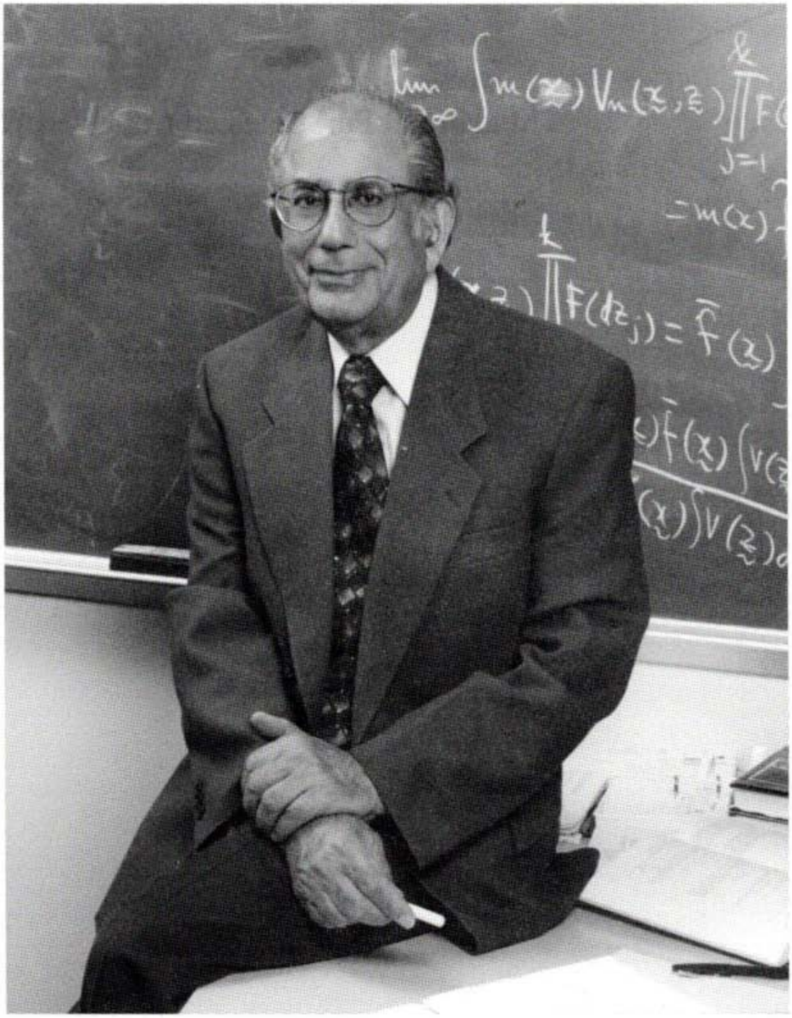
Madan Lal Puri Selected Collected Works
Volume 2



Madan Lal Puri

Dedication

These three volumes of "Selected Collected Works of Madan Lal Puri" will serve to preserve, in a unified and easily accessible form, the knowledge and wisdom conveyed in his many research papers, so as to aid its dissemination to future generations.

The Editors

MADAN LAL PURI
SELECTED COLLECTED WORKS

VOLUME 2

PROBABILITY THEORY AND EXTREME VALUE THEORY

Editors
Peter G. Hall, Marc Hallin and George G. Roussas

///VSP///
Utrecht • Boston
2003

VSP
P.O. Box 346
3700 AH Zeist
The Netherlands

Tel: +31 30 692 5790
Fax: +31 30 693 2081
vsppub@compuserve.com
www.vsppub.com

First published in 2003

ISBN 90-6764-385-8 (Volume 2)
ISBN 90-6764-387-4 (Set)

Printed in The Netherlands by Ridderprint bv, Ridderkerk

CONTENTS

Part I

Limit Theorems, Rates of Convergence, and Related Topics

(Independent Case)

ORDER OF NORMAL APPROXIMATION FOR RANK TEST STATISTICS DISTRIBUTION

JANA JUREČKOVÁ and MADAN L. PURI[1]

Charles University and Indiana University

0. Summary. Under suitable assumptions, it is established that the rate of convergence of the cdf (cumulative distribution function) of the simple linear rank statistics

$$S_N = \sum_{i=1}^N C_{Ni}\varphi\left(\frac{R_{Ni}}{N+1}\right)$$

to the normal one is $O(N^{-\frac{1}{2}+\delta})$ for any $\delta > 0$. Here $C_{N1}, \cdots, C_{NN}$ are known constants, $R_{N1}, \cdots, R_{NN}$ are the ranks of independent observations $X_{N1}, \cdots, X_{NN}$, and φ is a score generating function defined in Section 1.

1. Introduction. Let X_{Ni}, $i = 1, \cdots, N$ be independent rvs distributed according to the cdf $F_i(x) = F(x - \Delta d_{Ni})$, $i = 1, \cdots, N$. We assumed that $F(x)$ is absolutely continuous having the density function $f(x)$ whose derivative $f'(x)$ exists. Furthermore, $F(x)$ is assumed to have the finite Fisher information, that is,

$$I(f) = \int_{-\infty}^{\infty} [f'(x)/f(x)]^2 f(x)\,dx < \infty\,. \tag{1.1}$$

Δ is an unknown parameter, and d_{Ni}, $i = 1, \cdots, N$ are known constants. Let R_{Ni} be the rank of X_{Ni} among $X_{N1}, \cdots, X_{NN}$. Setting $u(x) = 1$ if $x \geqq 0$, and $u(x) = 0$ otherwise, we can write

$$R_{Ni} = \sum_{j=1}^N u(X_{Ni} - X_{Nj})\,, \qquad i = 1, \cdots, N\,. \tag{1.2}$$

Consider now the simple linear rank statistics

$$S_N = \sum_{i=1}^N C_{Ni}\, a_N(R_{Ni}) \tag{1.3}$$

where $C_{N1}, \cdots, C_{NN}$ are known constants, and $a_N(i)$, $i = 1, \cdots, N$ are "scores" generated by a function $\varphi(t)$ in the following manner:

$$a_N(i) = \varphi\left(\frac{i}{N+1}\right), \qquad 1 \leqq i \leqq N\,. \tag{1.4}$$

Statistics of the type (1.3) play an important role in the theory of nonparametric inference. For example, in the two sample problem where $F_1 = \cdots = F_m \equiv F$, and

$$F_{m+1} = \cdots = F_N \equiv G\,,$$

Received March 26, 1974; revised October 1, 1974.

[1] Work supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR 71-2009. Reproduction in whole or in part is permitted for any purpose of the U.S. Government. Part of this work was done while the author was the recipient of a Senior U.S. Scientist Award from the Alexander-von-Humboldt-Foundation.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 60F05, 60F99.

Key words and phrases. Linear rank statistics, score generating function, rate of convergence.

for testing the hypothesis H_0: $F \equiv G$, many rank tests are based on the statistic

$$S_N' = \sum_{i=1}^{m} a_N(R_{Ni})$$

which is a special case of (1.3) when $C_{N1} = \cdots = C_{Nm} = 1$ and $C_{Nm+1} = \cdots = C_{NN} = 0$. It is well known (see e.g., Capon (1961)) that the statistics of the form (1.3) for different score functions yield locally most powerful rank tests. Under suitable assumptions on the C's and the score generating function φ, Hájek (1962) [see also Hájek–Šidák (1967)] established the asymptotic normality of S_N. However, the problem of determining the rate of convergence of the cdf of S_N to the limiting normal distribution has remained open. This problem is investigated in this paper for the case $\Delta = 0$ as well as for $\Delta \neq 0$. In both cases, the rate of convergence is proved to be $O(N^{-\frac{1}{2}+\delta})$ for $\delta > 0$. For the case $\Delta = 0$, the result is valid for the φ functions having the bounded first derivative, and for the case $\Delta \neq 0$, it is necessary to assume the boundedness of the fourth derivative of φ.

Throughout the paper, we shall make the following assumptions on C's and d's.

(1.5) $\qquad \sum_{i=1}^{N} C_{Ni} = \sum_{i=1}^{N} d_{Ni} = 0\,, \qquad \sum_{i=1}^{N} C_{Ni}^2 = \sum_{i=1}^{N} d_{Ni}^2 = 1\,,$

(1.6) $\qquad \max_{1 \le i \le N} C_{Ni}^2 = O(N^{-1}\log N)\,, \qquad \max_{1\le i\le N} d_{Ni}^2 = O(N^{-1}\log N)\,.$

It may be noted that the assumption (1.5) can be made without any loss of generality. Furthermore, it may be noted [cf. Hájek–Šidák (1967)] that if φ is the difference of two non-decreasing, square integrable functions in (0, 1), then S_N has asymptotically $\eta(0, \sigma^2)$ distribution under $\Delta = 0$, and $\eta(ES_N, \sigma^2)$ or

$$\eta(\Delta \textstyle\sum_{i=1}^{N} C_{Ni} d_{Ni} \int_0^1 \varphi(t)\varphi(t, f)\, dt, \sigma^2)$$

distribution under $\Delta \neq 0$. Here

$$\sigma^2 = \int_0^1 (\varphi(t) - \bar\varphi)^2\, dt\,, \qquad \bar\varphi = \int_0^1 \varphi(t)\, dt\,, \qquad \varphi(t, f) = \frac{-f'(F^{-1}(t))}{f(F^{-1}(t))}$$

and $\eta(\xi, \sigma^2)$ stands for the normal distribution with mean ξ and variance σ^2.

2. Rate of convergence for $\Delta = 0$. The main result of this section is the following theorem.

THEOREM 2.1. *Let $\Delta = 0$ and the first derivative of $\varphi(t)$ exist and be bounded in* (0, 1). *Then, under the assumptions of Section* 1, *corresponding to any $\delta > 0$, there exists a constant $A(\delta) > 0$, and a positive integer N_δ such that for all $N > N_\delta$,*

(2.1) $\qquad \sup_{-\infty<x<\infty} |F_N(x) - \Phi(x)| \leqq A(\delta) N^{(-\frac{1}{2}+\delta)}$

where $F_N(x)$ is the cdf *of $\sigma^{-1}S_N$ and $\Phi(x)$ is the standard normal* cdf.

The proof of this theorem is based on the following two lemmas, the second of which is a consequence of Theorem 6, Chapter 5 of Petrov (1972).

LEMMA 2.1. *Under the assumptions of Theorem* 2.1, *corresponding to any positive integer k, where $2k + 1 < N$, there exists a constant $B(k) > 0$ and a positive integer N_k such that for all $N > N_k$,*

(2.2) $\qquad E(S_N - T_N)^{2k} \leqq B(k) N^{-k}$

where

$$T_N = \sum_{i=1}^N C_i \varphi(F(X_i)) . \tag{2.3}$$

LEMMA 2.2. *Under assumptions of Section 2 and Theorem 2.1, for any positive integer N,*

$$\sup_{-\infty<x<\infty} |F_N^*(x) - \Phi(x)| \leqq A \int_0^1 |\varphi(t) - \bar{\varphi}|^3 \, dt \cdot \sum_{i=1}^N |C_{Ni}|^3 \tag{2.4}$$

where $A > 0$ is a constant independent of N, and F_N^* is the cdf *of $\sigma^{-1}T_N$ under $\Delta = 0$.*

In what follows, we shall suppress the subscript N in C_{Ni}, d_{Ni}, R_{Ni}, etc. whenever there is no confusion.

PROOF OF LEMMA 2.1. Set $U_i = F(X_i)$, $i = 1, \cdots, N$. Denoting $Y_i = a_N(R_i) - \varphi(U_i)$, $i = 1, \cdots, N$, we get

$$\begin{aligned} E[(S_N - T_N)^{2k}] &= E\{(\textstyle\sum_{i=1}^N c_i Y_i)^{2k}\} \\ &= \sum \frac{(2k)!}{p_1! \cdots p_N!} c_1^{p_1} \cdots c_N^{p_N} E(\textstyle\prod_{i=1}^N Y_i^{p_i}) \end{aligned} \tag{2.5}$$

where the sum extends over the set A of vectors $(p_1, \cdots, p_N)$ of integers such that $0 \leqq p_i \leqq 2k$, $i = 1, \cdots, N$, $\sum_{i=1}^N p_i = 2k$.

Each point of A could have at most $2k$ positive components. Noting this fact, we may decompose A into $2k$ disjoint parts such that the jth part consists of those points which have just j positive components. Thus we may rewrite (2.5) as

$$\begin{aligned} E[(S_N - T_N)^{2k}] = \textstyle\sum_{i=1}^N c_i^{2k} E Y_i^{2k} + \cdots \\ + \textstyle\sum_{1 \leq p_1, \cdots, p_m < 2k, p_1 + \cdots + p_m = 2k} \dfrac{(2k)!}{p_1! \cdots p^m!} \\ \times \textstyle\sum_{i_1, \cdots, i_m = 1, \text{different}}^N c_{i_1}^{p_1} \cdots c_{i_m}^{p_m} E(Y_{i_1}^{p_1} \cdots Y_{i_m}^{p_m}) + \cdots \\ + \textstyle\sum_{i_1, \cdots, i_{2k} = 1, \text{different}}^N c_{i_1} \cdots c_{i_{2k}} E(Y_{i_1} \cdots Y_{i_{2k}}) . \end{aligned} \tag{2.6}$$

In view of (1.5) and (1.6), it follows that

$$|\textstyle\sum_{i_1, \cdots, i_m = 1, \text{different}}^N c_{i_1}^{p_1} \cdots c_{i_m}^{p_m}| \leqq K \qquad \text{for } N > N_k \tag{2.7}$$

for any $m = 1, \cdots, 2k$ and any p_i, $0 < p_i < 2k$, $i = 1, \cdots, m$, $\sum_{i=1}^m p_i = 2k$, $K > 0$ is a constant dependent only on k. Actually, if $p_i \geqq 2$ for $i = 1, \cdots, m$, then

$$|\textstyle\sum_{i_1, \cdots, i_m = 1, \text{different}}^N c_{i_1}^{p_1} \cdots c_{i_m}^{p_m} \leqq |\prod_{j=1}^m (\sum_{i=1}^N |c_i|^{p_j}) \leqq \max_{1 \leq i \leq N} |c_i|^{2(k-m)} .$$

On the other hand, suppose that some of p_i's are equal to one, say $p_m = 1$. Then in view of (1.5)

$$\begin{aligned} &\textstyle\sum_{i_1, \cdots, i_m = 1, \text{different}}^N c_{i_1}^{p_1} \cdots c_{i_1}^{p_m} \\ &\quad = \textstyle\sum_{i_1, \cdots, i_{m-1} = 1, \text{different}}^N c_{i_1}^{p_1} \cdots c_{i_{m-1}}^{p_{m-1}} (-c_{i_1} - \cdots - c_{i_{m-1}}) \end{aligned} \tag{2.8}$$

so that we get $m - 1$ sums of similar type; each of them sums the products of $(m - 1)$ factors. Considering any of these sums, we may have again two cases:

either all exponents are at least two, so that we are in the first case; or some of them equal one and we may write an equality analogous to (2.8). We continue in this way until after a finite number of steps (in which we decompose the original expression into at most $m!$ sums) we get only the sums with exponents greater than or equal to two. Actually, the extreme case is the sum of the type

$$\sum_{i_1,i_2=1,i_1\neq i_2}^{N} c_{i_1}^{2k-1} c_{i_2} = -\sum_{i_1=1}^{N} c_{i_1}^{2k},$$

so that (2.7) is proved.

