sigma_1^n G_{in}(t) - nt$$

is a non-increasing function on [0,1]. Then $\forall \varepsilon > 0$

(4.3) $$\lim_{\delta\to 0} \limsup_{n\to\infty} P[\sup_{|t-s|\le\delta} |W_d(t) - W_d(s)| \ge \varepsilon] = 0 .$$

Proof: The proof is presented in a lemma and remark.

Lemma 4.1. Assume the conditions of Proposition 4.1. Then $\forall\ 0 \le s,t \le 1$

(4.4) $$E|W_d(t) - W_d(s)|^4 \le k_d^4\{3(t - s)^2 + n^{-1}|t - s|\}.$$

Proof: Without loss of generality assume $0 \le s \le t \le 1$. Let $\xi_i = I(s \le Y_i \le t) - p_i$ where $p_i = G_i(t) - G_i(s)$, $i = 1,\dots,n$. Then using the independence of Y_i and the fact that $E\xi_i = 0$, we have by (4.1)

$$E|W_d(t) - W_d(s)|^4 = \sigma_d^{-4}E|\sum_i d_i\xi_i|^4$$

$$= \sigma_d^{-4}[\sum_i d_i^4E\xi_i^4 + 3\sum\sum_{i\neq j} d_i^2d_j^2E\xi_i^2E\xi_j^2]$$

$$= \sigma_d^{-4}[3(\sum_i d_i^2E\xi_i^2)^2 + \sum_i d_i^4(E\xi_i^4 - 3E^2(\xi_i^2))]$$

$$= \sigma_d^{-4}[3(\sum_i d_i^2p_i(1 - p_i))^2 + \sum_i d_i^4(p_i(1 - p_i)(1 - 6p_i + 6p_i^2))]$$

$$\le \sigma_d^{-4}[3 \max_{1\le i\le n} |d_i|^4(\sum_i p_i)^2 + \max_{1\le i\le n} |d_i|^4 \sum_i p_i]$$

$$\le \sigma_d^{-4} \max_{1\le i\le n} |d_i^4|n^2[3(n^{-1}\sum_i p_i)^2 + n^{-2} \sum_i p_i]$$

$$\le k_d^4[3(t - s)^2 + n^{-1}|t - s|]$$

which is (4.4). The last inequality follows because (4.2) implies

$$\Sigma(G_i(t) - G_i(s)) \le nt - ns.$$

This completes the proof of Lemma 4.1.

Remark 7. In Koul (1974), Proposition 4.1 was proved with (4.2) replaced with a stronger condition. Since (4.4), which was basic to the proof in Koul (1974), is also true under the weaker condition, Proposition 4.1 also holds under the weaker condition.

We now present an application of Proposition 4.1 which is used in proving Theorem 4.1. Whenever possible we suppress the subscript n used in the notation of Chapter 2 and denote $\underset{\sim}{K}_n$ be K.

Define for $0 \le t \le 1$

$$V_n(t) = \sigma_n^{-1}\Sigma_1^n c_i[I(|X_i| \le K^{-1}(t))s(X_i) - \mu_i^+(t)],$$

$$W_n(t) = n^{\frac{1}{2}}[K_n(K^{-1}(t)) - t] = n^{-\frac{1}{2}} \Sigma_1^n\{I(|X_i| \le K^{-1}(t)) - L_i^+(t)\},$$

(4.5) $$V_{n1}(t) = \sigma_n^{-1}\Sigma_1^n c_i[I(0 \le X_i \le K^{-1}(t)) - L_{i1}^+(t)]$$

and

$$V_{n2}(t) = \sigma_n^{-1}\Sigma_1^n c_i[I(-K^{-1}(t) \le X_i \le 0) - L_{i2}^+(t)].$$

Note that for $0 \le t \le 1$

(4.6) $$V_n(t) = V_{n1}(t) - V_{n2}(t)$$

and when $c_i \equiv 1$

(4.7) $$W_n(t) = V_{n1}(t) + V_{n2}(t) \quad \text{w.p. } 1.$$

Corollary 4.1. Assume $\{c_i\}$ satisfy (2.2). Then $\forall\ \varepsilon > 0$

$$(4.8)\qquad \lim_{\delta\to 0}\limsup_{n} P[\sup_{|t-s|\le\delta} |V_n(t) - V_n(s)| \ge \varepsilon] = 0$$

and

$$(4.9)\qquad \lim_{\delta\to 0}\limsup_{n} P[\sup_{|t-s|\le\delta} |W_n(t) - W_n(s)| \ge \varepsilon] = 0 .$$

Proof: In Proposition 4.1 take $Y_i \equiv K(X_i)I(0 \le X_i)$ and $d_i \equiv c_i$. Then $G_{in}(t) \equiv L^+_{in1}(t)$ and from (2.8) $\Sigma_1^n L^+_{in1}(t) - nt = -\Sigma_1^n L^+_{in2}(t)$. Hence (4.2) is satisfied for these G_{in}'s. Moreover $V_{n1}(t) = \mathcal{W}_d(t)$. Hence V_{n1} satisfies (4.3). Similar application of Proposition 4.1 to $Y_i \equiv K(-X_i)I(X_i < 0)$ and $d_i \equiv c_i$ will yield (4.8) in view of (4.6). Similarly by (4.7), (4.9) follows when $d_i \equiv 1$.

The following results which appear in Koul-Staudte (1972b) and Koul (1972) are presented here without proof.

Lemma 4.2. The random variable $\{\sup_{0\le t\le 1} |W_n(t)|\}$ is bounded in probability as $n \to \infty$. Also

$$\sup_{0\le t\le 1} |\underline{K}_{a_r}(K^{-1}_{a_r}(t)) - t| = o_p(1) \quad \text{and}$$

(4.10)

$$\sup_{0\le t\le 1} |\underline{H}_{a_r}(\underline{H}^{-1}_{N_r}(t)) - t| = o_p(1).$$

In view of (4.10) and the fact that $\underline{K}_n(x) = \underline{H}_n(x) - \underline{H}_n(-x)$ we have

$$(4.11)\qquad \sup_{0\le t\le 1} |\underline{K}_{a_r}(\underline{K}^{-1}_{N_r}(t)) - t| = o_p(1) .$$

We now state the main theorem of the chapter. Its proof will be presented in a series of lemmas which establish the asymptotic normality of the random weighted empirical cumulatives based on signed ranks by first establishing the asymptotic normality of non-random weighted empirical cumulatives and then establishing that the random and non-random weighted empirical cumulatives have the same asymptotic distribution. Similar results are then presented for random weighted empirical cumulatives based on ranks.

For signed rank statistics introduce for a fixed point $\nu \in [0,1]$

$$T_n^+(\nu) = \sigma_n^{-1}\Sigma_1^n c_i\{I(|X_i| \le K_n^{-1}(\nu))s(X_i) - \mu_i^+(\nu)\} \tag{4.12}$$

where K_n is the empirical cumulative of $\{|X_i|, 1 \le i \le n\}$ defined by (2.7). Also define for $\nu \in (0,1)$

$$Z_n^+(\nu) = \sigma_n^{-1}\{\Sigma_1^n c_i[I(|X_i| \le K^{-1}(\nu))s(X_i) - \mu_i^+(\nu)] - n^{\frac{1}{2}}\hat{c}_n^+(\nu)W_n(\nu)\} \tag{4.13}$$

where $\hat{c}_n^+$ is defined by (2.12).

<u>Theorem 4.1</u>. <u>Assume</u> $\{a_n(i)\}$, $\{N_r\}$ <u>and</u> $\{c_i\}$ <u>satisfy</u> (1.4), (2.1) <u>and</u> (2.2) <u>and for fixed</u> ν, $0 < \nu < 1$, <u>there exists numbers</u> $\ell_{ijk}^+(\nu)$, $1 \le i \le j$, $a_r \le j \le b_r$, $k = 1,2$, <u>such that for any</u> d, $0 < d < \infty$ <u>and</u> $k = 1,2$,

$$\max_{1\le i\le n} \sup_{|t-\nu|\le dn^{-\frac{1}{2}}} |L_{ink}^+(t) - L_{ink}^+(\nu) - (t-\nu)\ell_{ink}^+(\nu)| = o(n^{-\frac{1}{2}}) \tag{4.14}$$

<u>and</u>

(4.15) $\quad \sigma_{a_r}^{-1} a_r^{\frac{1}{2}} \max_{a_r \le j \le b_r} |\hat{c}_j^+(\nu) - \hat{c}_{a_r}^+(\nu)| \to 0 \quad \text{as} \quad r \to \infty.$

For any n define

$$\tau_{n+}^2(\nu) = \Sigma_1^n [c_i^2 [L_{in}^+(\nu) - (\mu_{in}^+(\nu))^2] + (\hat{c}_n^+(\nu))^2 L_{in}^+(\nu)[1 - L_{in}^+(\nu)] \tag{4.16}$$
$$- 2c_i \hat{c}_n^+(\nu)[\mu_{in}^+(\nu)(1 - L_{in}^+(\nu))]] \quad .$$

Then

$$\liminf_{r\to\infty} \tau_{a_r+}^2(\nu)/\sigma_{a_r}^2 > 0 \tag{4.17}$$

implies

$$\tau_{N_r+}^{-1}(\nu)\sigma_{N_r} T_{N_r}^+(\nu) \underset{D}{\to} N(0,1) \quad \text{r.v.} \tag{4.18}$$

The proof of Theorem 4.1 utilizes Lemmas 4.3 through 4.6 and is completed after them.

