

AV9U 180

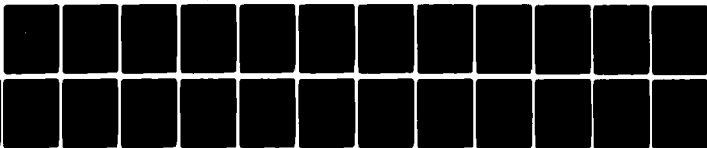
INDIANA UNIV AT BLOOMINGTON DEPT OF MATHEMATICS F/G.12/1
ASYMPTOTIC NORMALITY OF SIGNED RANK STATISTICS WITH DISCONTINUO--ETC(U)
AUG 80 M L PURI AFOSR-76-2927

AFOSR-TR-80-0869

NL

UNCLASSIFIED

[of]
ADD
000-80



END
DATE
FILMED
11-80
DTIC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

4

1. REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS - BEFORE COMPLETING FORM	
18	REPORT NUMBER AFOSR-TR-80-0869	2. GOVT ACCESSION NO. AD-A090180	3. RECIPIENT'S CATALOG NUMBER
6	TITLE (and Subtitle) Asymptotic normality of signed rank statistics with "discontinuous" score generating function.	5. TYPE OF REPORT & PERIOD COVERED Interim rept.	
7	AUTHOR(s) Madan L. Puri	6. PERFORMING ORG. REPORT NUMBER	
10	PERFORMING ORGANIZATION NAME AND ADDRESS Indiana University Department of Mathematics Bloomington, Indiana 47401	8. CONTRACT OR GRANT NUMBER(s) ✓ AFOSR-76-2927	
11	CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6110ZF 2304/A5	
12	MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	12. REPORT DATE August 1980	
12	28	13. NUMBER OF PAGES 26	
LEVEL II		15. SECURITY CLASS. (of this report) UNCLASSIFIED	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Asymptotic normality, signed rank statistic, score-generating function;			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let X_{N1}, \dots, X_{NN} , $N \geq 1$ be independent random variables with continuous distribution functions F_{N1}, \dots, F_{NN} ; and let R_{Ni}^+ be the rank of $ X_{Ni} $ among $ X_{N1} , \dots, X_{NN} $. Under different sets of conditions, we derive the asymptotic normality of simple linear signed rank statistic $S_N^+ = \sum_{i=1}^N c_{Ni} a_N(R_{Ni}^+) \text{sgn } X_{Ni}$ where (over)			

AD A090180

JDC FILE COPY

EXCLUDED
OCT 14 1980

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. (cont.)

C_{N1}, \dots, C_{NN} are known regression constants, $\text{sgn } x = 1$ if $x \geq 0$, $\text{sgn } x = -1$ if $x < 0$, and $a_{N(1)}, \dots, a_{N(N)}$ are scores generated by a function $\psi(t)$, $0 < t < 1$ which in contradistinction to the earlier literature is no longer assumed to be continuous. The results obtained are generalizations of the earlier results on limit theorems due to Hájek (1968, Ann. Math. Statist. 325-346) and Hušková (1970, Z. Wahrscheinlichkeitstheorie. Verw. Geb., 308-322), among others.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Special and/or
A	

Unclassified

Asymptotic normality of signed rank statistics with discontinuous score-generating function.

By Madan L. Puri

Indiana University, Bloomington

Dedicated to Professor C.R. Rao on the occasion of his 60th birthday.

The object of this paper is to derive the asymptotic distributions of simple linear signed rank statistic considered by Hájek (1968) and Hušková (1970) for the case when the score generating function is discontinuous.

1. Preliminaries

Let X_{N1}, \dots, X_{NN} , $N \geq 1$ be independent random variables with continuous distribution functions F_{N1}, \dots, F_{NN} ; and let R_{Ni}^+ be the rank of $|X_{Ni}|$ among $|X_{N1}|, \dots, |X_{NN}|$. We shall be concerned with the asymptotic distribution of the statistic

$$S_N^+ = \sum_{i=1}^N c_{Ni} a_N(R_{Ni}^+) \operatorname{sgn} X_{Ni} \quad (1.1)$$

where c_{N1}, \dots, c_{NN} are known regression constants; $a_N(1), \dots, a_N(N)$ are scores, and $\operatorname{sgn} x = 1$ if $x \geq 0$, $\operatorname{sgn} x = -1$ if $x < 0$. The results that we obtain are derived using the projection method together with a separate study of the case of the score generating function which has just one jump and is constant otherwise (see

Work partially supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR-76-2927. Reproduction in whole or in part permitted for the purpose of the United States Government.

Approved for public release
distribution unlimited.

80 9 22 157

**AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC**

This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.

A. D. BLOSE
Technical Information Officer

also Dupač and Hájek (1969).

We assume that the c_{Ni} 's satisfy the Noether condition

$$\lim_{N \rightarrow \infty} \max_{1 \leq i \leq N} c_{Ni}^2 / \sum_{i=1}^N c_{Ni}^2 = 0 \quad (1.2)$$

and the scores are generated by a function $\psi(t)$, $0 < t < 1$, either by interpolation

$$a_N(i) = \psi(i/(N+1)), \quad 1 \leq i \leq N \quad (1.3)$$

or by a procedure satisfying

$$\sum_{i=1}^N |a_N(i) - \psi(i/(N+1))| = O(1) \quad (1.4)$$

Set

$$H_N^*(x) = N^{-1} \sum_{i=1}^N F_{Ni}^*(x) \quad (1.5)$$

where F_{Ni}^* is the distribution function of $|X_i|$.

$$H_N^{*-1}(t) = \inf \{x : H_N^*(x) \geq t\}, \quad 0 < t < 1 \quad (1.6)$$

$$L_{Ni}(t) = F_{Ni}^*(H_N^{*-1}(t)), \quad 0 < t < 1 \quad (1.7)$$

$$M_{Ni}(t) = -F_{Ni}^*(-H_N^{*-1}(t)), \quad 0 < t < 1 \quad (1.8)$$

$$G_{Ni}(t) = F_{Ni}^*(H_N^{*-1}(t)) = L_{Ni}(t) + M_{Ni}(t), \quad 0 < t < 1 \quad (1.9)$$