Further, using the generalized Cauchy-Schwarz inequality

$$E|\prod_{i=1}^{n} Z_i| \leqq (\prod_{i=1}^{n} E|Z_i^n|)^{1/n}, \qquad n = 2, 3, \cdots \tag{2.9}$$

we see that

$$\begin{aligned} E|Y_{i_1}^{p_1} \cdots Y_{i_m}^{p_m}| &\leqq (\prod_{j=1}^{m} E|Y_{i_j}^{mp_j}|)^{1/m} < (\prod_{j=1}^{m} E|Y_{i_j}^{2kp_j}|)^{1/2k} \\ &= (\prod_{j=1}^{m} E|a_N(R_1) - \varphi(U_1)|^{2kp_j})^{1/2k} \end{aligned} \tag{2.10}$$

holds for any $m = 1, \cdots, 2k$ and any p_i, $0 < p_i \leqq 2k$, $\sum_{i=1}^{m} p_i = 2k$. Finally, the expression

$$\sum_{m=1}^{2k} \sum_{1 \leq p_1,\cdots,p_m \leq 2k, p_1+\cdots+p_m=2k} \frac{(2k)!}{p_1! \cdots p_m!} \tag{2.11}$$

depends only on k.

Now, if $a_N(i) = \varphi(i/(N+1))$, $i = 1, \cdots, N$, where φ has a bounded derivative we get the inequality

$$E[a_N(R_1) - \varphi(U_1)]^{2kp_j} \leqq B_2(k) E\left[\frac{R_1}{N+1} - U_1\right]^{2kp_j} \tag{2.12}$$

which is varied for $j = 1, \cdots, m$; $m = 1, \cdots, 2k$.

U_1 being fixed, R_1 is the sum of independent zero-one random variables (see (1.2)) so that

$$E\left(\frac{R_{N1}}{N+1} - U_{N1}\right)^{2kp_j} \leqq B_3(k) N^{-kp_j}. \tag{2.13}$$

(2.6), (2.7), (2.10), (2.11), (2.12) and (2.13) then prove the lemma.

Proof of Theorem 2.1. Since for any $\varepsilon > 0$ and any N, we have

$$P\{\sigma^{-1}S_N \leqq x\} \leqq P\{\sigma^{-1}T_N \leqq x + \varepsilon\} + P\{\sigma^{-1}|S_N - T_N| \geqq \varepsilon\} \tag{2.14}$$

and analogously

$$P\{\sigma^{-1}S_N \leqq x\} \geqq P\{\sigma^{-1}T_N \leqq x - \varepsilon\} - P\{\sigma^{-1}|S_N - T_N| \geqq \varepsilon\}, \tag{2.15}$$

it follows using Lemmas 2.1 and 2.2, that

$$\sup_{-\infty<x<\infty} |F_N(x) - \Phi(x)| \leqq (\varepsilon\sigma)^{-2k} B(k) N^{-k} + c_2 \sum_{i=1}^{N} |c_{Ni}|^3 + O(\varepsilon) \tag{2.16}$$

holds for any $\varepsilon > 0$, any k and for $N > N_k$.

For $\delta > 0$ being fixed, take k such that $2k + 1 > 1/2\delta \geqq 2k$ and put $\varepsilon = N^{-\frac{1}{2}(1-1/(2k+1))}$. The theorem then follows from (2.13) and from the assumption (1.6).

3. Rate of convergence for $\Delta \neq 0$. Without loss of generality, we assume that $\Delta > 0$. For convenience we shall use the following representation in this section. Let X_{Ni}, $i = 1, \cdots, N$ be independent and identically distributed rvs each having the cdf $F(x)$ such that $I(f) < \infty$. Let R_{Ni}^{Δ} be the rank of $X_{Ni} + \Delta d_{Ni}$, that is

$$R_{Ni}^{\Delta} = \sum_{j=1}^{N} u(X_{Ni} - X_{Nj} + \Delta(d_{Ni} - d_{Nj}))\,.$$

Consider now the statistics

$$S_{\Delta N} = \sum_{i=1}^{N} c_{Ni}\varphi\left(\frac{R_{Ni}^{\Delta}}{N+1}\right).$$

The asymptotic distribution of $S_{\Delta N} - S_{0N}$ was investigated by Jurečková for Wilcoxon scores in (1973a) and for general score function φ in (1973b). In the case of general scores function φ, it was assumed that the φ function has the four bounded derivatives in (0, 1).

Suppose now that the vectors $(c_{N1}, \cdots, c_{NN})$ and $(d_{N1}, \cdots, d_{NN})$ satisfy (1.5), (1.6) and the following:

$$\lim_{N\to\infty} \sum_{i=1}^{N} c_{Ni} d_{Ni} = a^2\,, \qquad 0 < a^2 < \infty\,, \tag{3.1}$$

$$\lim_{N\to\infty} [\max_{1\le i\le N} (c_{Ni}^2 d_{Ni}^2)(\sum_{i=1}^{N} c_{Ni}^2 d_{Ni}^2)^{-1}] = 0\,, \tag{3.2}$$

and

$$\lim_{N\to\infty} [N^{-1}(\sum_{i=1}^{N} c_{Ni} d_{Ni})^2(\sum_{i=1}^{N} c_{Ni}^2 d_{Ni}^2)^{-1}] = \gamma \geqq 0\,. \tag{3.3}$$

Then, [cf. Jurečková (1973b)] for φ having four bounded derivatives in (0, 1), the asymptotic distribution of

$$A_N^{-1}(S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N) \tag{3.4}$$

is $\eta(0, \Delta^2\rho^2)$ where

$$A_N^2 = \sum_{i=1}^{N} c_{Ni}^2 d_{Ni}^2 + 3N^{-1}(\sum_{i=1}^{N} c_{Ni} d_{Ni})^2 \tag{3.5}$$

$$a_N = \sum_{i=1}^{N} c_{Ni} d_{Ni} \int \varphi'(F(x)) f^2(x)\, dx = \sum_{i=1}^{N} c_{Ni} d_{Ni} \int_0^1 \varphi(t)\varphi(t, f)\, dt \tag{3.6}$$

$$b_N = \tfrac{1}{2} \sum_{i=1}^{N} c_{Ni} d_{Ni}^2 \int \varphi''(F(x)) f^3(x)\, dx \tag{3.7}$$

and

$$\begin{aligned} \rho^2 = {} & \int [\varphi'(F(x))]^2 f^3(x)\, dx - (\int [\varphi'(F(x))]^2 f^2(x)\, dx)^2 + 2\gamma(1+3\gamma)^{-1} \\ & \times [\iint_{x<y} F(x)(1-F(y))\varphi''(F(x))\varphi''(F(y)) f^2(x) f^2(y)\, dx\, dy \\ & + \iint_{x<y} \varphi'(F(x))\varphi''(F(y)) f^2(x) f^2(y)\, dx\, dy \\ & - \int \varphi'(F(x)) f(x)\, dx \cdot \int \varphi''(F(x)) F(x) f^2(x)\, dx\,. \end{aligned} \tag{3.8}$$

Let $F_{N\Delta}$ denote the cdf of $\sigma^{-1}(S_{\Delta N} - \Delta a_N)$. Then we have the following theorem.

THEOREM 3.1. *Suppose that c_{Ni}, d_{Ni}, $i = 1, \cdots, N$ satisfy* (1.5), (1.6), (3.1)—(3.3) *and that the score-generating function has four bounded derivatives on* (0, 1).

Then

$$\sup_x |F_{N\Delta}(x) - \Phi(x)| = O(N^{-\frac{1}{2}+\delta}) \tag{3.9}$$

holds for any $\delta > 0$ and any fixed Δ.

PROOF. We may write for any $\varepsilon > 0$ and for any x

$$\begin{aligned} P\{\sigma^{-1}(S_{\Delta N} - \Delta a_N - \Delta^2 b_N) \leqq x\} \\ &\leqq P\{\sigma^{-1}S_{0N} \leqq x + \varepsilon\} \\ &\quad + P\{\sigma^{-1}|S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N| \geqq \varepsilon\} \end{aligned} \tag{3.10}$$

and analogously

$$\begin{aligned} P\{\sigma^{-1}(S_{\Delta N} - \Delta a_N - \Delta^2 b_N) \leqq x\} \\ \geqq P\{\sigma^{-1}S_{0N} \leqq x - \varepsilon\} - P\{\sigma^{-1}|S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N| \geqq \varepsilon\}\,. \end{aligned}$$

Then by Theorem 2.1,

$$\begin{aligned} \sup_x |F_{N\Delta}(x + \sigma^{-1}\Delta^2 b_N) - \Phi(x)| \\ &\leqq C \cdot \varepsilon + P\{\sigma^{-1}|S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N| \geqq \varepsilon\} \\ &\quad + A(\delta)N^{-\frac{1}{2}+\delta} \end{aligned} \tag{3.11}$$

holds for any $\delta > 0$ and $N > N_\delta$.

Let us consider the third member of the right-hand side of (3.11). We shall use the following theorem:

THEOREM 3.2 (Petrov). *Let $H(x)$ be any* cdf *and $\Phi(x)$* cdf *of the normal* (0, 1) *distribution.*

Let

$$\nu = \sup_{-\infty<x<\infty} |H(x) - \Phi(x)|$$

and let M_p denote the set of distribution functions possessing the finite absolute moment of order $p > 0$. Then, if $0 < \nu \leqq e^{-\frac{1}{2}}$ and $H(x) \in M_p$, there exists a constant C_p depending on p only such that

$$|H(x) - \Phi(x)| \leqq \frac{C_p \nu \left(\log \dfrac{1}{\nu}\right)^{p/2} + \lambda_p}{1 + |x|^p} \tag{3.12}$$

holds for all real x; here

$$\lambda_p = \left|\int |x|^p \, dH(x) - \int |x|^p \, d\Phi(x)\right|\,.$$

For the proof, see Petrov (1972).

Let us denote by $G_{N\Delta}$ the cdf of $\Delta^{-1}A_N{}^{-1}\rho^{-1}(S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N)$. On account of the boundedness of φ, $G_{N\Delta}$ has finite absolute moments of any order for any fixed N and any fixed Δ. On the other hand, it follows from Theorem 2.1 of [6] (see (3.1)—(3.8) of the present paper) that $\lim_{N\to\infty} \sup_x |G_{N\Delta}(x) - \Phi(x)| = 0$ for any fixed Δ and that for $N > N_\Delta$

$$\sup_x |G_{N\Delta}(x) - \Phi(x)| < e^{-\frac{1}{2}}\,.$$

The assumptions of Theorem 3.2 are satisfied for any $p = k = 1, 2, \cdots$, so that there exists a constant C_k^* to any k such that

$$|G_{N\Delta}(x) - \Phi(x)| \leqq C_k^*(1 + |x|^k)^{-1} \tag{3.13}$$

holds for all $x \in (-\infty, \infty)$.

We have

$$P\{\sigma^{-1}|S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N| \geqq \varepsilon\} = 2[1 - G_{N\Delta}(\Delta^{-1}\rho^{-1}\sigma A_N{}^{-1}\varepsilon)] \tag{3.14}$$

so that (3.13) implies that

$$\begin{aligned} P\{\sigma^{-1}|S_{\Delta N} - S_{0N} - \Delta a_N - \Delta^2 b_N| \geqq \varepsilon\} \\ \leqq 2[1 - \Phi(\Delta^{-1}\rho^{-1}\sigma A_N{}^{-1}\varepsilon)] + 2C_k^*[1 + (\Delta^{-1}\rho^{-1}\sigma)^k A_N{}^{-k}\varepsilon^k]^{-1} \end{aligned} \tag{3.15}$$

holds for any $\varepsilon > 0$, any $k = 1, 2, \cdots$ and for $N > N_\Delta$.

Let us fix δ, $\delta > 0$ and put $\varepsilon = A_N \cdot N^{\delta/2}$. Then in view of (3.15) and Lemma 2, Chapter VII of Feller (1957) we have that for any $N > N_\Delta$ and sufficiently large k

$$\sup |F_{N\Delta}(x + \sigma^{-1}\Delta^2 b_N) - \Phi(x)| \leqq C_\delta'' N^{-\frac{1}{2}+\delta} + O(N^{-1+2\delta})\,. \tag{3.16}$$

Thus

$$\begin{aligned} \sup_{-\infty<x<\infty} &|F_{N\Delta}(x) - \Phi(x)| \\ &\leqq \sup_x |F_{N\Delta}(x) - \Phi(x + \sigma^{-1}\Delta^2 b_N)| \\ &\qquad + \sup_x |\Phi(x + \sigma^{-1}\Delta^2 b_N) - \Phi(x)| \\ &\leqq \sup_x |F_{N\Delta}(x - \sigma^{-1}\Delta^2 b_N) - \Phi(x)| + K \cdot \sigma^{-1}\Delta^2 b_N\,. \end{aligned} \tag{3.17}$$

(3.16) and (3.17) together with assumption (1.5) complete the proof of the Theorem.