Lemma 4.3. Assume $\{c_i\}$ satisfy (2.2) and $\ell_{irk}^+(\nu)$, $1 \le i \le r$, $k = 1,2$, satisfy (4.10). Then (4.17) implies

$$\tau_{r+}^{-1}(\nu)\sigma_r T_r^+(\nu) \underset{D}{\to} N(0,1). \tag{4.19}$$

Proof: (4.19) will follow when we show

$$|T_r^+(\nu) - Z_r^+(\nu)| = o_p(1) \tag{4.20}$$

and $\tau_{r+}^{-1}\sigma_r Z_r^+(\nu) \underset{D}{\to} N(0,1)$.

$$T_r^+(\nu) = V_r(K(K_r^{-1}(\nu))) + \sigma_r^{-1}[\bar{\mu}_r^+(K(K_r^{-1}(\nu))) - \bar{\mu}_r^+(\nu)]. \tag{4.21}$$

(4.8) and (4.10) imply

(4.22) $$\sup_{0\le t\le 1} |V_r(K(K_r^{-1}(t))) - V_r(t)| = o_p(1) .$$

Now since $|K_r(K_r^{-1}(t)) - t| \le r^{-1}$, we have by (4.5) and (4.9)

(4.23) $$r^{\frac{1}{2}}[K(K_r^{-1}(t)) - t] = r^{\frac{1}{2}}[K(K_r^{-1}(t)) - K_r(K_r^{-1}(t))] + O(r^{-\frac{1}{2}})$$

$$= -W_r(K(K_r^{-1}(t))) + O(r^{-\frac{1}{2}}) = -W_r(t) + o_p(1).$$

Since the sequence $\{\|W_r\|\}$ is bounded in probability in the limit $\exists$ d $\ni$.

(4.24) $$P[B_r] \to 1$$

where

$$B_r = \{\sup_t |K(K_r^{-1}(t)) - t| \le dr^{-\frac{1}{2}}\}.$$

Combining Assumption (4.14) and the above observations yield

(4.25) $$\sigma_r^{-1}|\bar{\mu}_r^+(K(K_r^{-1}(\nu))) - \bar{\mu}_r^+(\nu) + r^{\frac{1}{2}}\hat{c}_r^+(\nu)W_r(\nu)| = o_p(1).$$

Finally, combining (4.21), (4.22) and (4.25), we have shown

$$|T_r^+(\nu) - V_r^-(\nu) + \sigma^{-1}r^{\frac{1}{2}}\hat{c}_r^+(\nu)W_r(\nu)| = o_p(1).$$

Combining (4.22) and (4.13), we note that

(4.26) $$V_r(\nu) - \sigma^{-1}r^{\frac{1}{2}}\hat{c}_r^+(\nu)W_r(\nu) = Z_r^+(\nu)$$

which completes the proof of (4.20).

To complete the lemma, write

$$Z_r^+(\nu) = \sigma_r^{-1}\Sigma_1^r\{(c_i - \hat{c}_r^+)[I(0 \le X_i \le K^{-1}(\nu)) - L_{ir1}^+(\nu)]$$

$$- (c_i + \hat{c}_r^+)[I(-K^{-1}(\nu) \le X_i < 0) - L_{ir2}^+(\nu)]\}$$

$$= \Sigma_1^r Z_{ri}^+(\nu) \quad \text{(say)}.$$

Then since $|\hat{c}_r^+(\nu)| \le \max|c_i|$ follows from (2.20), we have

$$\max_{1 \le i \le r} |Z_{ri}^+(\nu)| \le 4 \max|c_i|\sigma_r^{-1} = o(1).$$

Therefore $\forall\ \varepsilon > 0$ and r sufficiently large, $\max_{1 \le i \le r} |Z_{ri}^+(\nu)| < \varepsilon$ and the Lindeberg-Feller theorem plus (4.17) yield that $\tau_{r+}^{-1}\sigma_r Z_r^+(\nu)$ is asymptotically normal with parameters $(0,1)$. Therefore (4.19) follows from (4.20).

Before proceeding with Lemma 4.4, we state an inequality which is given in Fernandez (1970).

Inequality. Let $D[0,1]$ be the space of functions on $[0,1]$, the elements of which are right continuous and have left limits. Let $\{Y_i\}$ be sequence of independent random variables in $D[0,1]$ and $\zeta_j = \Sigma_{i=1}^j Y_i$. Then $\forall\ \varepsilon > 0$

$$\text{(4.27)} \qquad P[\max_{1 \le j \le n} \|\zeta_j\| > 2\varepsilon] \le P[\|\zeta_n\| > \varepsilon](1 - \eta_n)^{-1}$$

where

$$\eta_n = \max_{1 \le j \le n} P[\|\zeta_n - \zeta_j\| > \varepsilon] \quad .$$

Lemma 4.4. Assume $\{N_r\}$ and $\{c_i\}$ satisfy (2.1) and (2.2). Then $\forall\ \varepsilon > 0$

(4.28) $$\|V_{N_r} - V_{a_r}\| = o_p(1),$$

(4.29) $$\|W_{N_r} - W_{a_r}\| = o_p(1)$$

and

(4.30) $$\|V^*_{N_r}\| = o_p(1)$$

where for any n, $V^*_n(t) = V_n(K_n(K_n^{-1}(t))) - V_n(t)$.

Finally the sequence $\{\|W_{N_r}\|\}$ is bounded in probability in the limit.

Proof: Unless otherwise specified a_r, N_r and b_r will be denoted by r, N and b. First note that (4.29) and the fact that $\{\|W_r\|\}$ is bounded in probability imply that the sequence $\{\|W_N\|\}$ is bounded in probability. Since (4.29) can be proved by the same methodology used to show (4.28), only the proof of (4.28) and (4.30) will be presented here.

Define for $1 \leq i \leq b$

$$Y_i(t) = \sigma_r^{-1} c_i [I(|X_i| \leq K_r^{-1}(t)) s(X_i) - \mu^+_{ir}(t)]$$

and

$$V'_r(t) = \Sigma_1^r \, Y_i(t) \;.$$

To prove (4.28) we will first show that

(4.31) $$\|V'_N - V_r\| = o_p(1) \;.$$

Note $\{Y_i\}$ are independent random variables and $Y_i \in D[0,1]$. Then by (4.27), we have

$$P[\|V_N' - V_r\| > 2\varepsilon] \le P[\max_{r \le j \le b} \|V_j' - V_r\| > 2\varepsilon] + P(A_r^c)$$

(4.32)

$$\le P[\|V_b' - V_r\| > \varepsilon](1 - \max_{r \le j \le b} P[\|V_b' - V_j'\| > \varepsilon])^{-1} + o_p(1) \to 0$$

as $r \to \infty$ as the following argument shows. For any t and any j, $r \le j \le b$, we have

$$E(|V_b'(t) - V_j'(t)|^2)$$

(4.33) $$= \sigma_r^{-2} \Sigma_{j+1}^{b} c_i^2 E[I\{|X_i| \le \underline{K}_r^{-1}(t)\} s(X_i) - \mu_{ir}^{+}(t)]^2$$

$$\le \sigma_r^{-2}(\sigma_b^2 - \sigma_r^2) \to 0 \quad \text{as} \quad r \to \infty \qquad \text{by (2.3).}$$

Noting that

$$\sup_{0 \le t \le 1} |\underline{K}_j(\underline{K}_r^{-1}(t)) - t| = o(1)$$

and

$$V_j'(t) = \sigma_j^{-1}\sigma_r V_j(\underline{K}_j(\underline{K}_r^{-1}(t))), \; r \le j \le b,$$

we have by (4.8) that

(4.34) $$P[\sup_{|t-s|<\delta} |(V_b'(t) - V_j'(t)) - (V_b'(s) - V_j'(s))| > \varepsilon] \to 0$$

as $r \to \infty$ and then $\delta \to 0$.

Combining (4.33) and (4.34), we have

$$\max_{r \le j \le b} P[\|V_b' - V_j'\| > \varepsilon] \to 0 \quad \text{as} \quad r \to \infty.$$

Therefore (4.32) and consequently (4.31) hold. To complete the proof of (4.28), we start by combining (4.31) with (4.8) to get

$$(4.35) \qquad P[\sup_{|t-s| \le \delta} |V_N'(t) - V_N'(s)| > \varepsilon] \to 0 .$$

Noting that $V_N(t) = \sigma_r \sigma_N^{-1} V_N'(\underline{K}_r(\underline{K}_N^{-1}(t)))$, we have

$$(4.36) \qquad \|V_N - V_r\|$$

$$\le |\sup_{0 \le t \le 1} \sigma_r \sigma_N^{-1}(V_N'(\underline{K}_r(\underline{K}_N^{-1}(t))) - V_N'(t))|$$

$$+ \sigma_r \sigma_N^{-1} \|V_N' - V_r\| + \|V_r\| |\sigma_r \sigma_N^{-1} - 1|$$

$$= R_1 + R_2 + R_3 \quad \text{(say)}.$$

(4.37) $R_3 = o_p(1)$ by (2.4) and the fact that $\{\|V_r\|\}$ are bounded in probability.

$$(4.38) \qquad R_2 = o_p(1) \quad \text{by (4.31) and (2.4)}.$$

Finally define $D_r^+ = [\sup_{0 \le t \le 1} |\underline{K}_r(\underline{K}_N^{-1}(t)) - t| < \delta]$ and note from (4.11)

$$(4.39) \qquad P(D_r^+) \to 1.$$

Then by (4.35) and (4.39)

$$(4.40) \quad P[R_1 > \varepsilon] \le P[\sup_{|t-s|<\delta} \sigma_r \sigma_N^{-1} |V_N'(t) - V_N'(s)| > \varepsilon] + P(D_{r+}^c) \to 0.$$

Combining (4.36) through (4.40) yields (4.28).