It is easy to check that

$$N^{-1} \sum_{i=1}^N G_{Ni}(t) = t, \quad 0 \leq t \leq 1, \quad N > 1 \quad (1.10)$$

The price for allowing discontinuous scores is the following differentiability conditions:

Let v denote a jump point of the score generating function ψ ; $v \in (0,1)$. Then, for every $K > 0$, $K' > 0$, we impose the following conditions:

$$\max_{1 \leq i \leq N} \sup_{KN^{-\frac{1}{2}} \leq |t-v| \leq K'N^{-\frac{1}{2}}} Lg^{\frac{1}{2}} N \left| \frac{L_{Ni}(t) - L_{Ni}(v)}{t-v} \right| = o(1) \quad (1.11)$$

$$\max_{1 \leq i \leq N} \sup_{KN^{-\frac{1}{2}} \leq |t-v| \leq K'N^{-\frac{1}{2}}} Lg^{\frac{1}{2}} N \left| \frac{M_{Ni}(t) - M_{Ni}(v)}{t-v} \right| = o(1) \quad (1.12)$$

Furthermore, we assume that there exist real numbers $l_{Ni}(v)$, $m_{Ni}(v)$, ($1 \leq i \leq N$) such that for every $K > 0$,

$$\max_{1 \leq i \leq N} \sup_{|t-v| \leq KN^{-\frac{1}{2}}} |L_{Ni}(t) - L_{Ni}(v) - (t-v) l_{Ni}(v)| = o(N^{-\frac{1}{2}}) \quad (1.13)$$

$$\max_{1 \leq i \leq N} \sup_{|t-v| \leq KN^{-\frac{1}{2}}} |M_{Ni}(t) - M_{Ni}(v) - (t-v) m_{Ni}(v)| = o(N^{-\frac{1}{2}}) \quad (1.14)$$

From (1.11) - (1.14) it follows that $G_{Ni}(t)$ satisfies the same type of conditions with $g_{Ni}(v) = l_{Ni}(v) + m_{Ni}(v)$. Also it is easy to see that without loss of generality one can assume that $l_{Ni}(v) \geq 0$, $m_{Ni}(v) \geq 0$, and

$$N^{-1} \sum_{i=1}^N l_{Ni}(v) = 1 = N^{-1} \sum_{i=1}^N m_{Ni}(v)$$

Finally, note that the numbers $l_{Ni}(v)$, $m_{Ni}(v)$ considered as function of (i, N) , $1 \leq i \leq N$, are bounded. Another condition concerning the G_{Ni} 's that we use is

$$\liminf_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N G_{Ni}(v)(1 - G_{Ni}(v)) > 0 \quad (1.15)$$

Some times, mainly for purposes of applications, we replace (1.11) - (1.15) by more feasible conditions easier to verify:

Suppose, for example, that each F_{Ni} has a density f_{Ni} such that for some α , $(0 < \alpha \leq \infty)$, we have

(a) $f_{Ni}(x) = 0$ for $x \notin (-\alpha, \alpha)$ if α is finite,

(b) $f_{Ni}(x)$ is continuous on every compact sub interval of $(-\alpha, \alpha)$ uniformly in (x, N, i) ,

(1.16)

(c) for every compact interval $C \subset (0, \alpha)$ there exists an $\epsilon_C > 0$ such that for all $N \geq 1$,

$$N^{-1} \text{Card} \{1 \leq i \leq N : \inf_{x \in C} f_{Ni}^*(x) > \epsilon_C\} > \epsilon_C,$$

where f_{Ni}^* is the density of F_{Ni}^* .

(d) $0 < \liminf_{N \rightarrow \infty} H_N^{*-1}(t) \leq \limsup_{N \rightarrow \infty} H_N^{*-1}(t) < \alpha$ for all $t \in (0, 1)$.

We will see that the condition (1.16) is satisfied in particular if

- (a) $f_{N_i}(x) = e^{-d_{Ni}} f(x e^{-d_{Ni}})$
- (b) f is uniformly continuous and positive on $(-\infty, \infty)$
- (c) $\sup_N \max_{1 \leq i \leq N} |d_{Ni}| < \infty$
- (1.17)

The last condition that we require concerns the nondegeneration of $\text{Var } S_N^+$ in the form

$$\liminf_{N \rightarrow \infty} \text{Var}(S_N^+) / \sum_{i=1}^N c_{Ni}^2 > 0 \quad (1.18)$$

2. Main Theorems.

Theorem 2.1. Consider the Statistic (1.1) with scores satisfying (1.4), where

$$\psi(t) = \begin{cases} 0 & \text{if } 0 < t < v \\ 1 & \text{if } v \leq t < 1 \end{cases}$$

Then S_N^+ is asymptotically normal with natural parameters $(E(S_N^+), \text{Var}(S_N^+))$ if any of the following sets of conditions is satisfied:

$$(C_1^+) : (1.2), (1.11), (1.12), (1.13), (1.14), (1.15), (1.18)$$

$$(C_2^+) : (1.2), (1.16), (1.18)$$

$$(C_3^+) : (1.2), (1.17), (1.18)$$

Proof. We show that S_N^+ is asymptotically equivalent to its projection \hat{S}_N^+ onto the space of linear statistics and then that \hat{S}_N^+ is asymptotically equivalent to a sum of independent random variables to which the Lindeberg Central Limit theorem applies. (For the ease of convenience we shall from now suppress the first subscript N from X_{Ni} , R_{Ni}^+ , etc).