Acknowledgment. It is a pleasure to express our appreciation to Dr. Søren Johansen and Martin Jacobsen for some helpful discussions, and to the referee for some very valuable comments. We would also like to thank the referee for bringing to our attention a paper of Bickel (1974) where he established that a Berry–Esseén bound of order $N^{-\frac{1}{2}}$ was valid for the Wilcoxon statistic. This suggests that the bound in the present paper may be improved upon. The reader is also referred to Vizková, (1974) for a related problem.

REFERENCES

[1] Bickel, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

[2] Capon, J. (1961). Asymptotic efficiency of certain locally most powerful rank tests. *Ann. Math. Statist.* **32** 88–100.

[3] Feller, W. (1957). *An Introduction to Probability Theory and its Applications.* Wiley, New York.

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

[5] Hájek, J. and Sidác, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[6] Jurečková, J. (1973 a). Central limit theorem for Wilcoxon rank statistics process. *Ann. Statist.* **1** 1046–1060.

[7] JUREČKOVÁ, J. (1973 b). Central limit theorem for rank statistics process. (Submitted).

[8] PETROV, V. V. (1972). ***The Sums of Independent Random Variables*** (in Russian). Nauka, Moscow.

[9] VÍZKOVÁ, Z. (1974). Local limit theorems for rank statistics. Ph. D. dissertation, Charles Univ., Prague.

JANA JUREČKOVÁ
CHARLES UNIVERSITY
DEPARTMENT OF STATISTICS
SOKOLOVSKA UL. 83
PRAGUE 8, CZECHOSLOVAKIA

MADAN L. PURI
INSTITUT FUR MATHEMATISCHE STATISTICK
UND WIRTSCHAFTSMATIK
UNIVERSITAT GOTTINGEN
34 GOTTINGEN
LOTZESTRABE 13, WEST GERMANY

CONVERGENCE AND REMAINDER TERMS IN LINEAR RANK STATISTICS[1]

HARALD BERGSTRÖM and MADAN L. PURI

University og Göteborg, Chalmers University of Technology, and Indiana University

A new approach to the asymptotic normality of simple linear rank statistics for the regression case studied earlier by Hájek (1968) is provided along with the estimation of the remainder term in the approximation to normality.

1. Introduction and summary. Let $X_1, \cdots, X_n$ be independent random variables having continuous cdf's (cumulative distribution functions) $F_1(x), \cdots, F_n(x)$ respectively. Consider a statistic $S_n = s(X_1, \cdots, X_n)$ with $ES_n = 0$ and $ES_n^2 < \infty$. Then, to prove the asymptotic normality of S_n (as $n \to \infty$), Hájek (1968) uses the method of projection which gives to the statistic S_n, the approximation of the form

$$\hat{S}_n = \sum_{j=1}^{n} E[S_n \mid X_j] \,. \tag{1.1}$$

Consider now the simple linear rank statistic S_n introduced by Hájek (1962, 1968)

$$S_n = \sum_{j=1}^{n} c_j\{\psi(R_j/n) - E[\psi(R_j/n)]\} \tag{1.2}$$

where the c's are known constants, R_j is the rank of X_j among $(X_1, \cdots, X_n)$ and $\psi(\cdot)$ is a score generating function defined on (0, 1). Hájek (1962) [see also Hájek–Šidák (1967)] established the asymptotic normality of S_n in (1.2) under the assumption that the F_i are contiguous, e.g., when $F_i(x) = F(x - \Delta d_{ni})$ where Δ is the unknown parameter and the d's are the known constants. Later on Hájek (1968) studied the asymptotic normality of S_n for the general $F_i(x)$ (the noncontiguous case). Under the setup of Hájek (1962), Jurečková and Puri (1975), referred to hereafter as JP, studied the problem of determining the rate of convergence of the cdf of S_n to the limiting normal cdf and established it of order $O(N^{-\frac{1}{2}+\delta})$ for $\delta > 0$. In this paper we not only give a new approach to the asymptotic normality of S_n for the general F_i (i.e., not necessarily contiguous) but improve the results of JP in providing a sharper bound (for the general F_i's). In the passing, we may also mention that whereas JP requires ψ to have a bounded fourth derivative, here we only require the boundedness of the second

---

Received May 1975; revised January 1977.

[1] Work supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFOSR 76-2927. Reproduction in whole or in part is permitted for any purpose of the U.S. Government. Part of this work was done while the author was the recipient of a Senior U.S. Scientist Award from the Alexander-von-Humboldt-Foundation.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 60F05, 60F99.

Key words and phrases. Linear rank statistics, score generating functions, rate of convergence.

derivative. Furthermore whereas this paper gives more explicit error bounds than the JP paper, the latter gives more information on the limiting behavior of ES_n and Var S_n.

We now introduce some notations. We define $\psi(\cdot) = 0$ outside $(0, 1)$. Then, we can use the supremum norm

$$(1.3) \qquad ||\psi|| = \sup_{t\in(-\infty,\infty)} |\psi(t)| \,.$$

Set

$$(1.4) \qquad \rho_i = R_i/n\,, \qquad \rho_{ii} = E[\rho_i \,|\, X_i]\,, \qquad u(x) = 1 \quad \text{if} \quad x \geqq 0 \quad \text{and} \quad u(x) = 0 \quad \text{otherwise.}$$

Then

$$(1.5) \qquad R_i = \textstyle\sum_{j=1}^n u(X_i - X_j)\,.$$

In this paper, we shall deal with the following approximation of S_n.

$$(1.6) \qquad T_n = \textstyle\sum_{i=1}^n c_i\{\psi(\rho_{ii}) - E[\psi(\rho_{ii})] + (\rho_i - \rho_{ii})\psi'(\rho_{ii})\}\,,$$

assuming that ψ' exists on $(0, 1)$ and

$$(1.7) \qquad \hat{T}_n = \textstyle\sum_{j=1}^n E[T_n \,|\, X_j]\,.$$

Since $E[(\rho_i - \rho_{ii})\psi'(\rho_{ii})] = 0$, it follows that

$$(1.8) \qquad \hat{T}_n = \textstyle\sum_{i=1}^n c_i\{\psi(\rho_{ii}) - E[\psi(\rho_{ii})] + \sum_{j\neq i}^n E[(\rho_i - \rho_{ii})\psi'(\rho_{ii}) \,|\, X_j]\}\,.$$

Let H_n, G_n and $\hat{G}_n$ be the cdf's of S_n, T_n and $\hat{T}_n$ respectively, and put

$$(1.9) \qquad \sigma_n^2 = E[S_n^2]\,, \qquad \hat{\delta}_n^2 = E[\hat{T}_n^2]\,, \qquad \Gamma_{nr}^{2r} = \frac{1}{n}\textstyle\sum_{i=1}^n c_i^{2r}\,, \qquad \Gamma_{nr} > 0\,.$$

Then our theorems are the following:

THEOREM 1.1. *If ψ has a derivative on* $(0, 1)$ *then*

$$(1.10) \qquad ||\hat{G}_n(\hat{\delta}_n\cdot) - \Phi(\cdot)|| \leqq 4C[2||\psi||^3 + ||\psi'||^3]\textstyle\sum_{i=1}^n |c_i|^3\hat{\delta}_n^{-3}\,;$$
$$\Phi(x) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^x e^{-t^2/2}\,dt$$

where C is the constant in Berry–Esseen's inequality (Zolotarev (1967) gives the approximation 0.9051). *Further,*

$$(1.11) \qquad |\hat{\delta}_n - \sigma_n| \leqq C_1(||\psi'|| + ||\psi''||)\Gamma_{n,1}$$

with an absolute constant C_1, provided ψ'' exists on $(0, 1)$.

THEOREM 1.2. *If ψ has a second order derivative on* $(0, 1)$, *then for any positive integers n and r such that $n^{-1}r^3 \leqq \frac{3}{8}$,*

$$(1.12) \qquad ||H_n(\hat{\delta}_n\cdot) - \Phi(\cdot)|| \leqq 4C(2||\psi||^3 + ||\psi'||^3)\textstyle\sum_{i=1}^n |c_i|^3\hat{\delta}_n^{-3} + C_2[\hat{\delta}_n^{-1}(||\psi'|| + ||\psi''||)r\Gamma_{nr}]^{2r/(2r+1)}\,,$$

where C_2 is an absolute constant.

REMARK. If the c_i are chosen such that $|c_i| \leqq a/n^{\frac{1}{2}}$ with constant a for all

i and n, then

$$\Gamma_{nr} \leqq a/n^{\frac{1}{2}} ,$$

and for $r = [\log n]$, $[r\Gamma_{nr}]^{2r/(2r+1)} \leqq a\sqrt[4]{e}\,(\log n)\, n^{-\frac{1}{2}}(1 + O(1/\log n))$.

Note that $\hat{\delta}_n^{-1}c_i$ is invariant and thus also $\hat{\delta}_n^{-1}\Gamma_{nr}$ is invariant under the transformation $c_i \to \gamma c_i$, $i = 1, 2, \cdots$.

2. Some lemmas.

LEMMA 2.1. *For any positive integers r and n, $2r \leqq n$, we have*

$$E[(\rho_i - \rho_{ii})^{2r}] \leqq b(r)n^{-r} \tag{2.1}$$

with

$$b(r) \leqq n^{-r} \sum_{t=1}^{r} \binom{n-1}{t} \frac{(2r)!}{(2r-2t)!}\, t^{2r-2t} \cdot 2^{-3t} \tag{2.2}$$

and for $n^{-1}r^3 \leqq \frac{3}{4}$

$$b(r) \leqq 2^{-3r} \frac{(2r)!}{r!} [1 + 8n^{-1}r^3] . \tag{2.3}$$

PROOF. By (1.4) we obtain

$$\rho_i - \rho_{ii} = \frac{1}{n} \sum_{j\neq i}^{n} [u(X_i - X_j)F_j(X_i)] .$$

By the polynomial theorem we then get

$$E[(\rho_i - \rho_{ii})^{2r}] = n^{-2r} \sum \frac{(2r)!}{s_1! \cdots s_n!} E \prod_{j\neq 1}^{n} [u(X_i - X_j) - F_j(X_i)]^{s_j} , \quad s_1 + \cdots + s_n = 2r . \tag{2.4}$$

We claim that any term in this sum is equal to zero if $s_{j_0} = 1$ for some j_0. Indeed we find that the conditional expection of the product with respect to all X_j, $j \neq j_0$ is equal to 0 if $s_{j_0} = 1$. Hence we have only to regard terms with $s_j = 0$ or $\geqq 2$ for any j, and there can be at most $t \leqq r$ exponents s_j different from 0. If $s_j \geqq 2$, $j = 1, 2, \cdots, t$, $s_j = 0$ for $j > t$, $i > t$ we obtain, observing that

$$|u(X_i - X_j) - F_j(X_i)| \leqq 1$$

$$\begin{aligned} E[\textstyle\prod_{j=1}^{t} [u(X_i - X_j) - F_j(X_i)]^{s_j} &\leqq E \textstyle\prod_{j=1}^{t} [u(X_i - X_j) - F_j(X_i)]^2 \\ &= E[\textstyle\prod_{j=1}^{t} [F_j(X_i) - F_j^2(X_i)]] \leqq 4^{-t} . \end{aligned} \tag{2.5}$$

This inequality remains true for all permutations of the indices $1, \cdots, n$. Put

$$\gamma(t) = \sum_{s_1+\cdots+s_t=2r;\, s_j\geqq 2,\, j=1,\cdots,t} \frac{(2r)!}{s_1! \cdots s_t!} . \tag{2.6}$$

Since t indices out of $n-1$ indices can be chosen in $\binom{n-1}{t}$ different ways we obtain from (2.4) through (2.6),

$$E[(\rho_i - \rho_{ii})^{2r}] \leqq n^{-2r} \sum_{t=1}^{r} \binom{n-1}{t} \gamma(t) 4^{-t} . \tag{2.7}$$

We claim that

$$\gamma(t) \leqq \frac{(2r)!}{(2r-2t)!} 2^{-t} t^{2r-2t} . \tag{2.8}$$

Indeed, differentiating the identity

$$(\sum_{j=1}^{t} y_j)^{2r} = \sum_{s_1+\cdots+s_t=2r} \frac{(2r)!}{s_1! \cdots s_t!} \prod_{j=1}^{t} y_j^{s_j}$$

twice with respect to all y_j and then putting all y_j equal to 1, we obtain

$$\frac{(2r)!}{(2r-2t)!} t^{(2r-2t)} = \sum_{s_1+\cdots+s_t=2r; s_j \geq 2, j=1\cdots t} \prod_{j=1}^{t} s_j(s_j-1) \frac{(2r)!}{s_1! \cdots s_t!} .$$

Now using (2.7) and (2.8), we get (2.1) and (2.2). We now estimate $b(r)$ further, mainly for use when n and r are large. Put $r - t = u$. Then we can write

$$b(r) \leqq 2^{-3r} \sum_{u=0}^{r-1} k(u) \tag{2.9}$$

with

$$k(u) = \frac{n^{-u}(2r)!\,(r-u)^{2u}2^{3u}}{(r-u)!\,(2u)!} .$$

Particularly

$$k(0) = \frac{(2r)!}{r!} , \qquad k(1) < 4n^{-1} r^3 \cdot \frac{(2r)!}{r!}$$

and for $u \geqq 1$

$$\frac{k(u+1)}{k(u)} = n^{-1}\left(1 - \frac{1}{r-u}\right)^{2u} \cdot 2^3 \cdot (r-u) \frac{(r-u-1)^2}{(2u+1)(2u+2)}$$
$$< \tfrac{2}{3} n^{-1} r^3 \leqq \tfrac{1}{2} \qquad \text{for} \quad n^{-1}r^3 \leqq \tfrac{3}{4} .$$

Hence

$$b(r) \leqq 2^{-3r} \cdot \frac{(2r)!}{r!} [1 + 8n^{-1}r^3]$$

for $n^{-1}r^3 \leqq \frac{3}{4}$.