Finally, the proof of the lemma is complete when we show that (4.30) holds. Since the sequence $\{\|W_N\|\}$ is bounded in probability, there exists d such that

$$C_r = \{\sup_t |\underline{K}_N(K_N^{-1}(t)) - t| \le dN^{-\frac{1}{2}}\}$$

and

$$P[C_r] \to 1 . \tag{4.41}$$

Then (4.30) follows by combining (4.8), (4.28) and (4.41)

$$\lim_{r\to\infty} P\{\|V_N^*\| > \varepsilon\}$$

$$\le \lim_{\delta\to 0} \lim_{r\to\infty} (P\{\sup_{|t-s|<\delta} |V_N(t) - V_N(s)| > \varepsilon\} + P[C_r^c]) = 0.$$

Lemma 4.5. Assume the conditions of Theorem 4.1. Then at ν

$$|\sigma_{N_r}^{-1}\, \bar{\mu}_{N_r}^{*}(\nu) - \sigma_{a_r}^{-1}\, \bar{\mu}_{a_r}^{*}(\nu)| = o_p(1) \tag{4.42}$$

where for any n

$$\bar{\mu}_n^{*}(\nu) = \bar{\mu}_n^{+}(\underline{K}_n(K_n^{-1}(\nu))) - \bar{\mu}_n^{+}(\nu). \tag{4.43}$$

Proof: Again we will write r, N, and b for a_r, N_r and b_r. We will repeatedly use the fact that since $\ell_{irk}^{+} \ge 0$, $i \le r$, $k = 1,2$, (2.20) implies for any $r \ge 1$

$$\sigma_r^{-1} r^{\frac{1}{2}} \|\hat{c}_r^{+}\| \le \sigma_r^{-1} r^{\frac{1}{2}} \max_{1\le i\le r} |c_i| r^{-1} \|\sum_i \sum_k \ell_{irk}^{+}\| \le k_c . \tag{4.44}$$

To prove (4.42) we will first show that

$$|\sigma_{N_r}^{-1} N_r^{\frac{1}{2}} \hat{c}_{N_r}^{+}(\nu) W_{N_r}(\nu) - \sigma_{a_r}^{-1} a_r^{\frac{1}{2}} \hat{c}_{a_r}^{+}(\nu) W_{a_r}(\nu)| = o_p(1) \tag{4.45}$$

In particular we will show that

$$R_1 = |\sigma_N^{-1}N^{\frac{1}{2}}\hat{c}_N^+(\nu) - \sigma_r^{-1}r^{\frac{1}{2}}\hat{c}_r^+(\nu)| = o_p(1) \tag{4.46}$$

which by (4.44) also implies

$$\lim_r P[\sigma_N^{-1}N^{\frac{1}{2}}|\hat{c}_N^+(\nu)| \le K < \infty] = 1 . \tag{4.47}$$

The proof of (4.45) will then be complete by the triangle inequality, (4.29) and the fact that the sequence $\{\|W_{a_r}\|\}$ is bounded in probability. We now prove (4.46). By (2.1), (4.15) and (4.44)

$$\begin{aligned}|R_1| &\le \sigma_N^{-1}N^{\frac{1}{2}}|\hat{c}_N^+(\nu) - \hat{c}_r^+(\nu)| + o_p(1) \\ &\le \sigma_r^{-1}b^{\frac{1}{2}} \max_{a_r\le j\le b_r} |\hat{c}_j^+(\nu) - \hat{c}_r^+(\nu)| + o_p(1) = o_p(1) .\end{aligned}$$

To prove (4.42) we first observe that (4.14) implies that $\forall\, \varepsilon > 0$ there exists $r_\varepsilon \ni \cdot$ if $r > r_\varepsilon$ and $|t - \nu| \le dj^{-\frac{1}{2}}$, $r \le j \le b$, then we have

$$\begin{aligned}&|\sigma_j^{-1} \Sigma_1^j c_i[\mu_{ij}^+(t) - \mu_{ij}^+(\nu)] - \sigma_r^{-1} \Sigma_1^r c_i[\mu_{ir}^+(t) - \mu_{ir}^+(\nu)]| \\ &\le |\sigma_j^{-1}j\hat{c}_j^+(\nu)(t - \nu) - \sigma_r^{-1}r\hat{c}_r^+(\nu)(t - \nu)| + 4\varepsilon k_c \,. \\ &\quad + 4j\sigma_j^{-1} \max_{1\le i\le j} |c_i| \max_{k=1,2} \max_{1\le i\le j} \sup_{|t-\nu|\le dj^{-\frac{1}{2}}} |L_{ijk}^+(t) - L_{ijk}^+(\nu) \\ &\qquad - (t - \nu)\ell_{ijk}^+(\nu)| \\ &\le |\sigma_j^{-1}j\hat{c}_j^+(\nu)(t-\nu) - \sigma_r^{-1}r\hat{c}_r^+(\nu)(t-\nu)| + 4\varepsilon k_c \,.\end{aligned} \tag{4.48}$$

To proceed let C_r be defined by (4.41). Note on $A_r \cap C_r$. we have for $r \le j \le b$

$$\sup_t |\underline{K}_j(K_j^{-1}(t)) - t| \le dj^{-\frac{1}{2}}.$$

We complete the proof of (4.42) by partitioning Ω based on $\{N_r\}$, applying (4.5), (4.23) and (4.48) and recombining. Let $\varepsilon_0 > 4k_c\varepsilon > 0$. Then

$$P[|\sigma_N^{-1}\bar{\mu}_N^*(\nu) - \sigma_r^{-1}\bar{\mu}_r^*(\nu)| > 3\varepsilon_0]$$

$$\leq P[\bigcup_{j=r}^{b}([|\sigma_j^{-1}\bar{\mu}_j^*(\nu) - \sigma_r^{-1}\bar{\mu}_r^*(\nu)| > 3\varepsilon_0]\cap[C_r]\cap[N_r = j])]$$

$$+ P(A_r^c) + P(C_r^c)$$

$$\leq P[\bigcup_{j=r}^{b}([|\sigma_j^{-1}j^{\frac{1}{2}}W_j(\nu)\hat{c}_j^+(\nu) - \sigma_r^{-1}r^{\frac{1}{2}}W_r(\nu)\hat{c}_r^+(\nu)| > \varepsilon_0]\cap[C_r]\cap[N_r = j])]$$

$$+ o_p(1)$$

$$\leq P[[|\sigma_N^{-1}N^{\frac{1}{2}}\hat{c}_N^+(\nu)W_N(\nu) - \sigma_r^{-1}r^{\frac{1}{2}}\hat{c}_r^+(\nu)W_r(\nu)| > \varepsilon_0] \cap [A_r]] + o_p(1)$$

$$= o_p(1) \quad \text{by (4.45)}.$$

<u>Lemma 4.6. Assume the conditions of Theorem 4.1. Then</u>

$$\tau_{N+}(\nu)\tau_{r+}^{-1}(\nu) \to 1 \text{ w.p. } 1 \quad \text{as} \quad r \to \infty.$$

<u>Proof</u>: As usual we will use r, N and b instead of a_r, N_r and b_r. Since ν is fixed we will simplify notation by not revealing ν when possible.

We will repeatedly use the observations that (4.17) insures that $\tau_r^{-2}\sigma_r^2$ is bounded, that (2.8) and (4.11) yield

$$(4.49) \quad r^{-1}|\Sigma_1^r(L_{iN}^+(\nu) - L_{ir}^+(\nu))| = |\underline{K}_r(\underline{K}_N^{-1}(\nu)) - \nu| = o_p(1)$$

and that

(4.50) $r^{-1}\Sigma_1^r|\mu_{iN}^+ - \mu_{ir}^+| \leq r^{-1}|\Sigma_1^r(L_{iN}^+ - L_{ir}^+)| = o_p(1)$.

To show that (4.50) holds assume $K_r^{-1} \geq K_N^{-1}$, then it is easily shown that

$$L_{iN}^+ - L_{ir}^+ \leq \mu_{iN}^+ - \mu_{ir}^+ \leq L_{ir}^+ - L_{iN}^+.$$

Similarly when $K_r^{-1} \leq K_N^{-1}$

$$L_{ir}^+ - L_{iN}^+ \leq \mu_{ir}^+ - \mu_{iN}^+ \leq L_{iN}^+ - L_{ir}^+.$$

Therefore

$$\Sigma_1^r|\mu_{iN}^+ - \mu_{ir}^+| \leq \Sigma_1^r|L_{ir}^+ - L_{iN}^+| = |\Sigma_1^r(L_{ir}^+ - L_{iN}^+)|.$$

Now by (2.1) and (4.16)

$$\tau_{N+}^2(\nu)\tau_{r+}^{-2}(\nu) = 1 + \Sigma_{m=1}^5 R_m + o_p(1) \tag{4.51}$$

where

$$R_1 = \tau_{r+}^{-2}\ \Sigma_1^r c_i^2(L_{iN}^+ - L_{ir}^+)$$

$$R_2 = \tau_{r+}^{-2}\ \Sigma_1^r c_i^2((\mu_{ir}^+)^2 - (\mu_{iN}^+)^2)$$

$$R_3 = \tau_{r+}^{-2}\ \Sigma_1^r[\hat{c}_{N+}^2 L_{iN}^+(1 - L_{iN}^+) - \hat{c}_{r+}^2 L_{ir}^+(1 - L_{ir}^+)]$$

$$R_4 = 2\tau_{r+}^{-2}\Sigma_1^r c_i[\hat{c}_r^+\mu_{ir}^+(1-L_{ir}^+) - \hat{c}_N^+\mu_{iN}^+(1 - L_{iN}^+)]$$

$$R_5 = \tau_{r+}^{-2}\ \Sigma_{r+1}^N[c_i^2(L_{iN}^+ - (\mu_{iN}^+)^2) + \hat{c}_{N+}^2 L_{iN}^+(1 - L_{iN}^+)$$

$$- 2c_i\hat{c}_N^+\mu_{iN}^+(1 - L_{iN}^+)].$$

In view of (4.49) and (4.50)

$$(4.52) \quad |R_1| \leq \tau_{r+}^{-2} \max_i c_i^2 |\Sigma_1^r (L_{iN}^+ - L_{ir}^+)| = o_p(1)$$

and

$$(4.53) \quad |R_2| \leq \tau_{r+}^{-2} \max_i c_i^2 \Sigma_1^r |\mu_{ir}^+ + \mu_{iN}^+||\mu_{ir}^+ - \mu_{iN}^+| = o_p(1).$$