First we would like to derive an upper bound for the residual variance $E(S_N^+ - \hat{S}_N^+)^2$, where

$$\hat{S}_N^+ = \sum_{i=1}^N E(S_N^+ | X_i) - (N-1)E(S_N^+).$$

By the Residual Variance Lemma (see Hájek (1968)), we have

$$\begin{aligned} E(S_N^+ - \hat{S}_N^+)^2 &\leq \sum_{i=1}^N c_i^2 E(a(R_i^+) - E(a(R_i^+) | X_i))^2 \\ &+ \sum_{i \neq j} c_i c_j \{ E(\text{sgn } X_i \text{sgn } X_j \text{Cov}(a(R_i^+), a(R_j^+) | X_i, X_j)) \\ &+ E\{\text{sgn } X_i \text{sgn } X_j [E(a(R_i^+) | X_i, X_j) - E(a(R_i^+) | X_i)] \\ &\quad \times [E(a(R_j^+) | X_i, X_j) - E(a(R_j^+) | X_j)]\} \\ &- \sum_{k \neq i, j} \text{Cov}\{E(\text{sgn } X_i a(R_i^+) | X_k), E(\text{sgn } X_j a(R_j^+) | X_k)\} \end{aligned}$$

We investigate each term in the above inequality. We begin by assuming that the scores are defined by (1.3) and that (C_1^+) holds. The proof is divided in several steps:

Lemma 2.1. The functions $L_{Ni}(t)$, $M_{Ni}(t)$, $G_{Ni}(t)$ satisfy
the following relations:

$$|L_{Ni}(t) - L_{Ni}(s)| \leq N|t - s| ,$$

$$|M_{Ni}(t) - M_{Ni}(s)| \leq N|t - s| ;$$

and

$$|G_{Ni}(t) - G_{Ni}(s)| \leq N|t - s| ; 0 < s, t < 1$$

Proof: It follows from the definitions and the fact that for
 $u > v > 0$ we have:

$$F_i(u) - F_i(v) \leq F_i^*(u) - F_i^*(v) , F_i(-v) - F_i(-u) \leq F_i^*(u) - F_i^*(v)$$

where we set $u = H^{*-1}(t)$, $v = H^{*-1}(s)$.

$$\text{Denote } V = \{(N+1)v\} , D^2 = N^{-1} \sum_{i=1}^N G_i(v)(1 - G_i(v))$$

([.] = integer part)

Lemma 2.2. Let $x, y \in \mathbb{R}$. Then to each $k_1 > 2$ there exist a
 $k_2 > 1$ such that for all $N > N_0(k_1)$ we have:

$$(i) \quad v - H_N^*(|x|) > k_1 N^{-\frac{1}{2}} \text{Lg}^{\frac{1}{2}} N \Rightarrow P(R_i^+ \geq V | X_i = x, X_j = y) < N^{-k_2}$$

$$(ii) \quad v - H_N^*(|x|) < -k_1 N^{-\frac{1}{2}} \text{Lg}^{\frac{1}{2}} N \Rightarrow P(R_i^+ < V | X_i = x, X_j = y) < N^{-k_2}$$

Furthermore, (i) and (ii) remain true even when the condition $X_j = y$ is omitted.

Lemma 2.3. Suppose that $|v - H^*(|x|)| \leq k_3 N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N$. Then,
for sufficiently large N , we have

$$(i) \quad \left| \sum_{i=1}^N F_i^*(|x|) (1 - F_i^*(|x|)) - ND^2 \right| \leq k_4 N^{\frac{1}{2}} Lg^{\frac{1}{2}} N$$

$$(ii) \quad \left| \phi(v; \sum_{i=1}^N F_i^*(|x|), \sum_{i=1}^N F_i^*(|x|) (1 - F_i^*(|x|))) \right. \\ \left. - \phi(Nv; \sum_{i=1}^N F_i^*(|x|), ND^2) \right| \leq k_5 N^{-1} Lg^{\frac{1}{2}} N$$

$$(iii) \quad \left| \phi(v; \sum_{i=1}^N F_i^*(|x|), \sum_{i=1}^N F_i^*(|x|) (1 - F_i^*(|x|))) \right. \\ \left. - \phi(Nv; \sum_{i=1}^N F_i^*(|x|), ND^2) \right| \leq k_6 N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N$$

where $\phi(x; \mu, \sigma^2)$, $\Phi(x; \mu, \sigma^2)$ denote the normal density, resp. the normal distribution function with parameters (μ, σ^2) .

The proofs of lemmas 2.2 and 2.3 are analogous to those of Lemmas 5 and 6 of Dupac^v and Hájek (1969), and are therefore omitted.

Lemma 2.4. For $N \rightarrow \infty$, we have

$$E(a(R_i^+) - E(a(R_i^+) | X_i))^2 = o(1)$$

uniformly in $1 \leq i \leq N$.

Proof: Let $\Omega^+(X_i) = E(a(R_i^+) | X_i) - [(E(a(R_i^+) | X_i))]^2$

Then, by conditioning we obtain:

$$E[a(R_i^+) - E(a(R_i^+) | X_i)]^2 = E(\Omega^+(X_i))$$

Now, by definition:

$$E(a(R_i^+) | X_i = x) = P(R_i^+ > v | X_i = x) .$$

Thus

$$\Omega^+(X_i = x) = P(R_i^+ > v | X_i = x) \cdot P(R_i^+ \leq v | X_i = x)$$

$$\text{Let } I = \{x \mid |H^*(|x|) - v| \leq k_1 N^{-\frac{1}{2}} \lg^{\frac{1}{2}} N\}$$

By Lemma 2, if $x \notin I$ we have:

$$\Omega^+(X_i = x) < N^{-k_2}, \text{ for every } N \geq N_0(k_2), k_2 > 1$$

On the other hand, if $x \in I$, then since $P(R_i^+ = k | X_i = x) = B^i(k, F_1^*(|x|), \dots, F_N^*(|x|))$ (in the notation used by Dupač and Hájek (1969)), we obtain using Lemmas 2.2 and 2.3, that

$$\begin{aligned} \Omega^+(X_i = x) &= \left\{ \sum_{k > v} B^i(k, F_1^*(|x|), \dots, F_N^*(|x|)) \right\} \\ &\quad \times \left\{ \sum_{\ell \leq v} B^i(\ell, F_1^*(|x|), \dots, F_N^*(|x|)) \right\} \end{aligned}$$

$$= \dots = \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right)\right\} + \theta_1 N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N$$

for sufficiently large N , $|\theta_1| \leq k_7$

We observe that the last equality remains true even if we enlarge I to

$$I' = \{x : |H^*(|x|) - v| \leq k_9 DN^{-\frac{1}{2}} Lg^{\frac{1}{2}} N\}$$

where k_9 is such that $k_9 = k_1/2k_8$ with $k_8 \leq D \leq \frac{1}{2}$ and $k_9 > 2$. (Here we use Condition (1.15)).