LEMMA 2.2. *For any positive integers r and n, $2r \leqq n$, we have*

$$E(T_n - \hat{T}_n)^{2r} \leqq c(r) \|\psi\|^{2r} \Gamma_{n,r}^{2r} \tag{2.10}$$

if ψ' exists on $(0, 1)$, *and if ψ'' exists on* $(0, 1)$

$$E[(S_n - T_n)^{2r}] \leqq b(2r) \|\psi''\|^{2r} \Gamma_{n,r}^{2r} , \tag{2.11}$$

$$E[(S_n - \hat{T}_n)^{2r}] \leqq d(r, \psi) \Gamma_{n,r}^{2r} \tag{2.12}$$

with

$$b(2r) \leqq n^{-2r} \sum_{t=1}^{2r} \binom{n-1}{t} \frac{(4r)!}{(4r-2t)!} t^{4r-2t} \cdot 2^{-3t}$$

$$c(r) \leqq 2^{2r} n^{-2r} \sum_{t=1}^{2r} \binom{n}{t} \frac{(4r)!}{(4r-2t)!} t^{4r-2t} \cdot 2^{-t}$$

$$d(r, \psi) \leqq [[b(2r)]^{1/2r} \|\psi''\| + [c(r)]^{1/2r} \|\psi'\|]^{2r} .$$

Further we have the estimates

$$b(2r) \leqq 2^{-6r}\frac{(4r)!}{(2r)!}[1 + 2^6 n^{-1} r^3] \tag{2.13}$$

for $2^3 n^{-1} r^3 \lesseqqgtr \frac{3}{4}$,

$$c(r) \leqq \frac{(4r)!}{(2r)!}[1 + 2^3 n^{-1} r^3] \qquad \textit{for} \quad n^{-1} r^3 \leqq \tfrac{3}{8}\,. \tag{2.14}$$

REMARK. By Stirling's approximation of the Γ-function we have

$$\frac{(4r)!}{(2r)!} \leqq 2^{6r+\frac{1}{2}} r^{2r} (\exp -2r) \exp \frac{1}{48r}\,.$$

PROOF. By (1.6) and (1.8) we get

$$T_n - \hat{T}_n = \textstyle\sum_{i=1}^n c_i \{(\rho_i - \rho_{ii})\psi'(\rho_{ii}) - \sum_{j=1; j\neq i}^n E[(\rho_i - \rho_{ii})\psi'(\rho_{ii}) \,|\, X_j]\} \tag{2.15}$$

and for $j \neq i$

$$\begin{aligned} E[(\rho_i - \rho_{ii})\psi'(\rho_{ii}) \,|\, X_j] &= \frac{1}{n} \textstyle\sum_{k\neq i}^n E\{[u(X_i - X_k) - F_k(X_i)]\psi'(\rho_{ii}) \,|\, X_j\} \\ &= \frac{1}{n} E[u(X_i - X_j) - F_j(X_i)]\psi'(\rho_{ii}) \,|\, X_j]\,, \end{aligned} \tag{2.16}$$

since the conditional expectations in the sum are zero for $j \neq k, i$. Now using the relation

$$(\rho_i - \rho_{ii})\psi'(\rho_{ii}) = \frac{1}{n} \textstyle\sum_{j\neq i}^n [u(X_i - X_j) - F_j(X_i)]\psi'(\rho_{ii})\,,$$

and noting that

$$E[(\rho_i - \rho_{ii})\psi'(\rho_{ii}) \,|\, X_i] = 0$$

we obtain from (2.15)

$$T_n - \hat{T}_n = \frac{1}{n} \textstyle\sum_{i=1}^n \sum_{j\neq i}^n c_i V_{ij} \tag{2.17}$$

with

$$\begin{aligned} V_{ij} &= [u(X_i - X_j) - F_j(X_i)]\psi'(\rho_{ii}) \\ &\quad - E\{[u(X_i - X_j) - F_j(X_i)]\psi'(\rho_{ii}) \,|\, X_j\}\,. \end{aligned} \tag{2.18}$$

Clearly

$$E[V_{ij} \,|\, X_j] = 0\,, \qquad E[V_{ij} \,|\, X_i] = 0\,. \tag{2.19}$$

By the polynomial theorem we get

$$\begin{aligned} E[(T_n - \hat{T}_n)^{2r}] &= n^{-2r} E[\textstyle\sum_{i=1}^n \sum_{j\neq i}^n c_i V_{ij}]^{2r} \\ &= n^{-2r} \sum \frac{(2r)!}{\prod_{i=1}^n \prod_{j\neq i}^n (s_{ij}!)} E\{\textstyle\prod_{i=1}^n \prod_{j\neq i}^n (c_i V_{ij})^{s_{ij}}\} \end{aligned} \tag{2.20}$$

where the sum should be taken over terms corresponding to different vector solutions $\{s_{ij}\}$, $i, j = 1, \cdots n$, $j \neq i$ of the equation

$$\textstyle\sum_{i=1}^n \sum_{j\neq i}^n s_{ij} = 2r\,. \tag{2.21}$$

$$(2.22) \qquad E[\prod_{i=1}^{n} \prod_{j\neq i}^{n} V_{ij}^{s_{ij}}]$$

is equal to 0 for some vector solutions of (2.21) since (2.19) holds, and we have only to regard those solutions for which the expectation (2.22) is not equal to 0.

We say that s_{ij} gives the contribution $\frac{1}{2}s_{ij}$ to the sum (2.21) from each of the indices i and j. Hence according to this notation an index k gives the contribution

$$(2.23) \qquad g(k) = \tfrac{1}{2} \sum_{j\neq k}^{n} s_{kj} + \tfrac{1}{2} \sum_{j\neq k}^{n} s_{jk}$$

to the sum (2.21). By conditioning with respect to all $X_j, j \neq k$ we easily find that the expectation (2.22) is equal to 0 if k gives the contribution $\frac{1}{2}$ to the sum (2.21), i.e., if $s_{kj} = 1$ for exactly one index $j \neq k$, and $s_{jk} = 0$ for $j \neq k$ or if $s_{jk} = 1$ for exactly one j and $s_{kj} = 0$ for $j \neq k$.

The sum $\sum$ on the right-hand side of (2.20) can be divided into partial sums as follows. Let C be a collection of different positive integers belonging to the set $1, \cdots, 2r$, say $C = (1, 2, \cdots, t)$. Let $\sum_C$ consist of all terms in (2.20) corresponding to the vector solutions of (2.21) such that

(a) $s_{ij} = 0$ if not both i and j belong to C;

(b) for any $k \in C$ the contribution to the sum (2.21) is larger than $\frac{1}{2}$. Note that C can contain at most 2r different integers since every $k \in C$ gives at least the contribution 1 to the sum (2.21). Clearly partial sums $\sum_{C_1}$ and $\sum_{C_2}$ contain no common terms if $C_1 \neq C_2$. Consider now the expectation

$$E[\prod_{i=1}^{t} \prod_{j\neq i}^{t} (c_i V_{ij})^{s_{ij}}]$$

where the i and j belong to the collection C. Note that s_{ij} may be equal to 0 for some pairs (i, j). By Hölder's inequality we get, using the fact that $|V_{ij}| \leqq 2||\psi'||$,

$$(2.24) \qquad \begin{aligned} |E \prod_{i=1}^{t} \prod_{j\neq i}^{t} (c_i V_{ij})^{s_{ij}}| &\leqq \prod_{i=1}^{t} \prod_{j\neq i}^{t} |c_i|^{s_{ij}} \{E[(V_{ij})^{2r}]\}^{s_{ij}/2r} \\ &\leqq 2^{2r} ||\psi'||^{2r} \prod_{i=1}^{t} |c_i|^{s_i} \end{aligned}$$

where

$$(2.25) \qquad s_i = \sum_{j=1}^{t} s_{ij}, \qquad \sum_{i=1}^{t} s_i = 2r .$$

The partial sum corresponding to C is then estimated by

$$(2.26) \qquad \sum_C' \frac{(2r)!}{\prod_{i=1}^{t} \prod_{i\neq j}^{t} (s_{ij})!} (2^{2r} ||\psi'||^{2r} \prod_{i=1}^{t} |c_i|^{s_i}) .$$

Note that $(2r)!/\prod_{i=1}^{t} \prod_{i\neq j}^{t} (s_{ij})!$ is an integer. Hence we have

$$N(t) = \sum_C' \frac{(2r)!}{\prod_{i=1}^{t} \prod_{j\neq i}^{t} (s_{ij})!}$$

terms in the class C which are estimated by (2.24). Let $\mathscr{C}_t$ be the set of all terms

$$\sum \prod_{i=1}^{n} \prod_{j\neq 1}^{n} (c_i V_{ij})^{s_{ij}}$$

$s_2, \cdots, s_t)$ in (2.26) be given, $0 \leqq s_1 \leqq s_2 < \cdots \leqq s_t$, $\sum_{i=1}^t s_i = 2r$. Then according to the symmetry the set $\mathscr{C}_t$ contains a sum of terms, each estimated by

$$2^{2r}\|\psi'\|^{2r} \prod_{i=1}^t |c_{k_i}|^{s_i} \tag{2.27}$$

where $(k_1 \cdots k_t)$ is any combination of numbers $1, 2, \cdots, n$ to the tth class and in any order within this class. Let the number of terms in C_t for a fixed vector $(s_1, s_2, \cdots, s_t)$ as above be $n(t)$ and the sum of terms (2.27) belonging to $(s_1, s_2, \cdots, s_t)$ be $A(s_1, s_2, \cdots, s_t)$. (Note that $n(t)$ depends on $s_1, \cdots, s_t$.) Then, since $A(s_1, \cdots, s_t)$ is a symmetrical function

$$A(s_1, s_2, \cdots, s_t) = \frac{n(t)}{n!} \sum{}' 2^{2r}\|\psi'\|^{2r} \prod_{i=1}^t |c_{k_i}|^{s_i} \tag{2.28}$$

where $\sum'$ is the sum all terms belonging to all permutations of the numbers $1, 2, \cdots, n$. By Hölder's inequality we get, observing that

$$|c_{k_i}|^{s_i} = [c_{k_i}^{2r}]^{s_i/2r}, \qquad \sum_{i=1}^t \frac{s_i}{2r} = 1,$$

$$\sum{}' \prod_{i=1}^t |c_{k_i}|^{s_i} \leqq \prod_{i=1}^t (\sum{}' c_{k_i}^{2r})^{s_i/2r} \tag{2.29}$$

and here

$$\sum{}' c_{k_i}^{2r} = \frac{n!}{n} \sum_{i=1}^n c_i^{2r}.$$

Hence we obtain by (2.28) and (2.29)

$$A(s_1, s_2, \cdots, s_t) \leqq 2^{2r}\|\psi'\|^{2r} \cdot n(t) \cdot \frac{1}{n} \sum_{i=1}^n c_i^{2r}.$$

Since $\mathscr{C}_t$ contains $\binom{n}{t}N(t)$ terms we then find that $\mathscr{C}_t$ gives at most the contribution

$$n^{-2r}2^{2r}\|\psi'\|^{2r}\binom{n}{t}N(t) \cdot \frac{1}{n} \sum_{i=1}^n c_i^{2r}$$

to the right-hand side of (2.20). Putting

$$\Gamma_{nr}^{2r} = \frac{1}{n} \sum_{i=1}^n c_i^{2r}, \qquad \Gamma_{nr} \geqq 0,$$

and regarding the sets $\mathscr{C}_t$ for $t = 1, 2, \cdots, 2r$, we obtain from (2.20) that

$$E[(T_n - \hat{T}_n)^{2r}] \leqq 2^{2r}n^{-2r}\|\psi'\|^{2r}\Gamma_{nr}^{2r} \sum_{t=1}^{2r}\binom{n}{t}N(t). \tag{2.30}$$

We estimate $N(t)$ in the following way. Consider the identity

$$(\sum_{i=1}^t \sum_{j=i}^t x_i x_j)^{2r} = \sum \frac{(2r)!}{\prod_{i=1}^t \prod_{j=1}^t (s_{ij})!} \prod_{i=1}^t \prod_{j\neq i}^t (x_i x_j)^{s_{ij}}. \tag{2.31}$$

If an index k gives the contribution $\geqq 1$ to the sum (2.21), i.e., to the sum

$$\sum_{i=1}^t \sum_{j\neq i}^t s_{ij} = 2r,$$

then the double product

$$\prod_{j=1}^{t} \prod_{j\neq i}^{t} (x_i x_j)^{s_{ij}}$$

contains x_k as factor at least in the power 2. Hence differentiating the identity twice with respect to each x_k, $k = 1, 2, \cdots, t$ and then putting all x_n equal to 1 we get the inequality

$$2^t N(t) \leqq \left\{\prod_{k=1}^{t} \frac{\partial^2}{\partial_{x_k}} \left(\sum_{i=1}^{t} \sum_{j\neq i}^{t} x_i x_j\right)^{2r}\right\}_{x_k=1, k=1,2,\cdots,t} . \tag{2.32}$$

The right-hand side, however, is at most equal to

$$\left\{\prod_{k=1}^{t} \frac{\partial^2}{\partial x_k} \left(\left(\sum_{i=1}^{t} x_i\right)^{4r}\right)\right\}_{x_k=1, k=1,\cdots,t} = \frac{(4r)!}{(4r-2t)!} t^{4r-2t} . \tag{2.33}$$

Combining (2.30), (2.32) and (2.33), we get

$$E[(T_n - \hat{T}_n)^{2r}] \leqq c(r) ||\psi'||^{2r} \Gamma_{nr}^{2r}$$

with

$$c(r) = 2^{2r} n^{-2r} \sum_{t=1}^{2r} \binom{n}{t} \frac{(4r)!}{(4r-2t)!} t^{4r-2t} \cdot 2^{-t}$$

$$\Gamma_{nr}^{2r} = \frac{1}{n} \sum_{i=1}^{n} |c_i|^{2r} .$$

We estimate $c(r)$ exactly in the same way as we have estimated $b(r)$ in Lemma 2.1 and then obtain for $u = 2r - t$

$$c(r) \leqq \sum_{u=0}^{r-1} k(u)$$

with

$$k(u) = n^{-u} \frac{(4r)!}{(2u)!\,(2r-u)!} (2r-u)^{2u} \cdot 2^u .$$

Hence

$$k(0) = \frac{(4r)!}{(2r)!}, \qquad k(1) < n^{-1} \cdot (2r)^3 \frac{(4r)!}{(2r)!}$$

and for $u \geqq 1$

$$\frac{k(u+1)}{k(u)} \leqq \tfrac{4}{3} n^{-1} r^3 \leqq \tfrac{1}{2} \qquad \text{for} \quad n^{-1} r^3 \leqq \tfrac{3}{8} .$$

Hence for $n^{-1} r^3 \leqq \frac{3}{8}$

$$c(r) \leqq \frac{(4r)!}{(2r)!} [1 + 8n^{-1} r^3] .$$

Thus we have proved (2.13) and (2.14) of the lemma.