Next

$$\begin{aligned} |R_3| &\leq \tau_{r+}^{-2} |(\hat{c}_N^+)^2 - (\hat{c}_r^+)^2||\Sigma_1^r L_{iN}^+(1 - L_{iN}^+)| \\ &+ \tau_{r+}^{-2}(\hat{c}_r^+)^2|\Sigma_1^r L_{iN}^+(1 - L_{iN}^+) - \Sigma_1^r L_{ir}^+(1 - L_{ir}^+)| \\ &= |R_{31}| + |R_{32}| \quad \text{(say)}. \end{aligned}$$

In view of (4.15), (4.44) and (4.49)

$$|R_{31}| \leq \tau_{r+}^{-2} \max_{r \leq j \leq b} |\hat{c}_j^+ - \hat{c}_r^+||\hat{c}_j^+ + \hat{c}_r^+||\Sigma_1^r L_{iN}^+| = o_p(1)$$

and

$$|R_{32}| \leq 3\tau_{r+}^{-2}(\hat{c}_r^+)^2|\Sigma_1^r (L_{iN}^+ - L_{ir}^+)| = o_p(1).$$

Therefore

$$(4.54) \quad |R_3| = o_p(1).$$

Similar techniques yield

$$(4.55) \quad |R_4| = o_p(1).$$

Finally in view of (4.47)

$$(4.56)\qquad |R_5| \le \tau_{r+}^{-2}\Sigma_{r+1}^{N}(\hat{c}_i^2 + (\hat{c}_N^+)^2 + 2c_i|\hat{c}_N^+|) + o_p(1)$$

$$\le \tau_{r+}^{-2}(N - r)[\max_{1\le i\le b} |c_i^2| + (\hat{c}_N^+)^2 + 2|\hat{c}_N^+| \max_{1\le i\le b} |c_i|]$$

$$+ o_p(1) = o_p(1).$$

Combining (4.51) through (4.56) completes the proof of the lemma.

Proof of Theorem 4.1. As usual we will use r, N and b instead of a_r, N_r and b_r.

We prove (4.18) by using the decomposition

$$T_n^+(\nu) = V_n(K(K_n^{-1}(\nu))) + \sigma_n^{-1}\bar{\mu}_n^*(\nu)$$

where $\bar{\mu}_n^*(\nu)$ is defined by (4.43). Note that from Lemmas 4.4 and 4.5 we have

$$(4.57)\qquad |T_N^+(\nu) - T_r^+(\nu)| \le |V_N^*(\nu) + V_N(\nu) - V_r^*(\nu) - V_r(\nu)|$$

$$+ |\sigma_N^{-1}\bar{\mu}_N^*(\nu) - \sigma_r^{-1}\bar{\mu}_r^*(\nu)| = o_p(1) .$$

Combining (2.4) and (4.57) with the results of Lemmas 4.3 and 4.6 completes the proof of (4.18).

For rank statistics we present Theorem 4.2 which is the analog of Theorem 4.1. The proof of Theorem 4.2 is similar to the proof of Theorem 4.1 with appropriate changes in notation. Define

$$T_n(\nu) = \sigma_n^{-1}\Sigma_1^n c_i\{I(X_i \le H_n^{-1}(\nu)) - L_{in}(\nu)\}$$

and

$$\tau_n^2(\nu) = \Sigma_1^n[(c_i - \hat{c}_n(\nu))^2(1 - L_i(\nu))L_i(\nu)].$$

<u>Theorem 4.2.</u> <u>Assume</u> $\{a_n(i)\}$, $\{N_r\}$ <u>and</u> $\{c_i\}$ <u>satisfy</u> (1.4), (2.1) <u>and</u> (2.2) <u>and there exist numbers</u> $\ell_{ij}(\nu)$, $1 \le i \le j$, $a_r \le j \le b_r$ <u>such that, at a fixed point</u> ν, <u>for any</u> d, $0 < d < \infty$

$$\max_{1\le i\le n} \sup_{|t-\nu|\le dn^{-\frac{1}{2}}} |L_{in}(t) - L_{in}(\nu) - (t - \nu)\ell_{in}(\nu)| = o(n^{-\frac{1}{2}}),$$

<u>and</u>

$$\sigma_{a_r}^{-1} a_r^{\frac{1}{2}} \max_{a_r\le j\le b_r} |\hat{c}_j(\nu) - \hat{c}_{a_r}(\nu)| = o(a_r^{-\frac{1}{2}}) \quad \underline{\text{as}} \quad r \to \infty .$$

<u>Then</u> $\liminf_{r\to\infty} \tau_{a_r}^2(\nu)/\sigma_r^2 > 0$ <u>implies</u>

$$\tau_{N_r}^{-1}(\nu)\sigma_{N_r} T_{N_r}(\nu) \xrightarrow[D]{} N(0,1) \quad \text{r.v.}$$

CHAPTER 5

PROOF OF THEOREMS 2.1 AND 2.2

Proof of Theorem 2.1. As before we will use r, N and b to represent a_r, N_r and b_r. By (2.6) we may assume φ is $\searrow$.

Define for any n

(5.1) $$T_n^+(\varphi) = \sigma_n^{-1}[S_n^+(\varphi) - \bar{\mu}_n^+(\varphi)] .$$

Since the assumptions of Theorem 2.1 are stronger than the assumptions of Proposition A.1, we have by (A.28) and Remark 9

$$\tau_{r+}^{-1}(\nu)\sigma_r T_r^+(\varphi) \xrightarrow[D]{} N(0,1) .$$

Thus to prove (2.15) it suffices, in view of (2.14), to prove

(5.2) $$|T_N^+(\varphi) - T_r^+(\varphi)| = o_p(1) .$$

To prove (5.2) note that for all $r \geq 1$, $T_r^+(\varphi) = \int_0^{r/r+1} \varphi(t)\,dT^+(r^{-1}(r+1)t)$ w.p. 1 where $T^+(t)$ is defined by (4.12).

Since $T^+(0) = 0$ and φ is bounded and $\searrow$, upon integration by parts one has

$$T_r^+(\varphi) = -\int_0^{r/r+1} T_r^+(r^{-1}(r + 1)t)d\varphi(t) + T_r^+(1)\varphi(1)$$

(5.3)
$$= -\int_0^1 T_r^+(t)d\varphi(t) + T_r^+(1)\varphi(1) + o_p(1) \ .$$

The latter follows from the boundedness of φ and the fact that $\sup_t |T_r^+(r^{-1}(r + 1)t) - T_r^+(t)| = o_p(1)$. Similar results hold for $T_N^+(\varphi)$ by (4.57).

Now by (4.28), the boundedness of φ, and the fact that for any n, $T_n^+(1) = V_n(1)$, we have

$$\varphi(1)|T_r^+(1) - T_N^+(1)| = o_p(1).$$

Thus to prove (5.2) it remains to show

$$|\int_0^1 (T_N^+ - T_r^+)d\varphi| = o_p(1) \ .$$

Defining for any n, V_n^* and $\bar{\mu}_n^*$ by (4.30) and (4.43) respectively and using the decomposition presented at (4.57) we have

$$|\int_0^1 (T_N^+ - T_r^+)d\varphi| \leq \int_0^1 |V_{N+}^*(t) + V_N(t) - V_{r+}^*(t) - V_r(t)|d\varphi(t)$$

(5.4)
$$+ |\int_0^1 \{\sigma_N^{-1}\bar{\mu}_N^*(t) - \sigma_r^{-1}\bar{\mu}_r^*(t)\}d\varphi(t)|$$

$$= R_1 + R_2 \qquad \text{(say)}.$$

(5.5)
$$|R_1| = o_p(1)$$

by (4.28), (4.30) and the boundedness of φ. Next we proceed to prove R_2 is $o_p(1)$. Since (4.44) holds for any $j \geq 1$

$$P[\|\sigma_N^{-1}N^{\frac{1}{2}}\hat{c}_N^+\| \le k_c]$$

$$(5.6)\qquad = P[\bigcup_{j=1}^{\infty}([N=j] \cup [\|\sigma_j^{-1}j^{\frac{1}{2}}\hat{c}_j^+\| \le k_c])] = 1 .$$

For any η, $0 < \eta < \infty$, r sufficiently large and E_r defined by (2.9) we have by (2.4) and (4.44)

$$\sup_{t\in E_r^c} |\sigma_N^{-1}N^{\frac{1}{2}}\hat{c}_N^+(t) - \sigma_r^{-1}r^{\frac{1}{2}}\hat{c}_r^+(t)|$$

$$\le \sup_{t\in E_r^c} \sigma_N^{-1}N^{\frac{1}{2}}|\hat{c}_N^+(t) - \hat{c}_r^+(t)| + o_p(1)$$

$$(5.7)\qquad \le \sigma_r^{-1}b^{\frac{1}{2}} \max_{r\le j\le b} \sup_{t\in E_r^c} |\hat{c}_j^+(t) - \hat{c}_r^+(t)| + o_p(1) \le 2\eta + o_p(1).$$

Define A_r and C_r as in (2.1) and (4.41). Let $G_r = A_r \cap C_r$. Note

$$(5.8)\qquad P(G_r^c) \to 0 \quad \text{as} \quad r \to \infty .$$

For an arbitrary $\varepsilon > 0$, $\eta > 0$ (a function of ε), we have for E_r

$$|\int_0^1\{\sigma_N^{-1}\bar{\mu}_N^* - \sigma_r^{-1}\bar{\mu}_r^*\}d\varphi|$$

$$(5.9)\qquad \le |\int_{E_r}\{\sigma_N^{-1}\bar{\mu}_N^* - \sigma_r^{-1}\bar{\mu}_r^*\}d\varphi| + |\int_{E_r^c}\{\sigma_N^{-1}\bar{\mu}_N^* - \sigma_r^{-1}\bar{\mu}_r^*\}d\varphi|$$

$$= R_{21} + R_{22} \qquad \text{(say)}.$$

We first show $R_{21} = o_p(1)$. For any j, $r \le j \le b$ and $t \in (0,1)$ we have on G_r