Now using (1.11) and (1.12) it is easy to show that

$$N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N \int_{I'} \theta_1 dF_i(x) = o(1)$$

and

$$\int_{I'} \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right)\right\} dF_i(x) = o(1)$$

Hence $E(\Omega^+(X_i)) = o(1)$ uniformly in $1 \leq i \leq N$ and the proof follows.

Lemma 2.5. For $N \rightarrow \infty$ we have:

$$E(\operatorname{sgn} X_i \operatorname{sgn} X_j [E(a(R_i^+) | X_i, X_j) - E(a(R_i^+) | X_i)] \cdot [E(a(R_j^+) | X_i, X_j)])$$

$$- E(a(R_j^+) | X_j)) = o(N^{-1})$$

uniformly in $1 \leq i, j \leq N$.

The proof of this Lemma is similar to that of Lemma 2.4 and is therefore omitted.

Lemma 2.6 For $N \rightarrow \infty$ we have:

$$\begin{aligned} E[\text{sgn } X_i \text{sgn } X_j \text{Cov}(a(R_i^+), a(R_j^+) | X_i, X_j)] \\ = N^{-1} D^2 (\ell_i(v) - m_i(v)) (\ell_j(v) - m_j(v)) + o(N^{-1}) \end{aligned}$$

uniformly in $1 \leq i, j \leq N$.

Proof: We have:

$$\begin{aligned} \Delta^+ &= \text{Cov}(a(R_i^+), a(R_j^+) | X_i = x, X_j = y) \\ &= \begin{cases} P(R_i^+ > v | X_i = x, X_j = y) \cdot P(R_j^+ \leq v | X_i = x, X_j = y) & \text{if } |x| < |y| \\ P(R_j^+ > v | X_i = x, X_j = y) \cdot P(R_i^+ \leq v | X_i = x, X_j = y) & \text{if } |x| \geq |y| \end{cases} \end{aligned}$$

Let $k_1 > 2$. Denote:

$$I = \{(x, y) : |H^*(|x|) - v| \leq k_1 N^{-\frac{1}{k_1}} Lg^{\frac{1}{k_1}} N, |H^*(|y|) - v| \leq k_1 N^{-\frac{1}{k_1}} Lg^{\frac{1}{k_1}} N\}$$

By considerations as used in the derivation of (4.11) and (4.12) in Dupač and Hájek (1969), we obtain

$$\begin{aligned}
 \Delta^+(x,y) & \left\{ \begin{aligned}
 & < N^{-k_2} \text{ for } (x,y) \notin I, N > N_0(k_2) \\
 & = \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|y|) - v}{DN^{-\frac{1}{2}}}\right)\right\} + \theta_2 N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N \\
 & \quad \text{for } N \text{ sufficiently large, } (x,y) \in I, |x| < |y|, \\
 & \quad \text{and } |\theta_2| \leq k_{10} \tag{2.1} \\
 & = \phi\left(\frac{H^*(|y|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right)\right\} + \theta_3 N^{-\frac{1}{2}} Lg^{\frac{1}{2}} N \\
 & \quad \text{for } N \text{ sufficiently large, } (x,y) \in I, |x| \geq |y| \\
 & \quad \text{and } |\theta_3| \leq k_{11}
 \end{aligned} \right.
 \end{aligned}$$

We note that the equality in (2.1) remains true even if we enlarge I to I'

$$I' = \{(x,y) : \max\{|H^*(|x|) - v|, |H^*(|y|) - v|\} \leq k_1' DN^{-\frac{1}{2}} Lg^{\frac{1}{2}} N\}$$

where k_1' is such that $k_1' = k_1/2k_8$, $k_8 \leq D \leq \frac{1}{2}$, $k_1' > 2$ (k_1' coincides, with k_9 in the notation of Lemma (2.4)).

We have, using (2.1) that

$$\begin{aligned}
 & E(\text{sgn } X_i \text{sgn } X_j \text{Cov}(a(R_i^+), a(R_j^+) | X_i, X_j)) \\
 & = \iint_{I \cap \{|x| < |y|\}} \text{sgn } x \text{sgn } y \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|y|) - v}{DN^{-\frac{1}{2}}}\right)\right\} dF_i(x) dF_j(y) \\
 & \quad + \iint_{I \cap \{|x| \geq |y|\}} \text{sgn } x \text{sgn } y \phi\left(\frac{H^*(|y|) - v}{DN^{-\frac{1}{2}}}\right) \left\{1 - \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{1}{2}}}\right)\right\} dF_i(x) dF_j(y)
 \end{aligned}$$

$$+ N^{-k} Lg^k N \int_{I'} \int \operatorname{sgn} x \operatorname{sgn} y \theta_4(x, y) dF_i(x) dF_j(y) + \theta_5 N^{-k_2} \quad (2.2)$$

with $|\theta_4| \leq k_{12}$, $|\theta_5| \leq 1$.

The last two terms are $o(N^{-1})$ uniformly in i, j as follows by using (1.11), (1.12) and $k_2 > 1$. It remains to estimate the first two terms.