It follows by the definition of T_n that

$$S_n - T_n = \sum_{i=1}^{n} c_i [\xi_i - E(\xi_i)]$$

with

$$|\xi_i| \leqq \tfrac{1}{2} (\rho_i - \rho_{ii})^2 ||\psi''|| .$$

Hence

$$E[(S_n - T_n)^{2r}] \leqq n^{2r-1} \sum_{i=1}^n c_i^{2r} E[(\xi_i - E\xi_i)^{2r}]$$

and by Lemma 2.1

$$E[(\xi_i - E(\xi_i))^{2r}] \leqq 2^{2r}E[\xi_i^{2r}] \leqq \|\psi''\|^{2r}E[(\rho_i - \rho_{ii})^{4r}] \leqq n^{-2r}b(2r)\|\psi''\|^{2r} .$$

Thus we get (2.11)

$$E[(S_n - T_n)^{2r}] \leqq b(2r)\Gamma_{nr} .$$

By Minkovski's inequality we obtain (2.12) from (2.10) and (2.11)

$$E^{1/2r}[(S_n - \hat{T}_n)^{2r}] \leqq E^{1/2r}[(S_n - T_n)^{2r}] + E^{1/2r}[(T_n - \hat{T}_n]^{2r}] .$$

LEMMA 2.3. $\hat{T}_n = \sum_{j=1}^n \hat{T}_n^{(j)}$ *with independent random variables*

(i) $$\hat{T}_n^{(j)} = c_j\{\psi(\rho_{jj}) - E[\psi(\rho_{jj})]\} + \frac{1}{n} \sum_{i\neq j}^n c_i[E(u(X_i - X_j) - F_j(X_i))\psi'(\rho_{ii}) \,|\, X_j] .$$

Further,

(ii) $$\sum_{j=1}^n [E|\hat{T}_n^{(j)}|^3] \leqq 4[2\|\psi\|^3 + \|\psi'\|^3\} \sum_{j=1}^n |c_i|^3 .$$

PROOF. We get the representation (i) by (2.16). Using well-known inequalities

$$|(a + b)^3| \leqq 4[|a|^3 + |b|^3] , \qquad |(\textstyle\sum_{i=1}^n a_i)^3| \leqq n^2 \sum_{i=1}^n |a_i|^3$$

we obtain

$$E[|\hat{T}_n^{(j)}|^3] \leqq 4|c_j|^3E[|[\psi(\rho_{jj})] - E\psi(\rho_{jj})|^3] + \frac{4}{n} \sum_{i\neq j}^n |c_i|^3\|\psi'\|^3 .$$

Here

$$E[|\psi(\rho_{jj}) - E[\psi(\rho_{jj})]|^3] \leqq 2\|\psi\|E(\psi(\rho_{jj}) - E(\psi(\rho_{jj}))^2 .$$

Thus we get (ii).

3. Proofs of the theorems.

(a) PROOF OF THEOREM 1.1. (1.10) follows from Berry-Esseen's inequality and Lemma 2.3 and (1.11) from Lemma 2.2 (2.12).

(b) PROOF OF THEOREM 1.2. For $h > 0$ we get

(3.1) $$P[S_n \leqq \hat{\delta}_n x] \leqq P(S_n \leqq \hat{\delta}_n x, |S_n - \hat{T}_n| < h\hat{\delta}_n) + P[|S_n - \hat{T}_n| \geqq h\delta_n] \leqq P[\hat{T}_n \leqq \hat{\delta}_n(x + h)] + P[|S_n - \hat{T}_n| \geqq h\hat{\delta}_n] .$$

Applying Theorem 1.1 we get

(3.2) $$P[\hat{T}_n \leqq \hat{\delta}_n(x + h)] \leqq \Phi(x + h) + 4C(2\|\psi\|^3 + \|\psi'\|^3) \cdot \sum_{i=1}^n |c_i|^3\hat{\delta}_n^{-3} .$$

Here

(3.3) $$\Phi(x + h) \leqq \Phi(x) + \|\Phi'(x)\| = \Phi(x) + \frac{h}{(2\pi)^{\frac{1}{2}}} .$$

By Chebyshev's inequality and the inequality (2.12) of Lemma 2.2 we get

(3.4) $$P[|S_n - \hat{T}_n| \geqq h\hat{\delta}_n] \leqq d(r, \psi)\Gamma_{nr}^{2r}(h\hat{\delta}_n)^{-2r} .$$

Now we choose n such that

$$\frac{h}{(2\pi)^{\frac{1}{2}}} = d(r, \psi)\Gamma_{nr}^{2r}(h\hat{\delta}_n)^{-2r},$$

i.e.,

(3.5) $$h = [(2\pi)^{\frac{1}{2}} d(r, \psi)\hat{\delta}_n^{-2r}\Gamma_{nr}^{2r}]^{1/(2r+1)}.$$

It follows by Lemma 2.2, (2.12), (2.13) and (2.14), and the remark made after Lemma 2.2 that for $n^{-1}r^3 \leqq \frac{3}{8}$

$$[d(r, \psi)]^{1/2r} \leqq C'r(\|\psi'\| + \|\psi''\|)$$

with an absolute constant C'. Then it follows by (3.4) and (3.5) that

$$\frac{h}{(2\pi)^{\frac{1}{2}}} + d(r, \psi)\Gamma_{nr}^{2r}(h\hat{\delta}_n)^{-2r} \leqq C_2[\hat{\delta}_n^{-1}(\|\psi'\| + \|\psi''\|)r\Gamma_{nr}]^{2r/(2r+1)}.$$

By (3.1)—(3.6) we get the inequality (1.12) in one direction. It follows for the other direction in the same way.

Acknowledgment. The authors would like to express their sincere appreciation to Professor I. R. Savage, the associate editor, and the referees for the critical examination of the original draft. Their constructive criticism, detailed comments and suggestions for improvements are gratefully acknowledged.

REFERENCES

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

[2] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math Statist.* **39** 325–346.

[3] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.

[4] Jurečková, J. and Puri, Madan L. (1975). Order of normal approximation for rank test statistics distribution. *Ann. Probability* **3** 526–533.

[5] Zolotarev, V. M. (1967). A sharpening of the inequality of Berry-Esseen. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 332–342.

Department of Mathematics
Indiana University
Swain Hall, East
Bloomington, IN 47401

INVARIANCE PRINCIPLES FOR RANK STATISTICS FOR TESTING INDEPENDENCE

MADAN L. PURI* and LANH T. TRAN†

ABSTRACT

In this paper we consider a general class of rank order statistics for testing independence in bivariate populations. Each statistic is represented as a sum of independent and identically distributed (i.i.d) random variables and a remainder term. Suitable order (a.s.) of the remainder term is found and then some invariance principles are obtained. The results obtained are extensions of the results of Chernoff and Savage (1958), Bhuchongkul (1964), Bahadur (1966), Ruymgaart *et al.* (1972), Sen and Ghosh (1974) and Lai (1975).

Key words and phrases: Invariance principles, linear rank statistics, score functions.

* Research supported by the Air Force Office of Scientific Research AFSC, USAF, under Grant No. AFOSR76-2927. This research was done at the University of Washington. during my leave of absence from Indiana University. I appreciate very much the use of facilities that were made available to me by the Department of Biostatistics at the University of Washington.

† Research supported by National Science Foundation Grant No. MCS76-00951. Reproduction in whole or part is permitted for any purpose of the U.S. Government.

1. Introduction

Let $\{(X_i, Y_i), 1 \le i \le N\}$ be N independent and identically distributed i.i.d. random vectors, each having a continuous cumulative distribution function (c.d.f.) $H(x, y)$. Let $F(x)$ and $G(y)$ denote the marginal c.d.f.'s of X_i and Y_i, respectively. Denote by $F_N(x)$, $G_N(y)$, and $H_N(x, y)$, the empirical c.d.f.'s of $\{X_i, 1 \le i \le N\}$, $\{Y_i, 1 \le i \le N\}$, and $\{(X_i, Y_i), 1 \le i \le N\}$, respectively. Finally, let R_{Ni} (and S_{Ni}) denote the rank of X_i (and Y_i) among $X_i, 1 \le i \le N$ (and Y_i, $1 \le i \le N$). Then many rank tests for the hypothesis of independence

$$H_0 : H(x, y) = F(x)G(y) \tag{1}$$

are based on the statistic

$$T_N = N^{-1} \sum_{i=1}^{N} J_N\left(\frac{R_{Ni}}{N}\right) L_N\left(\frac{S_{Ni}}{N}\right) = \int J_N[F_N(x)] L_N[G_N(y)]\, dH_N(x, y) \tag{2}$$

where $J_N(i/N) = EJ(U_{Ni})$ or $J(i/N + 1)$, $L_N(i/N) = EL(U_{Ni})$ or $L(i/(N + 1))$, U_{Ni}, $(1 \le i \le N)$ is the ith order statistic in a sample of size N from the uniform distribution over $(0, 1)$, and $J(u)$, $L(u)$, $0 < u < 1$, are nondecreasing, twice differentiable score functions [cf. Bhuchongkul (1964)].

For the case $J_N(i/N) = EJ(U_{Ni})$ and $L_N(i/N) = EL(U_{Ni})$, Sen and Ghosh (1974) have obtained some invariance principles for $\{T_N\}$ when the null hypothesis H_0 in (1) holds. Their results are based on a fundamental martingale property possessed by $\{T_N, \mathscr{F}_N\}$ when $\{R_{Ni}, 1 \le i \le N\}$ and $\{S_{Ni}, 1 \le i \le N\}$ are stochastically independent. Here $\mathscr{F}_N$ denotes the σ field generated by $\{R_{Ni}, S_{Ni}; 1 \le i \le N\}$.

In this paper the invariance principles are established for $\{T_N\}$ *under alternatives.* When $H(x, y) \ne F(x)G(y)$, the techniques of Sen and Ghosh (1974) are not applicable since $\{T_N, \mathscr{F}_N\}$ is not a martingale. Our methods are related to those of Chernoff and Savage (1958), Bhuchongkul (1964), Bahadur (1966), Sen and Ghosh (1973), Ruymgaart *et al.* (1972), and Lai (1975). The main argument is based on a representation of T_N as the sum of i.i.d. random variables and a remainder term which is shown to converge a.s. to zero at an appropriate rate. The contents of this paper are as follows:

In Sec. 2, assumptions on the score functions are stated and preliminary lemmas are presented. Section 3 deals with the order of magnitude of the remainder term. Some invariance principles are then established in Sec. 4. In what follows K is used as a generic constant whose values may differ from line to line.

2. Assumptions and Some Preliminary Lemmas

ASSUMPTION 2.1. $J(u)$ and $L(u)$, $0 < u < 1$, are absolutely continuous, twice differentiable score functions, with

$$|J^{(i)}(u)| \le K[u(1-u)]^{-\alpha-i}, \; |L^{(i)}(u)| \le K[u(1-u)]^{-\beta-i}, \qquad i = 0, 1, 2, \tag{3}$$

where

$$\alpha = (1 - 2\delta)/2p, \qquad \beta = (1 - 2\delta)/2q \tag{4}$$

for some $0 < \delta < \frac{1}{2}$ and some $p, q > 1$ with $p^{-1} + q^{-1} = 1$.

We shall start with the following lemmas which are slight variations or generalizations of some of the results of Bahadur (1966), Sen (1972), Sen and Ghosh (1974), and which we shall need in sequel.

LEMMA 2.1. *Let X_i, $1 \le i \le N$ be i.i.d. random variables, each having a continuous c.d.f. $F(x)$. Let $U_N(u)$ be the empirical c.d.f. of $\{F(X_i), 1 \le i \le N\}$. Then for every $\varepsilon > 0$,*

$$\sup N^{1/2}\{u(1-u)\}^{\varepsilon - 1/2}|U_N(u) - u| = o(\log N) \qquad \text{a.s.} \tag{5}$$

Proof. Follows from Lemma 2.1 of Sen and Ghosh (1974).

LEMMA 2.2. *Let* $F_{N,\theta}(x) = \theta N F_N(x)/(1 + N) + (1 - \theta)F(x), \quad 0 \le \theta \le 1.$ *Then*

$$1 - F_{N,\theta}(x) \ge \{1 - F(x)\}\{1 - 0(1)\} \text{ a.s.} = \{1 - F_N(x)\}\{1 - 0(1)\} \text{ a.s.} \tag{6}$$

as $N \to \infty$ for $F^{-1}(N^{-1+\lambda}) \le x \le F^{-1}(1 - N^{-1+\lambda})$, where λ is an arbitrary positive number < 1.