$|\bar{\mu}_j^*(t)|$

$$= |\Sigma_1^j c_i (L_{ij1}^+(\underline{K}_j(K_j^{-1}(t))) - L_{ij1}^+(t) - L_{ij2}^+(\underline{K}_j(K_j^{-1}(t))) + L_{ij2}^+(t))|$$

$$\leq |\Sigma_1^j c_i (L_{ij}^+(\underline{K}_j(K_j^{-1}(t))) - L_{ij}^+(t))|$$

$$\leq \max_{1\leq i\leq j} |c_i| |\Sigma_1^j (L_{ij}^+(\underline{K}_j(K_j^{-1}(t))) - L_{ij}^+(t))|$$

$$\leq d \max_{1\leq i\leq j} |c_i| j^{\frac{1}{2}} \qquad \text{by (2.8) and (4.41).}$$

Consequently on G_r we have

$$R_{21} \leq \sigma_r^{-1} \max_{1\leq i\leq b} |c_i| (db^{\frac{1}{2}} + dr^{\frac{1}{2}}) \varphi(E_r)$$

$$\leq 3dk_c \varphi(E_r) .$$

Therefore by (2.9) and (5.8)

$$R_{21} = o_p(1). \tag{5.10}$$

Next we show $R_{22} = o_p(1)$. On $G_r \cap [N_r = j]$

$$|R_{22}| \leq \sup_{t\in E_r^c} |\sigma_j^{-1} j^{\frac{1}{2}} W_j(t) \hat{c}_j^+(t) - \sigma_r^{-1} r^{\frac{1}{2}} W_r(t) \hat{c}_r^+(t)| \|\varphi\| + 4\eta k_c \|\varphi\|$$

$$\leq \sup_{t\in E_r^c} |\sigma_N^{-1} N^{\frac{1}{2}} W_N(t) \hat{c}_N^+(t) - \sigma_r^{-1} r^{\frac{1}{2}} W_r(t) \hat{c}_r^+(t)| \|\varphi\| + \tfrac{1}{4}\varepsilon .$$

The first inequality follows from (2.10) by an argument similar to the proof of (4.48). Therefore since $\|\varphi\|$ is bounded we have on G_r.

$$|R_{22}| \le \sup_{t\in E_r^c} |\sigma_N^{-1}N^{\frac{1}{2}}W_N(t)\hat{c}_N^+(t) - \sigma_r^{-1}r^{\frac{1}{2}}W_r(t)\hat{c}_r^+(t)|\,\|\varphi\| + \tfrac{1}{4}\varepsilon$$

$$\le \sigma_N^{-1}N^{\frac{1}{2}}\|\hat{c}_N^+\|\|W_N - W_r\|\,\|\varphi\| + \tfrac{1}{4}\varepsilon$$

$$+ \sigma_N^{-1}N^{\frac{1}{2}}\|W_r\|\sup_{t\in E_r^c}|\hat{c}_N^+(t) - \hat{c}_r^+(t)|\,\|\varphi\|$$

$$= o_p(1).$$

The first term on the right side is $o_p(1)$ since $\sigma_N^{-1}N^{\frac{1}{2}}\|\hat{c}_N^+\|$ is bounded w.p. 1 and $\|W_N - W_r\| = o_p(1)$. The second term is $o_p(1)$ because $\|W_r\|$ is bounded in probability in the limit and $\sup_{t\in E_r^c}|\hat{c}_N^+(t) - \hat{c}_r^+(t)| < \eta$ for η arbitrarily small. Consequently by (5.9)

$$R_{22} = o_p(1). \tag{5.11}$$

Combining (5.5) and (5.9) through (5.11) completes the proof of $|\int_0^1 (T_N^+ - T_r^+)d\varphi| = o_p(1)$ and the theorem

<u>Proof of Theorem</u> 2.2. With appropriate changes in notation the proof of Theorem 2.2 is similar to that of Theorem 2.1. The details are left to the reader.

<u>Proof of Corollary</u> 2.1. First note that $\{c_i\}$ defined by (2.21) satisfy conditions similar to (2.2) and (2.3) since $(a_r/n_r)^{\frac{1}{2}} = \lambda_0^{-\frac{1}{2}} + o_p(1)$ and $\sigma_{a_r}/\sigma_{b_r} = a_r'/b_r' + o_p(1) = 1 + o_p(1)$. Since the proof of Corollary 2.1 parallels the proof of Theorem 2.1 only the modifications will be discussed. Each place (2.2) or (2.3) are used in the proof of Theorem 2.1 the conditions above may be used. Similarly Corollary 4.1 and Lemma 4.6 hold as before for these $\{c_i\}$. Therefore (5.5) holds. Finally (5.6) and (5.7) hold since for all r

$$\|\sigma_r^{-1} a_r^{\frac{1}{2}} \hat{c}_r\| \leq (a_r/n_r)^{\frac{1}{2}} r^{-1} \|\Sigma_1^r \ell_{ir}\| \leq \lambda_0^{-\frac{1}{2}} + o_p(1) \ .$$

The remainder of the proof follows without change.

Similar arguments show that Theorem 2.2 also holds for these $\{c_i\}$.

CHAPTER 6

PROOF OF THEOREM 2.3

This proof is developed in Lemmas 6.1 through 6.3. Koul-Staudte (1972a) presents the asymptotic normality of the nonrandom linear sign-rank statistics. The specialization of their Theorem 2.2 to $\{X_i\}$ which are iid with symmetric distribution F is presented in Lemma 6.1. Lemmas 6.2 and 6.3 verify the methodology developed by Anscombe (1952), thereby completing the proof.

Lemma 6.1. *Let* $X_1,\ldots,X_n$ *be iid random variables with a continuous symmetric distribution function* F. *Let* $\psi(t) = \varphi_1(t) - \varphi_2(t)$, $t \in (0,1)$, *where* φ_i *are nondecreasing, square integrable. Let* $a_n(i)$ *and* $\{c_i\}$ *satisfy* (1.5) *and* (2.5) *and* $A^2 = \int_0^1 \psi^2(t)dt$. *Then*

$$A^{-1}\sigma_n^{-1}(S_n^+ - ES_n^+) \to N(0,1).$$

Lemma 6.2. *Define* $\{a_n(i)\}$ *by* (1.5). *Let* $\mathcal{F}_n$ *be the* σ-*field generated by* $\{(s(X_i), R_{in}^+), 1 \le i \le n\}$. *Then* $\{S_n^+,\mathcal{F}_n\}$ *is a martingale.*

Proof: Using the density of the ith order statistic of a set of independent observations each distributed uniformly on $(0,1)$ one can show that

$$i(n+1)^{-1}E(\psi(U_{n+1\ i+1})) + (n - i+1)(n+1)^{-1}E(\psi(U_{n+1i})) = E(\psi(U_{ni})).$$

Therefore

$$a_n(i) = [i/(n+1)]a_{n+1}(i+1) + [(n-i+1)/(n+1)]a_{n+1}(i) \ . \tag{6.1}$$

Since F is symmetric, $\{s(X_i),\ 1 \leq i \leq n\}$ and $\{R^+_{ni},\ 1 \leq i \leq n\}$ are independent. Therefore given $\{(s(X_i), R^+_{ni}),\ 1 \leq i \leq n\}$

(6.2) $$s(X_{n+1}) = \pm 1 \text{ with probability } \tfrac{1}{2}$$

and

(6.3) $$\begin{aligned} R^+_{n+1i} &= R^+_{ni} + 1 && \text{with probability } R^+_{ni}/(n+1) \\ &= R^+_{ni} && \text{with probability } (n+1 - R^+_{ni})/(n+1). \end{aligned}$$

Combining (6.1) and (6.3) we have for $i = 1,\ldots,n$

(6.4) $$\begin{aligned} E[a_{n+1}(R^+_{n+1i})|F_n] &= (R^+_{ni}/(n+1))a_{n+1}(R^+_{ni}+1) \\ &\quad + ((n+1 - R^+_{ni})/(n+1))a_{n+1}(R^+_{ni}) = a_n(R^+_{ni}). \end{aligned}$$

Now by (6.2) and (6.4)

$$\begin{aligned} E[S^+_{n+1}|F_n] &= E[\Sigma_1^{n+1} c_i s(X_i) a_{n+1}(R^+_{n+1i})|F_n] \\ &= c_{n+1}E[s(X_{n+1})|F_n]E[a_{n+1}(R^+_{n+1i})|F_n] + \Sigma_1^n c_i s(X_i) E[a_{n+1}(R^+_{n+1i})|F_n] \\ &= \Sigma_1^n c_i s(X_i) a_n(R^+_{ni}) = S^+_n . \end{aligned}$$

Therefore $\{S^+_n, F_n\}$ is a martingale. This completes the proof.

<u>Lemma 6.3.</u> <u>Assume the conditions of Theorem</u> 2.3. <u>For any</u> n <u>let</u> $T_n = S^+_n/A\sigma_n$. <u>Then</u> $\forall\ \varepsilon > 0$ <u>and</u> $\eta > 0$, <u>there exists</u> $\delta > 0$ <u>such that as</u> $r \to \infty$

$$P[\max_{|r-j| \leq \delta r} |T_r - T_j| > \eta] < \varepsilon .$$

Proof: Denote $r - [\delta r]$ by ℓ and $r + [\delta r]$ by u where $[r]$ is the largest integer smaller than r. Now since δ can be arbitrarily small, we have by (2.3)

$$(6.5) \qquad \max_{|j-r|<\delta r} |\sigma_r^{-2}(\sigma_j^2 - \sigma_r^2)| \le |\sigma_\ell^{-2}(\sigma_\ell^2 - \sigma_u^2)| \to 0 \quad \text{as} \quad r \to \infty .$$

Since $S_r^+ = A\sigma_r T_r$

$$(6.6) \qquad |T_r - T_j| \le |(\sigma_j - \sigma_r)\sigma_j^{-1}||T_r| + A^{-1}\sigma_j^{-1}|S_r^+ - S_j^+| .$$

Combining (6.5) and the fact that T_r is bounded in probability in the limit, the first term on the right hand side is $o_p(1)$.