Denote the first term by T . Consider:

$$T_1 = \int \int \phi\left(\frac{H^*(x) - v}{DN^{-k}}\right) \left\{1 - \phi\left(\frac{H^*(y) - v}{DN^{-k}}\right)\right\} dF_i(x) dF_j(y) \\ \{(x, y) : \begin{array}{l} x > 0 \\ y > 0 \\ x < y \end{array}, \max(|H^*(x) - v|, |H^*(y) - v|) \leq k_1' DN^{-k} Lg^k N\}$$

$$\text{Set } p = \frac{H^*(x) - v}{DN^{-k}}, \quad q = \frac{H^*(y) - v}{DN^{-k}} \text{ and}$$

$$I^* = \{(p, q) : \max(|p|, |q|) \leq k_1' Lg^k N\}. \text{ Then}$$

$$T_1 = \int \int_{I^* \cap \{p < q\}} \phi(p) (1 - \phi(q)) dL_i(v + DN^{-k} p) dL_j(v + DN^{-k} q)$$

As in the proof of Lemma 7 of Dupac^v and Hájek (1969) one can easily show that

$$T_1 = \frac{1}{2} N^{-1} D^2 \lambda_i(v) \lambda_j(v) + o(N^{-1})$$

uniformly in $1 \leq i, j \leq N$

(2.3)

Let

$$T_2 = \int \int -\phi\left(\frac{H^*(-x) - v}{DN^{-k}}\right) \left\{1 - \phi\left(\frac{H^*(y) - v}{DN^{-k}}\right)\right\} dF_i(x) dF_j(y) \\ \{(x, y) : \begin{array}{l} x < 0 \\ y > 0 \\ -x < y \end{array}, \max\{|H^*(-x) - v|, |H^*(y) - v|\} \leq k_1' DN^{-k} Lg^k N\}$$

Then

$$-T_2 = \iint_{I^n \cap \{p < q\}} \phi(p) (1 - \phi(q)) dM_1(v + DN^{-k}p) dL_j(v + DN^{-k}q)$$

Divide I^n into J and $I^n \setminus J$ where:

$$J = \{(p, q) : \max(|p|, |q|) \leq k_1^n\}$$

and in J make use of the expansions:

$$M_1(v + DN^{-k}p) = M_1(v) + m_1(v)DN^{-k}p + \Omega_1(p)$$

$$L_j(v + DN^{-k}q) = L_j(v) + l_j(v)DN^{-k}q + \Lambda_j(q)$$

where $\Omega_1(p)$ and $\Lambda_j(q)$ are absolutely continuous and are of order $o(N^{-k})$ in $[-k_1^n, k_1^n]$. This follows from (1.13) and (1.14).

Then considerations similar to the ones used in the derivation of (2.3) lead to:

$$T_2 = -kN^{-1}D^2m_1(v)l_j(v) + o(N^{-1}) \text{ uniformly in } 1 \leq i, j \leq N.$$

Consider now

$$T_3 = \iint \phi\left(\frac{H^*(x) - v}{DN^{-k}}\right) \{1 - \phi\left(\frac{H^*(-y) - v}{DN^{-k}}\right)\} dF_i(x) dF_j(y)$$

$$\{(x, y) : \begin{array}{l} x < 0 \\ y < 0 \\ -x < -y \end{array}, \max(|H^*(x) - v|, |H^*(-y) - v|) \leq k_1^n DN^{-k} Lg^{\frac{1}{2}} N\}$$

$$T_4 = \iint -\phi\left(\frac{H^*(x) - v}{DN^{-k}}\right) \left(1 - \phi\left(\frac{H^*(-y) - v}{DN^{-k}}\right)\right) dF_i(x) dF_j(y)$$

$$\{(x, y) : \begin{array}{l} x > 0 \\ y < 0 \\ x < -y \end{array} ; \max(|H^*(x) - v|, |H^*(-y) - v|) < k_1' DN^{-k} Lg^{\frac{1}{2}} N\}$$

Proceeding as before, (omitting the details of computations) it follows that

$$T_3 = kN^{-1} D^2 m_i(v) m_j(v) + o(N^{-1}) \quad \text{uniformly in } 1 \leq i, j \leq N$$

$$T_4 = -kN^{-1} D^2 l_i(v) m_j(v) + o(N^{-1}) \quad \text{uniformly in } 1 \leq i, j \leq N.$$

Thus,

$$T = T_1 + T_2 + T_3 + T_4 = kN^{-1} D^2 (l_i(v) - m_i(v))(l_j(v) - m_j(v))$$

$$+ o(N^{-1}) \quad \text{uniformly in } 1 \leq i, j \leq N. \quad (2.4)$$

Proceeding as above, it can be shown that the second term of (2.2) is the same as (2.4). The proof follows.

Lemma 2.7. For $N \rightarrow \infty$, we have (for $i \neq j$)

$$\sum_{k \neq i, j} \text{Cov} \{E(\text{sgn } X_i a(R_i^+) | X_k), E(\text{sgn } X_j a(R_j^+) | X_k)\} =$$

$$= N^{-1} D^2 (l_i(v) - m_i(v))(l_j(v) - m_j(v)) + o(N^{-1})$$

uniformly in $1 \leq i, j \leq N$.

Proof: By Lemma 3.2 in Hájek (1968) we have:

$$\begin{aligned} & E(a(R_i^+) \operatorname{sgn} X_i | X_i = x, X_k = z) - E(a(R_i^+) \operatorname{sgn} X_i | X_i = x) \\ &= \operatorname{sgn} x \{u(|x| - |z|) - F_k^*(|x|)\} P(R_i^+ = v+1 | X_i = x, |X_k| = |x| - 1) \end{aligned}$$

where $u(t) = 1$ for $t \geq 0$, and $u(t) = 0$ for $t < 0$.

From Lemma 2.2, we have

$$P(R_i^+ = v+1 | X_i = x, |X_k| = |x| - 1) < N^{-k_2} \quad (2.5)$$

for some $k_2 > 1$ and all $|H^*(|x|) - v| \geq k_1 N^{-k} Lg^k N$

Furthermore Lemmas 2.2 and 2.3 imply

$$\begin{aligned} P(R_i^+ = v+1 | X_i = x, |X_k| = |x| - 1) &= \phi(Nv; \sum_{j=1}^N F_j^*(|x|), ND^2) \\ &+ \theta_6 N^{-1} Lg^k N \end{aligned}$$

for some $|\theta_6| \leq k_{13}$ and all $|H^*(|x|) - v| \leq k_1 N^{-k} Lg^k N$.