Proof. Follows from Sen and Ghosh (1974, p. 164).

LEMMA 2.3. *Let $D = (0,1) \times (0,1)$ and let $\bar{w} = (u,v) \in D$. If $H(\bar{w})$ is a continuous c.d.f. with uniform $(0,1)$ marginal c.d.f.'s, then*

$$\begin{aligned}\sup_{\bar{w} \in D} \sup_{\bar{w}' \in D} &\{|H_N(\bar{w}') - H_N(\bar{w}) - H(\bar{w}') + H(\bar{w})| : |\bar{w}' - \bar{w}| \le N^{-1/2}\} \\ &= O(N^{-3/4} \log N) \text{ a.s.}\end{aligned} \tag{7}$$

Proof. By a straightforward generalization of Lemma 1 of Bahadur (1966), we have

$$\begin{aligned}\sup_{\bar{w}' \in D} &\{|H_N(\bar{w}') - H_N(\bar{w}) - H(\bar{w}') + H(\bar{w})| : |\bar{w}' - \bar{w}| \le N^{-1/2}\} \\ &= O(N^{-3/4} \log N) \text{ a.s.,}\end{aligned}$$

where $\bar{w} \in (0,1) \times (0,1)$ is an arbitrary fixed point. Then (7) can be obtained by the same line of argument as in Theorem 4.2 of Sen and Ghosh (1971).

LEMMA 2.4. *Let $0 < \lambda < 1$. Then*

$$|F_N(F^{-1}(1 - N^{-1+\lambda})) - (1 - N^{-1+\lambda})| = O(N^{-1+\lambda/2} \log N) \text{ a.s. as } N \to \infty.$$

Proof. Follows from a slight variation of Lemma 4.1 of Sen (1972).

3. Order of Magnitude of the Remainder Term

THEOREM 3.1. *Let T_N defined in* (2) *be written as*

$$T_N = \sum_{i=1}^{3} A_{iN} + R_N, \tag{8}$$

where

$$A_{1N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J[F(x)]L[G(y)]\,dH_N(x, y), \tag{9}$$

$$A_{2N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_N(x) - F(x)]J'[F(x)]L(G(y)]\,dH(x, y), \tag{10}$$

$$A_{3N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [G_N(y) - G(y)]L'[G(y)]J[F(x)]\,dH(x, y), \tag{11}$$

and R_N is defined below (13). *Then, under the assumption* 2.1, $\lim_{N\to\infty} N^{1/2}R_N = 0$ a.s.

Proof. Define a_{1N}, b_{1N}, c_{1N}, and d_{1N} by

$$F(a_{1N}) = N^{-1+\delta_1};\ F(b_{1N}) = 1 - N^{-1+\delta_1};$$
$$G(c_{1N}) = N^{-1+\delta_2},\ G(d_{1N}) = 1 - N^{-1+\delta_2},$$

where $\delta_1 = \delta p^{-1}$ and $\delta_2 = \delta q^{-1}$.

Let ξ be any positive number smaller than $\frac{1}{4}$. Pick γ_1 and γ_2 with $0 < \gamma_1 < \min(\xi(1+\alpha)^{-1}, (1-2\beta)(8\alpha)^{-1})$ and $0 < \gamma_2 < \min(\xi(1+\beta)^{-1}, (1-2\alpha)(8\beta)^{-1})$. Define a_{2N}, b_{2N}, c_{2N}, and d_{2N} by $F(a_{2N}) = N^{-\gamma_1}$, $F(b_{2N}) = 1 - N^{-\gamma_1}$, $G(c_{2N}) = N^{-\gamma_2}$ and $G(d_{2N}) = 1 - N^{-\gamma_2}$. Let

$$I_{1,N} = [a_{1N}, b_{1N}] \times [c_{1N}, d_{1N}],\ I_{2N} = [a_{2N}, b_{2N}] \times [c_{2N}, d_{2N}]. \tag{12}$$

and denote the complements of I_{1N} and I_{2N} by I_{1N}^c and I_{2N}^c, respectively. Then from the decomposition (8) we have

$$R_N = \sum_{i=1}^{17} B_{i,N}, \tag{13}$$

where

$$B_{1N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{L_N[G_N(y)] - L[NG_N(y)/(N+1)]\}\{J_N[F_N(x)] - J[NF_N(x)/(N+1)]\}\,dH_N(x, y),$$

$$B_{2N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L[NG_N(y)/(N+1)][J_N[F_N(x)] - J(NF_N(x)/(N+1))]\,dH_N(x, y),$$

$$B_{3N} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J[NF_N(x)/(N+1)]\{L_N[G_N(y)] - L[NG_N(y)/(N+1)]\}\,dH_N(x, y),$$

$$B_{4N} = \iint_{I_{1,N}} \{L[NG_N(y)/(N+1)] - L[G(y)]\}\{J[NF_N(x)/(N+1)] - J[F(x)]\}\, dH_N(x, y),$$

$$B_{5N} = \iint_{I_{1,N}^c} \{L[NG_N(y)/(N+1)] - L[G(y)]\}\{J[NF_N(x)/(N+1)] - J[F(x)]\}\, dH_N(x, y),$$

$$B_{6N} = \iint_{I_{1,N}} L[G(y)]\{J[NF_N(x)/(N+1)] - J[F(x)] - J'[F(x)]\}\{NF_N(x)/(N+1) - F(x)\}\, dH_N(x, y),$$

$$B_{7N} = \iint_{I_{1,N}} J[F(x)]\{L[NG_N(y)/(N+1)] - L[G(y)] - L'[G(y)][NG_N(y)/(N+1) - G(y)]\}\, dH_N(x, y),$$

$$B_{8N} = \iint_{I_{1,N}^c} L[G(y)]\{J[NF_N(x)/(N+1)] - J[F(x)]\}\, dH_N(x, y),$$

$$B_{9N} = \iint_{I_{1,N}^c} J[F(x)]\{L[NG_N(y)/(N+1)] - L[G(y)]\}\, dH_N(x, y),$$

$$B_{10N} = -\iint_{I_{2,N}^c} L[G(y)]J'[F(x)]\{F_N(x) - F(x)\}\, dH(x, y),$$

$$B_{11N} = -\iint_{I_{2,N}^c} J[F(x)]L'[G(y)]\{G_N(y) - G(y)\}\, dH(x, y),$$

$$B_{12N} = \iint_{I_{1,N}\cap I_{2,N}^c} L[G(y)]J'[F(x)]\{F_N(x) - F(x)\}\, dH_N(x, y),$$

$$B_{13N} = \iint_{I_{1,N}\cap I_{2,N}^c} J[F(x)]L'[G(y)]\{G_N(y) - G(y)\}\, dH_N(x, y),$$

$$B_{14N} = \iint_{I_{2,N}} L[G(y)]J'[F(x)][F_N(x) - F(x)]\, d[H_N(x, y) - H(x, y)],$$

$$B_{15N} = \iint_{I_{2,N}} J[F(x)]L'[G(y)][G_N(y) - G(y)]\, d[H_N(x, y) - H(x, y)],$$

$$B_{16N} = -(N+1)^{-1} \iint_{I_{1,N}} L[G(y)]J'[F(x)]F_N(x)\, dH_N(x, y),$$

$$B_{17N} = -(N+1)^{-1} \iint_{I_{1,N}} J[F(x)]L'[G(y)]G_N(y)\, dH_N(x, y).$$

For each $1 \le i \le 17$, we shall show that $|B_{i,N}| = O(N^{-1/2-\eta})$ for some $\eta > 0$.

For reasons of symmetry, we do not need to treat $B_{3N}, B_{7N}, B_{9N}, B_{11N}, B_{15N}, B_{17N}$.

First, consider B_{1N}. By Holder's inequality, $|B_{1N}|$ is bounded by

$$\left(\int_{-\infty}^{\infty} |J_N[F_N(x)] - J[NF_N(x)/(N+1)]|^p \, dF_N(x)\right)^{1/p}$$
$$\times \left(\int_{-\infty}^{\infty} |L_N[G_N(y)] - L[NG_N(y)/(N+1)]|^q \, dG_N(y)\right)^{1/q}. \qquad (14)$$

From Chernoff and Savage (1958, Theorem 2), it follows that

$$\begin{aligned} &|J_N(1/N) - J(1/(N+1))| < KN^{\alpha}; \qquad |J_N(i/N) - J(i/(N+1))| \\ &\quad < KN^{\alpha}[\Phi(-\sqrt{i}/K) + i^{-1-\alpha}], \qquad 1 < i \le N/2. \end{aligned} \qquad (15)$$

Thus

$$\int_{1 \le NF_N(x) \le N/2} |J_N[F_N(x)] - J[NF_N(x)/(N+1)]|^p \, dF_N(x)$$

is bounded by

$$N^{-1}\left(KN^{\alpha p} + \sum_{i=2}^{N/2} KN^{\alpha p}[\Phi(-\sqrt{i}/K) + i^{-1-\alpha}]^p\right) \le KN^{\alpha p - 1}.$$

By a symmetric argument we can cover the range $N/2 \le NF_N(x) \le N$. Hence the first factor of (14) is bounded by $KN^{(\alpha p - 1)p^{-1}}$. Similarly the second factor of (14) is bounded by $KN^{(\beta q - 1)q^{-1}}$. Thus

$$\begin{aligned} |B_{1N}| &= O(N^{(\alpha p - 1)p^{-1} + (\beta q - 1)q^{-1}}) \\ &= O(N^{-\delta_1 - \delta_2 - 1/2}) = O(N^{-\eta - 1/2}) \qquad \text{for} \quad \eta \le \delta_1 + \delta_2. \end{aligned} \qquad (16)$$

We now consider B_{2N}. By (3), it is clear that $L[NG_N(y)/(N+1)] \le KN^{\beta}$ for $N^{-1} \le G_N(y) \le 1$. Hence $|B_{2N}|$ is bounded by

$$KN^{\beta} \int_{-\infty}^{\infty} |J_N(F_N(x)] - J[NF_N(x)/(N+1)]| \, dF_N(x). \qquad (17)$$

Using (15), it follows that (17) is bounded by

$$KN^{\beta}N^{-1}\left(KN^{\alpha} + \sum_{i=2}^{N/2} KN^{\alpha}[\Phi(-\sqrt{i}/K) + i^{-1-\alpha}]\right) \le KN^{-\delta_1 - \delta_2 - 1/2}.$$

Consequently

$$|B_{2N}| = O(N^{-\eta - 1/2}) \qquad \text{for} \quad \eta < \delta_1 + \delta_2. \qquad (18)$$

We now consider B_{4N}. By Holder's inequality, $|B_{4N}|$ is bounded by

$$\left(\int_{a_{1N}}^{b_{1N}} |J[NF_N(x)/(N+1)] - J[F(x)]|^p \, dF_N(x)\right)^{p^{-1}}$$
$$\times \left(\int_{c_{1N}}^{d_{1N}} |L[NG_N(y)/(N+1)] - L[G(y)]|^q \, dG_N(y)\right)^{q^{-1}} \qquad (19)$$

By the mean value theorem, the integral in the first factor of (19) is bounded by

$$\int_{a_{1N}}^{b_{1N}} |NF_N(x)/(N+1) - F(x)|^p \, |J'[F_{N,\theta}(x)]|^p \, dF_N(x), \tag{20}$$

where $F_{N,\theta}(x)$ is defined in Lemma 2.2.

Let ε be a positive number to be specified later. By (3) and Lemma 2.1, (20) is bounded by

$$O(N^{-p/2}(\log N)^p) \times \int_{a_{1N}}^{b_{1N}} [F(x)(1-F(x)]^{p(-\varepsilon+1/2)}[F_{N,\theta}(x)(1-F_{N,\theta}(x))]^{-p(1+\alpha)} \, dF_N(x). \tag{21}$$

Pick a with $F(a) = \frac{1}{2}$. Then (21) is bounded by

$$O(N^{-p/2}(\log N)^p) \int_{a}^{b_{1N}} [1-F(x)]^{p(-\varepsilon+1/2)}[F_{N,\theta}(x)(1-F_{N,\theta}(x)]^{-p(1+\alpha)} \, dF_N(x) + O(N^{-p/2}(\log N)^p) \int_{a_{1N}}^{a} [F(x)]^{p(-\varepsilon+1/2)}[F_{N,\theta}(x)]^{-p(1+\alpha)} \, dF_N(x). \tag{22}$$

By Lemma 2.2, the first term of (22) is bounded by

$$O(N^{-p/2}(\log N)^p) \int_{a}^{b_{1N}} [1-F_N(x)]^{-p(\varepsilon+\alpha+1/2)} \, dF_N(x). \tag{23}$$

Observe that the integral in (23) tends to infinity as $N \to \infty$ if and only if $(\frac{1}{2}+\alpha+\varepsilon)p \geq 1$. For simplicity, we pick $\varepsilon > \delta_1$. Then $(\frac{1}{2}+\alpha+\varepsilon)p > 1$ and only the case where this integral tends to infinity has to be considered. By a simple integration, we note that (23) is bounded by

$$O(N^{-p/2}(\log N)^p)(1-F_N(b_{1N}))^{-(1/2+\alpha+\varepsilon)p+1}. \tag{24}$$