Define $J_r^2 = r^{-1}\Sigma_1^r a_r^2(i)$. Note that $ES_r^+ = 0$, $V(S_r^+) = J_r^2 \Sigma_1^r c_i^2 = \sigma_r^2 J_r^2$ and $J_r^2 = A^2 + o(1)$. Since $\sigma_\ell^{-2}\sigma_u^2 \to 1$ as $r \to \infty$, (6.6) implies for a suitably chosen δ and $\eta_1 < \frac{1}{2}A\eta$

$$P[\sup_{|r-j|<\delta r} |T_r - T_j| > \eta] \le P[\sup_{|r-j|<\delta r} \sigma_j^{-1}|S_r^+ - S_j^+| > \eta_1] + o_p(1)$$

$$\le (\eta_1\sigma_\ell)^{-2}E(S_u^{+2} - S_\ell^{+2}) + o_p(1) \le (\eta_1\sigma_\ell)^{-2}[\sigma_u^2 J_u^2 - \sigma_\ell^2 J_\ell^2] < \varepsilon .$$

The third inequality follows by the Kolmogorov inequality for martingales (Loève, 1963, pg. 386). This completes the proof of Anscombe's condition and therefore the proof of Theorem 2.3.

APPENDIX

APPENDIX

Proposition A.1 presents the asymptotic normality for signed rank statistics for φ satisfying Hoeffding's condition. These φ may be discontinuous. This proposition retains a simpler centering constant similar to the centering constants used for rank statistics presented in Dupač (1969), Hoeffding (1973) and Koul (1974). However, it eliminates the condition D-H (2.12) which Dupač-Hájek (1969) and Dupač (1969) used and relaxes the boundedness condition which Koul (1974) used. No conditions are required on the alternatives for the absolutely continuous part of φ as was shown by Hájek (1968) for rank statistics and Husková (1970) for signed rank statistics.

To prove Proposition A.1 we first present a lemma for signed rank statistics. This lemma is an analog of a similar lemma for rank statistics in Hoeffding (1973).

Define

(A.1) $$J(\varphi) = \int_0^1 t^{\frac{1}{2}}(1 - t)^{\frac{1}{2}} d\varphi(t) .$$

<u>Lemma A.1</u>. <u>If</u> φ <u>is non-decreasing, then</u>

(A.2) $$\Sigma_1^n \left| E\varphi(R_i^+/(n+1))s(X_i) - \int_0^1 \varphi(t)d\mu_i^+(t) \right| \leq 8n^{\frac{1}{2}} J(\varphi) .$$

<u>Proof</u>: For $x \in (-\infty,\infty)$ define

$$\mathcal{W}_i(x) = I(|X_i| > |x|) + nK_n(|x|).$$

Then

(A.3) $$E\{\varphi(R_i^+/(n+1)) \,|\, |X_i| = x\} = E\varphi(\mathcal{W}_i(x)/(n+1)).$$

Consequently for $1 \leq i \leq n$

$$|E\varphi(R_i^+/(n+1))s(X_i) - \int\varphi(t)d\mu_i(t)|$$

$$= |E\varphi(R_i^+/(n+1))I(X_i > 0) - \int\varphi(t)dL_{in1}^+(t)$$

(A.4) $$- (E\varphi(R_i^+/(n+1))I(X_i < 0) - \int\varphi(t)dL_{in2}^+(t))|$$

$$\leq \int E|\varphi(\mathcal{W}_i(K^{-1}(t))/(n+1)) - \varphi(t)|dL_{in}^+(t) \;.$$

Since $\mathcal{W}_i(x) \leq nK_n(x) + 1$ and $\mathcal{W}_i(x) \leq n$, we have

(A.5) $$\varphi(\mathcal{W}_i(x)/(n+1)) \leq g(nK_n(x)) \leq g(\mathcal{W}_i(x))$$

where

$$g(i) = \min\{\varphi((i+1)/(n+1)), \varphi(n/(n+1))\}, \; 0 \leq i \leq n \;.$$

Now by (A.4)

(A.6) $$\Sigma_1^n|E\varphi(R_i^+/(n+1))s(X_i) - \int_0^1\varphi(t)d\mu_i^+(t)|$$

$$\leq \Sigma_1^n \int E|\varphi(\mathcal{W}_i(K^{-1}(t))/(n+1)) - g(nK_n(K^{-1}(t)))|dL_{in}^+(t)$$

$$+ \Sigma_1^n \int E|g(nK_n(K^{-1}(t))) - \varphi(t)|dL_{in}^+(t)$$

$$= R_1 + R_2 \quad \text{(say)} \;.$$

But by (A.3) and (A.5)

$$R_1 \leq \Sigma_1^n \int E[g(\mathcal{W}_j(K^{-1}(t))) - \varphi(\mathcal{W}_i(K^{-1}(t))/(n+1))]dL_{in}^+(t)$$

$$\text{(A.7)} \qquad = \Sigma_1^n E[g(R_i^+) - \varphi(R_i^+/(n+1))]$$

$$= \Sigma_1^n \{g(i) - \varphi(i/(n+1))\}$$

$$= \varphi(n/(n+1)) - \varphi(1/(n+1)) \;.$$

Now the remainder of Hoeffding's proof of Proposition 2 holds without change by considering K and K_n to be the distribution and empirical distribution of the random variables $|X_1|, |X_2|, \ldots, |X_n|$ respectively. In particular by (2.8)

$$R_2 + \varphi(n/(n+1)) - \varphi(1/(n+1))$$

$$\text{(A.8)} \qquad = n\int E|g(nK_n(K^{-1}(t))) - \varphi(t)|dt + \varphi(n/(n+1)) - \varphi(1/(n+1))$$

$$\leq 8n^{\frac{1}{2}}J(\varphi).$$

The last inequality follows from Hoeffding's proof. Combining (A.6) through (A.8) yields (A.2), which completes the proof.

The asymptotic normality of signed rank statistics will be established by approximating them by asymptotically equivalent sums of independent random variables. For this purpose define

$$Z_n^+(\varphi) = \sigma_n^{-1}[\Sigma_1^n c_i \int \varphi(t)d[I(|X_i| \leq K^{-1}(t))s(X_i) - \mu_i^+(t)]$$

$$\text{(A.9)} \qquad + n^{\frac{1}{2}}\int \hat{c}_n^+(t)W_n(t)d\varphi(t)],$$

where $\hat{c}_n^+$ is defined by (2.12). Obviously $Z_n^+(\varphi)$ is a sum of independent random variables.

Lemma A.2. *Let* $\{a_n(i)\}$ *and* $\{c_i\}$ *satisfy* (1.4) *and* (2.2). *Let* φ *be* $\searrow$ *and bounded. Assume there exist measurable functions on* $(0,1)$, $\{\ell^+_{ink},\ k = 1,2,\ 1 \le i \le n\}$ *such that* $\forall\, d,\ 0 < d < \infty$, and $\forall\ \eta > 0$, $\lim_{n\to\infty} \varphi(E_n) = 0$ *where*

$$E_n = \{s;\ \max_{k=1,2}\ \max_{1\le i\le n}\ \sup_{|t-s|\le dn^{-\frac12}} n^{\frac12}|L^+_{ink}(t) - L^+_{ink}(s) - (t-s)\ell^+_{ink}(t)| > \eta\}.$$

Define $T^+_n(\varphi)$ *by* (5.1).

(A.10) *Then* $|T^+_n(\varphi) - Z^+_n(\varphi)| = o_p(1)$.

Proof: As in (5.3) we have

$$T^+_n(\varphi) = -\int_0^1 T^+_n(t)d\varphi(t) + T^+_n(1)\varphi(1) + o_p(1)$$

where $T^+_n(t)$ is defined by (4.12). By integration by parts

$$Z^+_n(\varphi) = -\int Z^+_n d\varphi + \sigma_n^{-1}\{\Sigma_1^n c_i[s(X_i) - \mu^+_i(1)]\}\varphi(1)$$

where $Z^+_n(t)$ is defined by (4.13) on $(0,1)$. Then

$$T^+(\varphi) - Z^+(\varphi) = \int_0^1 (Z^+(t) - T^+(t))d\varphi(t) + o_p(1)\ .$$

Define $\bar{\mu}^*_n(t)$ by (4.43). By the decompositions (4.21) and (4.26)

$$\begin{aligned} T^+(\varphi) - Z^+(\varphi) &= \int[-V_n(K(K_n^{-1}(t))) + V_n(t)]d\varphi(t) \\ &\quad - \sigma_n^{-1}\int[\bar{\mu}^*_n(t) - n\hat{c}^+_n(t)(K(K_n^{-1}(t)) - t)]d\varphi(t) \\ &\quad - \sigma_n^{-1}\int n^{\frac12}\hat{c}^+_n(t)\{W_n(t) + n^{\frac12}(K(K_n^{-1}(t)) - t)\}d\varphi(t) \qquad \text{(A.11)} \\ &\quad + o_p(1) \\ &= R_1 - R_2 - R_3 + o_p(1) \quad \text{(say)}. \end{aligned}$$