As before, the last equality remains true even if

$$|H^*(|x|) - v| \leq k_1' DN^{-k} Lg^k N$$

Let $I' = \{x : |H^*(|x|) - v| \leq k_1' DN^{-k} Lg^k N\}$.

Then

$$\begin{aligned}
 & E(a(R_i^+) \operatorname{sgn} X_i | X_k = z) - E(a(R_i^+) \operatorname{sgn} X_i) = \\
 & = \int \operatorname{sgn} x [u(|x| - |z|) - F_k^*(|x|)] P(R_i^+ = v+1 | X_i = x, |X_k| = |x|-1) dF_i(x) \\
 & = \int_{I'} (\dots) dF_i(x) + \int_{\mathbb{R} \setminus I'} (\dots) dF_i(x)
 \end{aligned}$$

The second integral is $o(N^{-1})$ by (2.5), while the first is equal to

$$\begin{aligned}
 & \int_{I'} \operatorname{sgn} x [u(|x| - |z|) - F_k^*(|x|)] \phi(Nv, \sum_{j=1}^N F_j^*(|x|), ND^2) dF_i(x) \\
 & + \int_{I'} \operatorname{sgn} x [u(|x| - |z|) - F_k^*(|x|)] \theta_6 N^{-1} Lg^k N dF_i(x)
 \end{aligned}$$

In the last expression, let us denote by T_5 the first term and T_6 the second.

Then, we have

$$T_5 = D^{-1} N^{-\frac{k}{2}} \int_{I'} \operatorname{sgn} x [u(|x| - |z|) - F_k^*(|x|)] \phi\left(\frac{H^*(|x|) - v}{DN^{-\frac{k}{2}}}\right) dF_i(x)$$

Using the fact that

$$\begin{aligned}
 & D^{-1} N^{-\frac{k}{2}} \int_{|p| \leq k_1' Lg^{\frac{k}{2}} N} G_k(v + DN^{-\frac{k}{2}} p) \phi(p) dL_1(v + DN^{-\frac{k}{2}} p) \\
 & = D^{-1} N^{-\frac{k}{2}} \int_{|p| \leq k_1' Lg^{\frac{k}{2}} N} G_k(v) \phi(p) dL_1(v + DN^{-\frac{k}{2}} p) + o(N^{-1})
 \end{aligned}$$

uniformly in i and k (which follows from (1.11)), we can show that

$$D^{-1} N^{-k} \int [u(x - |z|) - F_k^+(x)] \phi\left(\frac{H^*(x) - v}{DN^{-k}}\right) dF_i(x) \\ (x > 0 : |H^*(x) - v| \leq k_1' DN^{-k} Lg^k N)$$

$$= N^{-1} l_1(v) [1 - \phi(q) - G_k(v)] + o(N^{-1})$$

uniformly in z , $1 \leq i \leq N$, where $q = \frac{H^*(|z|) - v}{DN^{-k}}$

Similarly we estimate the integral over $\{x < 0 : |H^*(-x) - v| \leq k_1' DN^{-k} Lg^k N\}$ and we obtain

$$E(a(R_i^+) \operatorname{sgn} X_i | X_k = z) - E(a(R_i^+) \operatorname{sgn} X_i) \\ = N^{-1} [1 - \phi(q) - G_k(v)] (l_1(v) - m_1(v)) + o(N^{-1}) \quad (2.6)$$

uniformly in $-\infty < z < \infty$.

Thus, denoting

$$K_k = \operatorname{Cov}(E(a(R_i^+) \operatorname{sgn} X_i | X_k), E(a(R_j^+) \operatorname{sgn} X_j | X_k)) \quad (2.7)$$

we obtain

$$K_k = \int_{-\infty}^{\infty} [E(a(R_i^+) \operatorname{sgn} X_i | X_k = z) - E(a(R_i^+) \operatorname{sgn} X_i)] \\ \times [E(a(R_j^+) \operatorname{sgn} X_j | X_k = z) - E(a(R_j^+) \operatorname{sgn} X_j)] dF_k(z) \\ = N^{-2} (l_1(v) - m_1(v)) (l_j(v) - m_j(v)) \\ \times \int_{-v/DN^{-k}}^{(1-v)/DN^{-k}} [1 - \phi(q) - G_k(v)]^2 dG_k(v + DN^{-k}q) + o(N^{-2})$$

Note that

$$\int_{-v/DN^{-\frac{1}{2}}}^{(1-v)/DN^{-\frac{1}{2}}} (1 - \phi(q)) dG_k(v + DN^{-\frac{1}{2}}q) = G_k(v) + o(1)$$

and

$$\int_{-v/DN^{-\frac{1}{2}}}^{(1-v)/DN^{-\frac{1}{2}}} (1 - \phi(q))^2 dG_k(v + DN^{-\frac{1}{2}}q) = G_k(v) + o(1)$$

Hence:

$$K_k = N^{-2}(\ell_i(v) - \ell_j(v))(m_i(v) - m_j(v)) [G_k(v)(1 - G_k(v))] \\ + N^{-2}(\ell_i(v) - \ell_j(v))(m_i(v) - m_j(v))o(1) + o(N^{-2})$$

Summing over $1 \leq k \leq N$, $k \neq i, j$ ($i \neq j$), we finally obtain:

$$\sum_{k \neq i, j} K_k = N^{-1}(\ell_i(v) - \ell_j(v))(m_i(v) - m_j(v))D^2 + o(N^{-1})$$

uniformly in $1 \leq i \neq j \leq N$. The proof follows.