By Lemma 2.4 it can easily be seen that (24) is bounded by

$$\begin{aligned} &O(N^{-p/2}(\log N)^p)[N^{-1+\delta_1} \pm 0\,(N^{-1+\delta_1/2}\log N)]^{-(1/2+\alpha+\varepsilon)p+1} \\ &= O(N^{-p/2}(\log N)^p)N^{(-1+\delta_1)[-(1/2+\alpha+\varepsilon)p+1]} \\ &= O(N^{(-1+\delta_1)(-p\alpha+1)}(\log N)^p)N^{-p[2^{-1}\delta_1-\varepsilon(1-\delta_1)]} \\ &= O(N^{(-1+\delta_1)(-p\alpha+1)}) \end{aligned}$$

by choosing $\varepsilon < \delta_1\,(1-\delta_1)^{-1}2^{-1}$. Clearly, the second term of (22) is also bounded by $O(N^{(-1+\delta_1)(-p\alpha+1)})$. Hence, the first factor of (19) is bounded by $O(N^{(-1+\delta_1)(-p\alpha+1)p^{-1}})$. Similarly, the second factor of (19) is bounded by $O(N^{(-1+\delta_2)(-q\beta+1)q^{-1}})$. Finally,

$$\begin{aligned} |B_{4N}| &= O(N^{-1/2-\delta_1(1-2^{-1}p^{-1}-\delta_1)-\delta_2(1+2^{-1}q^{-1}+\delta_2)}) \\ &= O(N^{-\eta-1/2}) \qquad \text{for some } \eta > 0. \end{aligned}$$

Next, consider B_{5N}. Observe that $I_{1N}^c = \bigcup_{i=1}^8 I_{1,iN}$, where

$$I_{1,1N} = (b_{1N}, \infty) \times (d_{1N}, \infty), \qquad I_{1,2N} = (-\infty, a_{1N}) \times (-\infty, c_{1N}),$$
$$I_{1,3N} = (b_{1N}, \infty) \times (-\infty, c_{1N}), \qquad I_{1,4N} = (-\infty, a_{1N}) \times (d_{1N}, \infty),$$
$$I_{1,5N} = (a_{1N}, b_{1N}) \times (d_{1N}, \infty), \qquad I_{1,6N} = (a_{1N}, b_{1N}) \times (-\infty, c_{1N}),$$
$$I_{1,7N} = (b_{1N}, \infty) \times (c_{1N}, d_{1N}), \qquad I_{1,8N} = (-\infty, a_{1N}) \times (c_{1N}, d_{1N}).$$

Define

$$B_{5,iN} = \iint_{I_{1,iN}} \left\{ L\left[\frac{NG_N(y)}{(N+1)}\right] - L[G(y)] \right\} \left\{ J\left[\frac{NF_N(x)}{(N+1)}\right] - J[F(x)] \right\} dH_N(x, y).$$

Then $B_{5N} = \sum_{i=1}^8 B_{5,iN}$.

First, consider $B_{5,1N}$. By Holder's inequality, $|B_{5,1N}|$ is bounded by

$$\left(\int_{b_{1N}}^{\infty} |J[NF_N(x)/(N+1)] - J[F(x)]|^p \, dF_N(x) \right)^{1/p} \times \left(\int_{d_{1N}}^{\infty} |L[NG_N(y)/(N+1)] - L[G(y)]|^q \, dG_N(y) \right)^{1/q}. \tag{25}$$

The integral in the first factor of (25) is bounded by

$$K \int_{b_{1N}}^{\infty} |J[NF_N(x)/(N+1)]|^p \, dF_N(x) + K \int_{b_{1N}}^{\infty} |J[F(x)]|^p \, dF_N(x). \tag{26}$$

By (3), the first term of (26) is smaller than

$$K \int_{b_{1N}}^{\infty} [1 - NF_N(x)/(N+1)]^{-p\alpha} \, dF_N(x),$$

which upon integration equals

$$K(1 - N(N+1)^{-1} F_N(b_N))^{-p\alpha+1}. \tag{27}$$

Lemma 2.4 now implies that (27) equals

$$K(1 - N(N+1)^{-1}(F(b_N) \pm O(N^{-1+\delta_1/2} \log N))^{-p\alpha+1} = O(N^{(-1+\delta_1)(-p\alpha+1)}).$$

Next the second term of (26) which can be written as $-K \int_{b_{1N}}^{\infty} |J[F(x)]|^p \, d[1 - F_N(x)]$ is bounded by

$$K|J[F(b_N)]|^p [1 - F_N(b_N)] + K \int_{b_{1N}}^{\infty} [1 - F_N(x)] |J[F(x)]|^{p-1} |J'[F(x)]| \, dF(x) \tag{28}$$

upon integration by parts.

Now using (3) and some routine computations, it follows that the first term of (28) is bounded by $K(1 - F(b_N)]^{-p\alpha}[1 - F(b_N) + O(N^{-1+\delta_1/2} \log N)] = O(N^{(-1+\delta_1)(-p\alpha+1)})$.

By Lemma 2.1, the second term of (28) is bounded by

$$K \int_{b_{1N}}^{\infty} [1 - F(x)] |J(F(x))|^{p-1} |J'(F(x))|\, dF(x) + O(N^{-1/2} \log N) \times \int_{b_{1N}}^{\infty} [1 - F(x)]^{-\varepsilon - 1/2} |J(F(x))|^{p-1} |J'(F(x))|\, dF(x), \tag{29}$$

where ε is any arbitrary positive number.

Now using (3), it is easy to verify that (29) is bounded by

$$\begin{aligned} K \int_{b_{1N}}^{\infty} [1 - F(x)]^{-p\alpha} dF(x) &+ O(N^{-1/2} \log N) \int_{b_{1N}}^{\infty} [1 - F(x)]^{-1/2 - \alpha p - \varepsilon} dF(x) \\ &\le O(N^{(-1+\delta_1)(-p\alpha+1)}) + O(N^{-1/2} \log N)(N^{(-1+\delta_1)(1/2 - p\alpha - \varepsilon)}) \\ &= O(N^{(-1+\delta_1)(-p\alpha+1)}) \qquad \text{if} \quad \varepsilon < \delta_1 (1 - \delta_1)^{-1} 2^{-1}. \end{aligned}$$

It is now clear that the first factor of (25) equals $O(N^{(-1+\delta_1)(-p\alpha+1)p^{-1}})$. Similarly, the second factor equals $O(N^{(-1+\delta_2)(-q\beta+1)q^{-1}})$. Hence $|B_{5,1N}| = O(N^{-\eta - 1/2})$ for some $\eta > 0$.

Using the same arguments, $|B_{5,iN}|$ can be shown to be equal to $O(N^{-\eta - 1/2})$ for some $\eta > 0$, $i = 1, \ldots, 8$.

Next, consider $B_{6,N}$. By (3), it is easy to verify that $L[G(y)]$ is bounded by $KN^{\beta(1-\delta_2)}$ for $y \in [c_{1N}, d_{1N}]$. Hence, $|B_{6N}|$ is smaller than or equal to

$$KN^{\beta(1-\delta_2)} \int_{a_{1N}}^{b_{1N}} |J[NF_N(x)/(N+1)] - J[F(x)] - J'[F(x)][NF_N(x)/(N+1) - F(x)]|\, dF_N(x). \tag{30}$$

Using (30) and the mean value theorem,

$$|B_{6N}| \le KN^{\beta(1-\delta_2)} \int_{a_{1N}}^{b_{1N}} \{[NF_N(x)/(N+1)] - F(x)\}^2 |J''(F_{N,\theta}(x))]|\, dF_N(x), \tag{31}$$

where $F_{N,\theta}$ is defined in Lemma 2.2. Next, (3) and Lemma 2.2 imply that (31) is bounded by

$$N^{\beta(1-\delta_2)} O(N^{-1} (\log N)^2) \int_{a}^{b_{1N}} [1 - F(x)]^{1-2\varepsilon} [1 - F_{N,\theta}(x)]^{-2-\alpha} dF_N(x) + N^{\beta(1-\delta_2)} O(N^{-1} (\log N)^2) \int_{a_{1N}}^{a} [F(x)]^{1-2\varepsilon} [F_{N,\theta}(x)]^{-2-\alpha} dF_N(x) \tag{32}$$

for all $\varepsilon > 0$. By Lemma 2.2, the first term of (32) is bounded by

$$N^{\beta(1-\delta_2)} O(N^{-1} (\log N)^2) \int_{a}^{b_{1N}} [1 - F_N(x)]^{-1-\alpha-2\varepsilon} dF_N(x) \tag{33}$$

which, on integration, equals

$$N^{\beta(1-\delta_2)} O(N^{-1} (\log N)^2) O(N^{(1-\delta_1)(\alpha + 2\varepsilon)}) = O(N^{-\eta - 1/2})$$

for some $\eta > 0$ since ε can be chosen to be smaller than $(\delta_1 + \delta_2 + \alpha\delta_1 + \beta\delta_2)2^{-1}(1 - \delta_1)^{-1}$. Similarly, the second term of (32) equals $O(N^{-\eta-1/2})$ for some $\eta > 0$. Hence $|B_{6N}| = O(N^{-\eta-1/2})$ for some $\eta > 0$.

Consider B_{8N}. Define

$$B_{8,iN} = \iint_{I_{1,iN}} L[G(y)]\{J[NF_N(x)/(N+1)] - J[F(x)]\}\, dH_N(x, y).$$

Then $B_{8N} = \sum_{i=1}^{8} B_{8,iN}$.

By Holder's inequality,

$$|B_{8,1N}| \le \left(\int_{b_{1N}}^{\infty} |J[NF_N(x)/(N+1)] - J[F(x)]|^p\, dF_N(x)\right)^{1/p} \times \left(\int_{d_{1N}}^{\infty} |L[G(y)]|^q\, dG_N(y)\right)^{1/q}. \tag{34}$$

The first factor in (34) is identical to the first factor of (25), which has been shown to be equal to $O(N^{(-1+\delta_1)(-p\alpha+1)p^{-1}})$. Since the second term of (26) is bounded by $O(N^{(-1+\delta_1)(-p\alpha+1)})$, it follows easily that the second term of (34) is bounded by $O(N^{(-1+\delta_2)(-q\beta+1)q^{-1}})$. Thus $|B_{8,1N}| = O(N^{-\eta-1/2})$ for some $\eta > 0$. By similar treatments, we see that $|B_{8,2N}|$, $|B_{8,3N}|$, and $|B_{8,4N}|$ are all equal to $O(N^{-\eta-1/2})$ for sufficiently small η.

By Holder's inequality, $|B_{8,5N}|$ is bounded by

$$\left(\int_{a_{1N}}^{b_{1N}} |J[NF_N(x)/(N+1)] - J[F(x)]|^p\, dF_N(x)\right)^{1/p} \left(\int_{d_{1N}}^{\infty} |L[G(y)]|^q\, dG_N(y)\right)^{1/q}. \tag{35}$$

The first and second factor of (35) are respectively equal to the first factor of (19) and the second factor of (34). Therefore, $|B_{8,5N}| = O(N^{-\eta-1/2})$ for some $\eta > 0$. Similarly, $|B_{8,6N}| = O(N^{-\eta-1/2})$ for sufficiently small η.

Consider $B_{8,7N}$. By (3), $L[G(y)] \le KN^{\beta(1-\delta_2)}$ for $y \in [c_{1N}, d_{1N}]$. Hence $|B_{8,7N}|$ is bounded by

$$KN^{\beta(1-\delta_2)} \int_{b_{1N}}^{\infty} |J[NF_N(x)/(N+1)] - J[F(x)]|\, dF_N(x). \tag{36}$$

Recall that the first factor of (25) is equal to $O(N^{(-1+\delta_1)(-p\alpha+1)p^{-1}})$. It is easy to deduce that the integral in (26) equals $O(N^{(-1+\delta_1)(-\alpha p+1)})$. Hence $|B_{8,7N}| = O(N^{\beta(1-\delta_2)+(-1+\delta_1)(-\alpha p+1)}) = O(N^{-\eta-1/2})$ for some $\eta > 0$. One can treat $B_{8,8N}$ by using the same line of arguments. Finally, $|B_{8,N}| = O(N^{-\eta-1/2})$ for sufficiently small η.

Consider $B_{10,N}$. Write $I_{2,N}^c = \bigcup_{i=1}^{8} I_{2,iN}$, where $I_{2,1N} = (b_{2N,\infty}) \times (d_{2N,\infty})$, $I_{2,2N} = (-\infty, a_{2N}) \times (-\infty, c_{2N})$, $I_{2,3N} = (b_{2N,\infty}) \times (-\infty, c_{2N})$, $I_{2,4N} = (-\infty, a_{2N}) \times (d_{2N,\infty})$, $I_{2,5N} = (a_{2N}, b_{2N}) \times (d_{2N,\infty})$, $I_{2,6N} = (a_{2N}, b_{2N}) \times (-\infty, c_{2N})$, $I_{2,7N} = (b_{2N,\infty}) \times (c_{2N}, d_{2N})$, $I_{2,8N} = (-\infty, a_{2N}) \times (c_{2N}, d_{2N})$.