By (4.22) and the boundedness of φ it follows that

(A.12) $$|R_1| = o_p(1) .$$

By (4.23), (4.44), and the boundedness of φ,

(A.13) $$|R_3| = o_p(1) .$$

Now to show $|R_2| = o_p(1)$ define $c_i^+ = c_i I(c_i \geq 0)$ and $c_i^- = c_i I(c_i < 0)$. Also define B_n and d by (4.24). Observe on B_n that

$$\|K(K_n^{-1}(t)) - t\| \leq dn^{-\frac{1}{2}}$$

and

$$\begin{aligned}
&\sigma_n^{-1}\|\bar{\mu}_n^*(t)\| \\
&= \sigma_n^{-1}\|\Sigma_1^n c_i(\mu_{in}^+(K(K_n^{-1}(t)) - \mu_{in}^+(t))\| \\
&\leq \sigma_n^{-1}\|\Sigma_1^n c_i^+[L_{in}^+(K(K_n^{-1}(t))) - L_{in}^+(t)]\| \\
&+ \sigma_n^{-1}\|\Sigma_1^n c_i^-[L_{in}^+(K(K_n^{-1}(t))) - L_{in}^+(t)]\| \\
&\leq 2\sigma_n^{-1} \max_{1\leq i\leq n} |c_i| dn^{+\frac{1}{2}} .
\end{aligned}$$

The second inequality follows from (2.8) and the first from the fact that for $t > s$, $L_{ink}^+(t) - L_{ink}^+(s) \geq 0$, $k = 1,2$, $i \leq n$. Therefore on B_n

$$|R_2| \leq \int_{E_n} \sigma_n^{-1}[2 \max_{1\leq i\leq n} |c_i| dn^{\frac{1}{2}} + dn^{\frac{1}{2}}\hat{c}_n^+(t)]d\varphi + \int_{E_n^c} \eta d\varphi .$$

The first term on the right hand side is $o(1)$ since $\lim_{n\to\infty} \varphi(E_n) = 0$. Since φ is bounded the second term can be made arbitrarily small by making η arbitrarily small. Finally since $P(B_n) \to 1$ as $n \to \infty$

(A.14) $$|R_2| = o_p(1) .$$

Combining (A.11) through (A.14) yields (A.10).

Lemma A.3. *Let* $\{a_n(i)\}$ *and* $\{c_i\}$ *satisfy* (1.4) *and* (2.2). *Let* φ *be* $\searrow$, *absolutely continuous and satisfy the Hoeffding's condition*

$$\int |\varphi(t)| (t(1-t))^{-\frac{1}{2}} dt < \infty .$$

(A.15) *Then* $E(T_n^+(\varphi) - Z_n^+(\varphi))^2 \to 0$ *as* $n \to \infty$.

Proof:

By Lemma 1 of Hoeffding (1973), φ can be decomposed into

$$\varphi(t) = \gamma_1(t) + \gamma_2(t) - \gamma_3(t)$$

where γ_1 is a polynomial and γ_2 and γ_3 are $\searrow$ and

(A.16) $$J(\gamma_2) + J(\gamma_3) < \alpha$$

where $J(\gamma_m)$, $m = 2,3$, is defined by (A.1).

Now

$$E(T^+(\varphi) - Z^+(\varphi))^2 \le 2E(\sigma_n^{-1}(S^+(\varphi) - ES^+(\varphi)) - Z^+(\varphi))^2$$

(A.17) $$+ 2\sigma_n^{-2}(ES^+(\varphi) - \mu^+(\varphi))^2$$

$$= 2R_1 + 2R_2 \qquad \text{(say)}.$$

$$|R_1| \le 3E(\sigma_n^{-1}(S^+(\psi_1) - ES^+(\psi_1)) - Z^+(\psi_1))^2$$

$$+ 6\Sigma_{m=2}^{3}[\sigma_n^{-2}E(S^+(\psi_m) - ES^+(\psi_m))^2 + E(Z^+(\psi_m))^2]$$

Since ψ_1 has a bounded second derivative the first term on the right hand side goes to zero as $n \to \infty$ by applying Lemma 2 of Husková (1970). The second term goes to zero by applying Theorem 4 and Lemma 5 of Husková (1970). Therefore

$$(A.18) \qquad |R_1| \to 0 \quad \text{as} \quad n \to \infty$$

$$(A.19) \qquad \begin{aligned} &|ES^+(\varphi) - \mu^+(\varphi)| \\ &\le \Sigma_1^3 |ES^+(\gamma_m) - \mu^+(\gamma_m)| \\ &= R_{21} + R_{22} + R_{23} \quad \text{(say)}. \end{aligned}$$

For $m = 22,23$

$$(A.20) \qquad \begin{aligned} R_m &= |\Sigma_1^n c_i \{E\gamma_m(R_i^+/(n+1))s(X_i) - \int \gamma_m(t)d\mu_i^+(t)\}| \\ &\le \max_{1\le i\le n} |c_i| \Sigma_1^n |E\gamma_m(R_i^+/(n+1))s(X_i) - \int \gamma_m(t)d\mu_i^+(t)| \ . \end{aligned}$$

Combining (A.2), (A.16) and (A.20) yields

$$(A.21) \qquad R_{22} + R_{23} < 8 \max_{1\le i\le n} |c_i| n^{1/2}\alpha \ .$$

Since γ_1 has a bounded second derivative, it follows from Lemma 2 of Husková (1970) that there exists K dependent only on γ_1 such that

$$(A.22) \qquad R_{21} \le Kn^{-1/2}\sigma \ .$$

Combining (A.19), (A.21), and (A.22), yields

$$|ES^+(\varphi) - \mu^+(\varphi)|/\sigma \le n^{-1/2}K + 8k_c\alpha \ .$$

Finally for any $\varepsilon > 0$, α is selected so that $16k_c\alpha < \varepsilon$. Now K is fixed. Then for $n \ge (2K/\varepsilon)^2$, we have

$$|ES^{+}(\varphi) - \mu^{+}(\varphi)|/\sigma < \varepsilon.$$

Therefore

(A.23) $$|R_2| \to 0 \quad \text{as} \quad n \to \infty .$$

Combining (A.17), (A.18) and (A.23) yields (A.15). This completes the proof of Lemma A.3.

In the following propositions φ is assumed to satisfy Hoeffding's condition on $(0,1)$. To simplify the notation we will assume that φ is non-decreasing on $(0,1)$. This can be done without loss of generality since φ may be written

$$\varphi = \varphi_1 - \varphi_2 \quad \text{where} \quad \varphi_1 \quad \text{and} \quad \varphi_2 \quad \text{are} \quad \searrow$$

and satisfying Hoeffding's condition on $(0,1)$.

Decompose φ into

(A.24) $$\varphi = \Phi_1 + \Phi_2$$

where Φ_1 and Φ_2 are $\searrow$ and satisfy Hoeffding's condition on $(0,1)$; Φ_1 is the absolutely continuous part of φ and satisfies $\int_{B^c} d\Phi_1 = 0$; Φ_2 is the singular part of φ and satisfies $\int_B d\Phi_2 = 0$; and B^c is a φ-measurable set containing the singular set of φ.

<u>Proposition A.1</u>. <u>Let</u> $\{a_n(i)\}$ <u>and</u> $\{c_i\}$ <u>satisfy</u> (1.4) <u>and</u> (2.2) <u>and</u> φ <u>be</u> $\searrow$ <u>and satisfy Hoeffding's condition</u>

(A.25) $$\int_0^1 |\varphi(t)|\,(t(1-t))^{-\frac{1}{2}}dt < \infty .$$

<u>Let</u> φ <u>be decomposed into</u> Φ_m, $m = 1,2$. <u>Assume that there exist measurable functions</u>

$\{\ell^+_{ink}\}$, $i = 1,\ldots,n$, $n \geq 1$, $k = 1,2$, on $(0,1)$ such that $\forall$ d, $0 < d < \infty$, and $\forall$ $\eta > 0$ and $k = 1,2$,

$$\text{(A.26)} \quad \lim_n \Phi_2\{s;\ \max_{1\leq i\leq n}\ \sup_{|t-s|\leq dn^{-\frac{1}{2}}} n^{\frac{1}{2}}|L^+_{ink}(t) - L^+_{ink}(s) - (t - s)\ell^+_{ink}(s)| > \eta\} = 0\ .$$

Define

$$\hat{c}^+_n(t) = n^{-1}\Sigma^n_1 c_i(\ell^+_{in1}(t) - \ell^+_{in2}(t)),\ t \in B^c$$

and

$$\hat{c}^+_n(t) = n^{-1}\ {}^n_1 c_i L^{+'}_i(t),\ t \in B$$

when the derivative $L^{+'}_i(t)$ exists.

Then

$$\text{(A.27)} \qquad \liminf_n \tau^2_{n+}(\varphi)\sigma^{-2}_n > 0$$

implies

$$\text{(A.28)} \qquad \tau^{-1}_{n+}(\varphi)[S^+_n(\varphi) - \bar{\mu}^+_n(\varphi)] \underset{D}{\to} N(0,1) \quad \text{r.v.}$$

Remark 8. Since the L^+_{ni} are absolutely continuous and (A.26) holds, $\hat{c}^+_n(t)$ is well defined almost everywhere with respect to φ. Consequently $\tau_{n+}(\varphi)$ is well defined. Also (2.8) implies $n^{-1}\Sigma^n_1 L^{+'}_i(t) = 1$.

Proof of Proposition A.1. Given $\alpha > 0$, there exists a decomposition

(A.29) $$\varphi = \Phi_1 + \psi_1 + \psi_2$$

where

$$\psi_1(t) = \begin{cases} \Phi_2(b) & 0 < t < b \\ \Phi_2(t) & b \le t \le 1 - b \\ \Phi_2(1-b) & 1 - b < t < 1, \end{cases}$$

$\psi_2(t) = \Phi_2(t) - \psi_1(t)$,

Φ_1 and Φ_2 are defined by (A.24) and

b is chosen so that $\int_0^1 t^{\frac{1}{2}}(1 - t)^{\frac{1}{2}} d\psi_2 < \alpha$ and

$$\int_0^1 \psi_2^2 dt < \alpha^2.$$

The choice of b is possible because φ satisfies Hoeffding's condition and is therefore square integrable. Now ψ_1 is bounded on $(0,1)$ and satisfies the conditions of Lemma A.2. Φ_1 is absolutely continuous and satisfies the conditions of Lemma A.3.