Using the Residual Variance Inequality and Lemmas 2.4 - 2.7, we obtain

Lemma 2.8. For $N \rightarrow \infty$, we have

$$E(S_N^+ - \hat{S}_N^+)^2 = o\left(\sum_{i=1}^N c_i^2\right). \quad (2.8)$$

Now from the definition of \hat{S}_N^+ , we have

$$\hat{S}_N^+ - E(\hat{S}_N^+) = \sum_{i=1}^N \sum_{j=1}^N c_j \{E(a(R_j^+) \operatorname{sgn} X_j | X_i) - E(a(R_j^+) \operatorname{sgn} X_j)\}$$

Set

$$Y_i = \sum_{j=1}^N c_j \{E(a(R_j^+) \operatorname{sgn} X_j | X_i) - E(a(R_j^+) \operatorname{sgn} X_j)\}, \quad 1 \leq i \leq N$$

and note that the Y_i , $1 \leq i \leq N$ are independent random variables with $E(Y_i) = 0$ and $\operatorname{Var}(\hat{S}_N^+) = \sum_{i=1}^N \operatorname{Var} Y_i$.

Define

$$Z_i = N^{-1} \left(\sum_{\substack{j=1 \\ j \neq i}}^N c_j (\ell_j(v) - m_j(v)) [u(v - H^*(|X_i|)) - G_i(v)] \right. \\ \left. + c_i [E(\operatorname{sgn} X_i a(R_i^+) | X_i) - E(\operatorname{sgn} X_i a(R_i^+))] \right), \quad 1 \leq i \leq N$$

and note that the Z_i , $1 \leq i \leq N$ are independent random variables with $E(Z_i) = 0$, $1 \leq i \leq N$.

Now for $j \neq i$, $1 \leq i, j \leq N$ we have

$$E(\operatorname{sgn} X_j a(R_j^+) | X_i) - E(\operatorname{sgn} X_j a(R_j^+)) \\ = \frac{1}{N} [\ell_j(v) - m_j(v)] \left[\Phi \left(\frac{v - H^*(|X_i|)}{DN^{-1/2}} \right) - G_i(v) \right] + \eta_j^i$$

where the η_j^i are random variables such that $|\eta_j^i| \leq \epsilon_N^i$ for some sequence of constants ϵ_N^i satisfying $N\epsilon_N^i \rightarrow 0$.

Now proceeding as in the derivation of (5.4) in Dupač and Hájek (1969), we obtain

$$E(Y_i - Z_i)^2 = o(N^{-1} \sum_{i=1}^N c_i^2). \quad (2.9)$$

Using

$$\operatorname{Var}(S_N^+ - \sum_{i=1}^N Z_i) \leq 2 \sum_{i=1}^N E(Y_i - Z_i)^2 + 2E(S_N^+ - \hat{S}_N^+)^2,$$

(2.8) and (2.9), we obtain

$$\text{Lemma 2.9. } \operatorname{Var}(S_N^+ - \sum_{i=1}^N Z_i) = o\left(\sum_{j=1}^N c_j^2\right). \quad (2.10)$$

Lemma 2.10. (1.18) holds if and only if

$$\liminf_{N \rightarrow \infty} \sigma_N^2 / \sum_{i=1}^N c_i^2 > 0 \quad (2.11)$$

where

$$\sigma_N^2 = \sum_{i=1}^N \text{Var}(Z_i)$$

In this case, $\lim_{N \rightarrow \infty} \text{Var}(S_N^+) / \sigma_N^2 = 1$.

Proof. It follows from the Minkowski inequality that

$$\left((\text{Var}(U_1) / \text{Var}(U_2))^{1/2} - 1 \right)^2 \leq \text{Var}(U_1 - U_2) / \text{Var}(U_2) \quad (2.12)$$

Let (1.18) be satisfied, then putting $U_1 = \sum_{i=1}^N Z_i$, $U_2 = S_N^+$ in (2.12) and using Lemma 2.9, we obtain (2.11). Let (2.11) be satisfied, then putting $U_1 = S_N^+$, $U_2 = \sum_{i=1}^N Z_i$, and using Lemma 2.9, we obtain (1.18).

Lemma 2.11. The random variables $\sum_{i=1}^N Z_i$ are asymptotically normal with parameters $(0, \sigma_N^2)$.

Proof. Since $\ell_i(v)$, $m_i(v)$ are bounded as functions of (i, N) , $1 \leq i \leq N$, it follows that

$$|Z_i| \leq C \max_{1 \leq j \leq N} |c_j| \quad \text{for some constant } C > 0. \quad (2.12')$$

Now (1.2) and (2.11) along with (2.12') imply

$$\max_{1 \leq i \leq N} |Z_i| / \sigma_N = o(1),$$

which by the Markov inequality, implies the Lindeberg condition for asymptotic normality.

Finally, since we have proved that

$$\sum_{i=1}^N z_i / \sigma_N \xrightarrow{D} N(0,1) , \quad (S_N^+ - E(S_N^+) - \sum_{i=1}^N z_i) / \sigma_N \xrightarrow{L^2} 0 , \quad (2.13)$$

and $\text{Var}(S_N^+) / \sigma_N^2 \rightarrow 1$,

we obtain

$$(S_N^+ - E(S_N^+)) / (\text{Var } S_N^+)^{1/2} \xrightarrow{D} N(0,1) \quad (2.14)$$

Remark 1. Suppose we want to relax the condition (1.3) to (1.4). Let us denote the statistic corresponding to (1.3) by S_N^+ and the statistic corresponding to (1.4) by S_N^{+*} . Then using (1.2) and (1.4), it follows that $\text{Var}(S_N^+ - S_N^{+*}) = o(\sum_{i=1}^N c_i^2)$. Consequently, the asymptotic normality of S_N^{+*} follows by using (2.11), (2.12), (2.13) and (2.14).

Remark 2. We have proved Theorem 2.1 under condition (C_1^+) . It remains to show that this set of conditions is implied by the conditions (C_2^+) and (C_3^+) . The proofs of these facts are similar to the implications $(C_3) \Rightarrow (C_1)$ and $(C_2) \Rightarrow (C_1)$ in Dupač and Hájek (1969, Section 5), and are therefore omitted.