Define $T^+(\Phi_1)$, $Z^+(\Phi_1)$, $T^+(\psi_m)$ and $Z^+(\psi_m)$, $m = 1,2$, as in (5.1) and (A.9).

(A.30) $$E|T^+(\psi_2) - Z^+(\psi_2)|^2 = 4\sigma^{-2}E|S^+(\psi_2) - ES^+(\psi_2)|^2$$

$$+ 4\sigma^{-2}|ES^+(\psi_2) - \bar{\mu}^+(\psi_2)|^2 + 2E(Z^+(\psi_2))^2 = 4R_1 + 4R_2 + 2R_3 \text{ (say)}.$$

By Theorem 4 in Husková (1970)

(A.31) $$R_1 \le 80\sigma_n^{-2} n \max_{1 \le i \le n} |c_i|^2 \int_0^1 \psi_2^2 dt \le 80k_c^2\alpha^2 .$$

Also by (A.2)

$$R_2 \le \sigma^{-2} \max_{1\le i\le n} |c_i|^2 (\Sigma_1^n |E\psi_2(R_i^+/(n+1))s(X_i) - \int_0^1 \psi_2(t)d\mu_i^+(t)|)^2$$

(A.32) $$\le 64\sigma^{-2} n\alpha^2 \max_{1\le i\le n} |c_i^2|$$

$$\le 64 k_c^2 \alpha^2 .$$

Now Lemma 5 of Husková (1970) implies there exists K such that

(A.33) $$R_3 = \mathrm{Var}(Z^+(\psi_2))^2 \le K\ k_c^2\alpha^2 .$$

Combining (A.30) through (A.33) yields

(A.34) $$E|T^+(\psi_2) - Z^+(\psi_2)|^2 \le (320 + 2K + 256)k_c^2\alpha^2 .$$

Since α can be made arbitrarily small, (A.34) implies

(A.35) $$|T^+(\psi_2) - Z^+(\psi_2)| = o_p(1).$$

Since Lemmas A.2 and A.3 imply $|T^+(\Phi_1) - Z^+(\Phi_1)| = o_p(1)$ and $|T^+(\psi_1) - Z^+(\psi_1)| = o_p(1)$, we have by (A.29) and (A.35)

(A.36) $$|T^+(\varphi) - Z^+(\varphi)| = o_p(1) .$$

Therefore $T^+(\varphi)$ has the same asymptotic distribution as $Z^+(\varphi)$.

If φ is bounded, write

$$Z_n^+(\varphi) = -\sigma_n^{-1}\Sigma_1^n\{\int[(c_i - \bar{c}_n^+)(I(0 \le X_i \le K^{-1}(t)) - L_{in1}(t))$$

$$- (c_i + \hat{c}_n^+)(I(-K^{-1}(t) \le X_i < 0) - L_{in2}(t))]d\varphi(t)$$

$$- c_i[s(X_i) - \mu_i^+(1)]\varphi(1)\}$$

$$= -\Sigma_1^n Z_{ni}^+(\varphi) \qquad \text{(say)}.$$

Since $\|\hat{c}_n^+\| \leq \max_{1\leq i\leq n} |c_i|$ and φ is bounded, we have

$$\max_{1\leq i\leq n} |Z_{ni}^+(\varphi)| \leq 5\|\varphi\|\sigma_n^{-1} \max_{1\leq i\leq n} |c_i| = o(1).$$

Therefore by (A.27), the Lindeberg-Feller theorem yields that $\tau_{n+}^{-1}(\varphi)\sigma_n Z_n^+(\varphi)$ is asymptotically $N(0,1)$. Then by (A.36) for φ bounded

(A.37) $\quad \tau_{n+}^{-1}(\varphi)\sigma_n T^+(\varphi)$ is asymptotically $N(0,1)$.

If φ is not bounded but is $\searrow$ and satisfies Hoeffding's condition, decompose φ into $\varphi = \varphi_1 + \varphi_2$ where φ_1 is bounded and $\int\varphi_2^2(t)dt < \varepsilon$. Then by (A.33)

(A.38) $\quad E(Z^+(\varphi) - Z^+(\varphi_1))^2 = \mathrm{Var}(Z^+(\varphi_2)) \leq Kk_c^2\varepsilon$

and

(A.39) $\quad ((\tau_{n+}(\varphi_1)/\tau_{n+}(\varphi)) - 1)^2 \leq [\tau_{n+}(\varphi_2)/\tau_{n+}(\varphi)]^2 \leq Kk_c^2\varepsilon\tau_{n+}^{-2}(\varphi)$.

Therefore for large n if ε is sufficiently small $\tau_{n+}(\varphi_1)/\tau_{n+}(\varphi)$ will be as close to 1 as we want.

Therefore combining (A.38) and (A.39) yields

(A.40) $\quad |Z^+(\varphi)\tau_{n+}^{-1}(\varphi) - Z^+(\varphi_1)\tau_{n+}^{-1}(\varphi_1)| = o_p(1)$.

Since φ_1 is bounded (A.27) and (A.40) imply $\tau_{n+}^{-1}(\varphi)\sigma_n Z^+(\varphi)$ is asymptotically $N(0,1)$. Then by (A.36), (A.37) holds for φ not bounded which completes the proof of (A.30).

Remark 9 . If $\varphi = \varphi_1 - \varphi_2$ where φ_1 and φ_2 are $\searrow$ and satisfy Hoeffding's condition, the same proof when applied to $T^+(\varphi_1)$ and $T^+(\varphi_2)$ will prove Proposition A.1 for φ which satisfy Hoeffding's condition.

Proposition A.2. Let $\{a_n(i)\}$ and $\{c_i\}$ satisfy (1.4) and (2.2) and φ be $\searrow$ and satisfy Hoeffding's condition (A.25). Let φ be decomposed into Φ_k, $k = 1,2$, defined by (A.24). Assume that there exist measurable functions $\{\ell_{in}\}$, $i = 1,\ldots,n$, $n \geq 1$, on $(0,1)$ such that $\forall\, d$, $0 < d < \infty$, and $\forall\, \eta > 0$,

$$\lim_{n\to\infty} \Phi_2\{s;\ \max_{1\leq i\leq n}\ \sup_{|t-s|\leq dn^{-\frac{1}{2}}} n^{\frac{1}{2}}|L_{in}(t) - L_{in}(s) - (t-s)\ell_{in}(s)| > \eta\} = 0 \ .$$

Define

$$\hat{c}(t) = n^{-1}\Sigma_1^n c_i \ell_{in}(t),\ t \in B^c$$

and

$$\hat{c}(t) = n^{-1}\Sigma_1^n c_i L_i'(t),\ t \in B$$

when the derivative $L_i'(t)$ exists.

Define $\mu(\varphi)$ and $\tau_n^2(\varphi)$ by (2.16) and (2.19) respectively. Then

$$\liminf_n \tau_n^2(\varphi)\sigma_n^{-2} > 0$$

implies

$$\tau_n^{-1}(\varphi)[S_n(\varphi) - \mu(\varphi)] \underset{D}{\rightarrow} N(0,1) \quad \text{r.v.}$$

The proof is similar to Proposition A.1 with appropriate changes and is not presented here.

REFERENCES

REFERENCES

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

Braun, H.I. (1976). Weak convergence of sequential linear rank statistics. Ann. Statist. 4 554-575.

Dupač, V. (1969). A contribution to the asymptotic normality of simple linear rank statistics. Nonparametric Techniques in Statistical Inference, ed. M.L. Puri, Cambridge University Press 75-88.

Dupač, V. and Hájek, J. (1969). Asymptotic normality of simple linear rank statistics under alternatives II. Ann. Math. Statist. 40 1992-2017.

Fernandez, P.J. (1970). A weak convergence theorem for random sums of independent random variables. Ann. Math. Statist. 41 710-712.

Ghosh, M. and Sen, P. (1972). On bounded length confidence intervals for the regression coefficient based on a class of rank statistics. Sankhyā: The Indian Journal of Statistics, Series A. 34 33-52.

Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist. 39 325-346.

Hoeffding, W. (1973). On the centering of a simple linear rank statistics. Ann. Statist. 1 54-66.

Husková, M. (1970). Asymptotic distribution of simple linear rank statistics for testing symmetry. Z. Wahrscheinlichkeitstheorie verw. Geb. 14 308-322.

Koul, H. (1970). Asymptotic normality of random rank statistics. Ann. Math. Statist. 41 2144-2149.

Koul, H. (1972) Some asymptotic results on random rank statistics. Ann. Math. Statist. 43 842-859.

Koul, H. (1974). Asymptotic normality under alternatives of linear rank statistics. RM-316, Department of Statistics and Probability, Michigan State University.

Koul, H. and Staudte, R. (1972a). Asymptotic normality of signed rank statistics. Z. Wahrscheinlichkeitstheorie verw. Geb. **22** 293-300.

Koul, H. and Staudte, R. (1972b) Weak convergence of weighted empirical cumulatives based on ranks. Ann. Math. Statist. **43** 832-841.

Loève, M. (1963). Probability Theory (3rd ed.). Van Nostrand, Princeton.

Pyke, R. and Shorack, G. (1968a). Weak convergence of a two sample empirical process and a new approach to Chernoff-Savage theorems. Ann. Math. Statist. **39** 755-771.

Pyke, R. and Shorack, G. (1968b). Weak convergence and a Chernoff-Savage theorem for random sample size. Ann. Math. Statist. **39** 1675-1685.

Sen, P. and Ghosh, M. (1971). On bounded length sequential confidence intervals based on one-sample rank order statistics. Ann. Math. Statist. **42** 189-203.

Sen, P. and Ghosh, M. (1972). On strong convergence of regression rank statistics. Sankhyā: The Indian Journal of Statistics, Series A. **34** 335-348.

Shorack, G. (1973). Convergence of reduced empirical and quantile processes with applications to functions of order statistics in the non-iid case. Ann. Statist. **1** 146-152.