The following theorem based on Theorem 2.1 and on Lemma 2 of Hušková (1970) combines unbounded c_{Ni} with a class of bounded score generating functions. The proof of this theorem is similar to that of Theorem 3 in Dupač and Hájek (1969) and is omitted.

Theorem 2.2. Let $S_N^+ = \sum_{i=1}^N c_{Ni} a_N(R_{Ni}^+) \text{sgn } X_{Ni}$ where

$a_N(i) = \psi(i/(N+1))$. Assume that $\psi = \psi_1 + \psi_2$, where ψ_1 is constant but for a finite number of jumps, and ψ_2 has a bounded second derivative. Assume that any one set of the conditions (C_1^+) , (C_2^+) , (C_3^+) holds along with (1.2). Then S_N^+ is asymptotically normal with natural parameters $(ES_N^+, \text{Var } S_N^+)$.

We now show that under slightly strengthened assumptions concerning the regression constants, S_N^+ is asymptotically normal with (simpler) parameters (μ_N^+, σ_N^2) where

$$\mu_N^+ = \sum_{i=1}^N c_i E[\text{sgn } X_i \psi(H^*(|X_i|))] \quad (2.15)$$

and

$$\sigma_N^2 = \sum_{i=1}^N \text{Var } Z_i \quad (2.16)$$

Theorem 2.3. Consider the statistic S_N^+ given by (1.1) with scores given by (1.4) where $\psi(t) = u(t-v)$. Assume that (C_1^+) or (C_2^+) or (C_3^+) holds. Then S_N^+ is asymptotically normal with parameters (μ_N^+, σ_N^2) defined in (2.15) and (2.16) if

$$\max_{1 \leq i \leq N} c_i^2 / \sum_{i=1}^N c_i^2 = O(N^{-\delta-1/2}) \text{ for some } \delta > 0. \quad (2.17)$$

Proof. Define

$$\begin{aligned} \Delta_i(X_i) = & \{E(\text{sgn } X_i a(R_i^+) | X_i) - E(\text{sgn } X_i a(R_i^+))\} \\ & - \{\text{sgn } X_i \psi(H^*(|X_i|)) - E(\text{sgn } X_i \psi(H^*(|X_i|)))\} \end{aligned}$$

Proceeding as in Dupač (1970), it can be shown (omitting the details of computation) that

$$E(\Delta_i^2) = O(N^{-1/2}), \text{ where } \Delta_i = \Delta_i(X_i).$$

This, together with (2.17) entails

$$c_i^2 E(\Delta_i^2) = o(N^{-1} \sum_{i=1}^N c_i^2)$$

Now, using the inequality,

$$(E(S_N^+) - \mu_N^+)^2 \leq \left(\sum_{i=1}^N c_i^2 \right) \left(\sum_{i=1}^N [E(\operatorname{sgn} X_i a(R_i^+)) - E(\operatorname{sgn} X_i \psi(H^*(|X_i|)))]^2 \right),$$

we obtain

$$(E(S_N^+) - \mu_N^+)^2 = o\left(\sum_{i=1}^N c_i^2 \right).$$

Now writing

$$E\left((S_N^+ - \mu_N^+ - \sum_{i=1}^N Z_i) / \sigma_N \right)^2 \leq 2E\left((S_N^+ - E(S_N^+) - \sum_{i=1}^N Z_i) / \sigma_N \right)^2 + 2(E(S_N^+) - \mu_N^+)^2 / \sigma_N^2$$

and proceeding as in Theorem 2.1, the proof follows.

Theorem 2.4. Consider the statistic S_N^+ given by (1.1) with the scores given by (1.3). Assume that $\psi = \psi_1 + \psi_2$ where

$\psi_1 = \sum_{j=1}^k \lambda_j \psi_{v_j}$ where $\psi_{v_j}(t) = u(t - v_j)$, $j = 1, \dots, k$, and

ψ_2 has a bounded second derivative. Let $Z_i = Z_i^{\psi_2} + \sum_{\ell=1}^k \lambda_\ell Z_i^{\psi_{v_\ell}}$,

where

$$Z_i^{\psi_{v_\ell}} = N^{-1} \left(\sum_{\substack{j=1 \\ j \neq i}}^N c_j(\ell_j(v_\ell)) - m_j(v_\ell) \right) (u(v_\ell - H^*(|X_i|)) - G_i(v_\ell)) + c_i [E(\operatorname{sgn} X_i a(R_i^+) | X_i) - E(\operatorname{sgn} X_i a(R_i^+))],$$

$$1 \leq i \leq N, \ell = 1, \dots, k$$

and

$$z_i^{\psi_2} = N^{-1} \sum_{j=1}^N c_j \int \operatorname{sgn} x [u(|x| - |x_i|) - F_i^*(|x|)] \psi_2'(H^*(|x|)) dF_j(x) \\ + c_i [\operatorname{sgn} x_i \psi_2(H^*(|x_i|)) - E(\operatorname{sgn} X_i \psi_2(H^*(|X_i|)))] ,$$

$$1 \leq i \leq N$$

(cf. Hušková (1970), p. 310)

Assume that (C_1^+) or (C_2^+) or (C_3^+) holds. Then the condition (2.17) implies the asymptotic normality of S_N^+ with parameters (μ_N^+, σ_N^2) where μ_N^+ and σ_N^2 are given by (2.15) and (2.16) respectively with Z_i given by (2.18).

The proof follows by combining Theorem 2.3, lemma 2 of Hušková (1970) and going through routine mathematical details.

REFERENCES

- Dupač, V. (1970). A contribution to the asymptotic normality of simple linear rank statistics. Nonparametric Techniques in Statistical Inference. 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.
- Hájek, J. (1968), Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist. 39, 325-346.
- Husková, M. (1970). Asymptotic distribution of simple linear rank statistics for testing symmetry. Z. Wahrscheinlichkeits theorie verw. Geb. 14, 308-322.