Define

$$B_{10,iN} = -\iint\limits_{I_{2,iN}} L[G(y)]J'[F(x)][F_N(x) - F(X)]\,dH(x, y).$$

Then

$$B_{10,N} = \sum_{i=1}^{8} B_{10,iN}.$$

By Holder's inequality, $|B_{10,1N}|$ is bounded by

$$\left(\int_{b_{2N}}^{\infty} |J'[F(x)][F_N(x) - F(x)]|^{2p/(p+1)}\,dF(x)\right)^{(p+1)/2p}$$
$$\times\left(\int_{d_{2N}}^{\infty} |L[G(y)]|^{2q}\,dG(y)\right)^{1/2q} \tag{37}$$

Let $v = (\alpha + \varepsilon + \frac{1}{2})2p/(p+1) = (\alpha + \varepsilon + \frac{1}{2})/(\alpha + \delta_1 + \frac{1}{2})$. Pick $0 < \varepsilon < \delta_1$ and note that $\frac{1}{2} < v < 1$. Lemma 2.1 and (3) imply that the first factor of (37) is bounded by

$$O(N^{-1/2}\log N)\left(\int_{b_{2N}}^{\infty} [1 - F(x)]^{-v}\,dF(x)\right)^{(p+1)/2p}, \tag{38}$$

which equals

$$O(N^{-1/2}\log N)(N^{-\gamma_1(1-v)(p+1)/2p}) = O(N^{-\eta-1/2}) \qquad \text{for some} \quad \eta > 0.$$

Again by (3), the second factor of (37) is bounded by

$$K\left(\int_{d_{2N}}^{\infty} [1 - G(y)]^{-2\beta q}\,dG(y)\right)^{1/2q},$$

which, on integration, equals $O(N^{-\gamma_2(1-2\beta q)/2q}) = O(N^{-\eta})$ for some $\eta > 0$. Hence $|B_{10,1N}| = O(N^{-\eta-1/2})$. By similar arguments, $|B_{10,iN}| = O(N^{-\eta-1/2})$ for $i = \ldots, 11$. We now consider $B_{12,N}$, which can be decomposed as

$$B_{12,N} = \sum_{i=1}^{8} B_{12,iN}, \tag{39}$$

where

$$B_{12,iN} = \iint\limits_{I_{1N}\cap I_{2,iN}} L[G(y)]J'[F(x)][F_N(x) - F(x)]\,dH_N(x, y). \tag{40}$$

It is easy to check (by Holder's inequality) that

$$|B_{12,1N}| \le \left(\int_{b_{2N}}^{b_{1N}} |J'[F(x)][F_N(x) - F(x)]|^{2p/(p+1)}\right)^{(p+1)/2p}$$
$$\times\left(\int_{d_{2N}}^{d_{1N}} L[G(y)]^{2q}\,dG_N(y)\right)^{1/2q}. \tag{41}$$

Using (3) and Lemma 2.1, the first factor of (41) equals

$$O(N^{-1/2}\log N)\left(\int_{b_{2N}}^{b_{1N}} [1-F(x)]^{-\nu}\,dF_N(x)\right)^{(p+1)/2p}, \tag{42}$$

i.e.,

$$O(N^{-1/2}\log N)\left(-\int_{b_{2N}}^{b_{1N}} [1-F(x)]^{-\nu}\,d[1-F_N(x)]\right)^{(p+1)/2p}. \tag{43}$$

Now integrating by parts and using Lemmas 2.1 and 2.4, we obtain, after routine computations, that (43) equals $O(N^{-\eta})$ for some η. Similarly the second factor on the right-hand side of (41) equals $O(N^{-\eta})$ for sufficiently small η. Thus $|B_{12,1N}| = O(N^{-\eta-1/2})$. The proof that the other terms in (39) equal $O(N^{-\eta-1/2})$ is similar and therefore omitted.

Next we consider $B_{14,N}$. First, observe that the rank of X_i (and Y_i) among $X_1, \ldots, X_N$ (and $Y_1, \ldots, Y_N$) is the same as the rank of $F(X_i)$ [and $F(Y_i)$] among $F(X_1), \ldots, F(x_N)$ [and $F(Y_1), \ldots, F(Y_N)$]. Following Groeneboom *et al.* (1976), we define $\bar{H}(u,v) = H(F^{-1}(u), G^{-1}(v))$ for $(u,v) \in (0,1) \times (0,1)$. Clearly, $\bar{H}(u,v) = P(\{F(X) \le u,\, G(Y) \le v\})$ so that it assigns mass 1 to the unit square and has uniform (0, 1) marginal distribution functions. Without loss of generality, we can then assume that $H(u,v)$ has uniform (0, 1) marginal distribution functions.

For reasons of symmetry, it is enough to consider

$$\left|\int_{1/2}^{1-N^{-\gamma_2}} \int_{1/2}^{1-N^{-\gamma_1}} L(v)J'(u)[F_N(u)-u]\,d[H_N(u,v)-H(u,v)]\right|. \tag{44}$$

We now show that (44) is equal to $O(N^{-\eta-1/2})$ for some $\eta > 0$ by extending a method used in Sen and Ghosh (1974).

Define $\bar{I}_{1,Ni} = [\frac{1}{2} + (i-1)N^{-1/2}, \frac{1}{2} + iN^{-1/2}]$, $i = 1, 2, \ldots, N_1^* - 1$, where N_1^* is the largest positive integer such that $\frac{1}{2} + (N_1^* - 1)N^{-1/2} < 1 - N^{-\gamma_1}$. Define $\bar{I}_{1,NN_1^*} = [\frac{1}{2} + (N_1^* - 1)N^{-1/2}, 1 - N^{-\gamma_1}]$, $u_{Ni} = \frac{1}{2} + iN^{-1/2}$, $0 \le i \le N_1^* - 1$, $u_{NN_1^*} = 1 - N^{-\gamma_1}$. Define $\bar{I}_{2,Nj} = [\frac{1}{2} + (j-1)N^{-\xi}, \frac{1}{2} + jN^{-\xi}]$, $j = 1, 2, \ldots, N_2^* - 1$, where N_2^* is the largest positive integer such that $\frac{1}{2} + (N_2^* - 1)N^{-\xi} < 1 - N^{-\gamma_2}$. Define $\bar{I}_{2,NN_2^*} = [\frac{1}{2} + (N_2^* - 1)N^{-\xi}, 1 - N^{-\gamma_2}]$, $v_{Nj} = \frac{1}{2} + jN^{-\xi}$, $0 \le j \le N_2^* - 1$, $v_{NN_2^*} = 1 - N^{-\gamma_2}$. Observe that $N_1^* = O(N^{1/2})$, $N_2^* = O(N^{\xi})$. Let $\bar{I}_{Nij} = \bar{I}_{1,Ni} \times \bar{I}_{2,Nj}$, $i = 1, \ldots, N_1^*$; $j = 1, \ldots, N_2^*$. If $u \in \bar{I}_{1,Ni}$, then $J'(u) = J'(u_{Ni}) + O(N^{-1/2})O(\{1 - u_{Ni}\}^{-2-\alpha})$ and by Lemma 2.3,

$$F_N(u) - u = F_N(u_{Ni}) + O(N^{-3/4}\log N) \text{ a.s.}$$

If $v \in \bar{I}_{2,Nj}$, then $L(v) = L(v_{Nj}) + O(N^{-\xi})O(\{1 - v_{Nj}\}^{-1-\beta})$. Note that for v, $v_{Nj} \in [\frac{1}{2}, 1 - N^{-\gamma_2}]$, $|L(v)| \le KN^{\gamma_2\beta}$ and $\{1 - v_{Nj}\}^{-1-\beta} \le N^{\gamma_2(1+\beta)}$. It is easy to see that (44) is bounded by

$$
\begin{aligned}
&\sum_{j=1}^{N_2^*}\sum_{i=1}^{N_1^*} |F_N(u_{Ni}) - u_{Ni}|\,|J'(u_{Ni})|\,|L(v_{Nj})| \left|\iint_{I_{Nij}} d[H_N(u,v) - H(u,v)]\right| \\
&\quad + \sum_{i=1}^{N_1^*} |F_N(u_{Ni}) - u_{Ni}|\,|J'(u_{Ni})|\,O(N^{-\xi})KN^{\gamma_2(1+\beta)} \\
&\quad \times \left|\int_{1/2}^{d_{2N}} \int_{\bar{I}_{1,Ni}} d[H_N(u,v) - H(u,v)]\right| \\
&\quad + \sum_{i=1}^{N_1^*} |F_N(u_{Ni}) - u_{Ni}|\,O(N^{-1/2})O(\{1 - u_{Ni}\}^{-2-\alpha})KN^{\gamma_2\beta} \\
&\quad \times \left|\int_{1/2}^{d_{2N}} \int_{\bar{I}_{1,Ni}} d[H_N(u,v) - H(u,v)]\right| \\
&\quad + \sum_{i=1}^{N_1^*} O(N^{-3/4}\log N)|J'(u_{Ni})|KN^{\gamma_2\beta} \int_{1/2}^{d_{2N}} \int_{\bar{I}_{1,Ni}} |d[H_N(u,v) - H(u,v)]| \\
&\quad + \sum_{i=1}^{N_1^*} O(N^{-5/4}\log N)O(\{1 - u_{Ni}\}^{-2-\alpha})KN^{\gamma_2\beta} \\
&\quad \times \int_{1/2}^{d_{2N}} \int_{\bar{I}_{1,Ni}} |d[H_N(u,v) - H(u,v)]|. \qquad (45)
\end{aligned}
$$

Consider the first term of (45). Observe that

$$
\begin{aligned}
&\left|\iint_{I_{Nij}} d[H_N(u,v) - H(u,v)]\right| \\
&\quad \le |H_N(u_{Ni}, v_{Nj}) - H_N(u_{Ni-1'}, v_{Nj}) - H(u_{Ni}, v_{Nj}) + H(u_{Ni-1'}, v_{Nj})| \\
&\qquad + |H_N(u_{Ni-1'} v_{Nj-1}) - H_N(u_{Ni'} v_{Nj-1}) - H(u_{Ni-1'} v_{Nj-1}) + H(u_{Ni'} v_{Nj-1})|. \qquad (46)
\end{aligned}
$$

Lemma 2.3 implies that (46) is bounded by $O(N^{-3/4}\log N)$.

Using (3) and Lemma 2.1, it is easily seen that the first term of (45) is bounded by

$$
O(N^{-5/4}(\log N)^2)\left(\sum_{i=1}^{N_1^*} [1 - u_{Ni}]^{-1/2-\alpha-\varepsilon}\right)\left(\sum_{j=1}^{N_2^*} [1 - v_{Ni}]^{-\beta}\right) \qquad (47)
$$

for any $\varepsilon > 0$. Note that if ε is small enough, then $\frac{1}{2} + \alpha + \varepsilon < 1$. Hence

$$
\sum_{i=1}^{N_1^*} [1 - u_{Ni}]^{-1/2-\alpha-\varepsilon} \le \sum_{i=1}^{N_1^*} (iN^{-1/2})^{-1/2-\alpha-\varepsilon} = O(N^{-1/2}). \qquad (48)
$$

Also note that

$$\sum_{j=1}^{N_2^*} [1 - v_{Nj}]^{-\beta} \le \sum_{j=1}^{N_2^*} (jN^{-\xi})^{-\beta} = O(N^{\xi}).$$

Since $\xi < \frac{1}{4}$, we see that (47) is equal to $O(N^{-\eta-1/2})$ for some $\eta > 0$. It can similarly be shown that the other terms in (45) also equal $O(N^{-\eta-1/2})$. Hence $|B_{14,N}| = O(N^{-\eta-1/2})$ for sufficiently small η.

Next consider $B_{16,N}$. For $y \in [c_{1N}, d_{1N}]$, $L[G(y)] \le KN^{\beta(1-\delta_2)}$,

$$\begin{aligned} |B_{16,N}| \le (N+1)^{-1} KN^{\beta(1-\delta_2)} \int_a^{b_{1N}} [1 - F(x)]^{-1-\alpha}\, dF_N(x) \\ + (N+1)^{-1} KN^{\beta(1-\delta_2)} \int_{a_{1N}}^{a} [F(x)]^{-1-\alpha}\, dF_N(x). \end{aligned} \tag{49}$$

Lemma 2.2 implies that the first term of (49) is bounded by

$$O(N^{-1+\beta(1-\delta_2)}) \int_a^{b_{1N}} [1 - F_N(x)]^{-1-\alpha}\, dF_N(x). \tag{50}$$

Also the integral inside (50) equals $O(N^{\alpha(1-\delta_1)})$. Thus Eq. (50) equals

$$O(N^{-1/2-\delta_1-\delta_2-\beta\delta_2-\alpha\delta_1}) = O(N^{-\eta-1/2}) \qquad \text{for} \quad \eta < \delta_1 + \delta_2 + \beta\delta_2 + \alpha\delta_1.$$

Similarly, the second term of (49) equals $O(N^{-\eta-1/2})$ for sufficiently small η. Hence $|B_{16,N}| = O(N^{-\eta-1/2})$. The proof of the theorem is now complete.

Now set

$$\mu = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J[F(x)]L[G(y)]\, dH(x, y) \tag{51}$$

and

$$\begin{aligned} \sigma^2 = \operatorname{var}\Big(& J[F(X)]L[G(Y)] \\ & + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\phi_X(x) - F(x)]J'[F(x)]L[G(y)]\, dH(x, y) \\ & + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\phi_Y(y) - G(y)]J[F(x)]L'[G(y)]\, dH(x, y)\Big), \end{aligned}$$

where $\phi_X(x) = 1$ ($\phi_Y(y) = 1$) if $X \le x$ ($Y \le y$) and is zero otherwise. Then we have the following result.

THEOREM 3.2. *Under the assumption* 2.1, $N^{1/2}(T_N - \mu) \xrightarrow{d} N(0, \sigma^2)$ *as* $N \to \infty$, *uniformly with respect to* $H(x, y)$.

Proof. From Ruymgaart *et al.* (1972), $N^{1/2}(\sum_{i=1}^{N} A_{iN} - \mu) \xrightarrow{d} N(0, \sigma^2)$ uniformly with respect to $H(x, y)$ as $N \to \infty$. Since (by Theorem 2.1), $N^{1/2}R_N \to 0$ a.s., the proof follows.

4. Invariance Principles

As a consequence of Theorem 3.1, we have the following law of the iterated logarithm.

THEOREM 4.1. *Under Assumption* 2.1,

$$\limsup_{N\to\infty} N^{1/2}(T_N - \mu)/(2\log\log N)^{1/2} = \sigma \text{ a.s.},$$
$$\liminf_{N\to\infty} N^{1/2}(T_N - \mu)/(2\log\log N)^{1/2} = -\sigma \text{ a.s.}$$

Proof: It can be shown that $\sum_{i=1}^{3} A_{i,N}$ is the average of N independent and identically distributed random variables, each having mean μ and variance σ^2. Hence

$$\limsup_{N\to\infty} N^{1/2}\left[\left(\sum_{i=1}^{3} A_{i,N}\right) - \mu\right]\Big/(2\log\log N)^{1/2} = \sigma \text{ a.s.},$$
$$\liminf_{N\to\infty} N^{1/2}\left[\left(\sum_{i=1}^{3} A_{i,N}\right) - \mu\right]\Big/(2\log\log N)^{1/2} = -\sigma \text{ a.s.}$$

Also by Theorem 3.1,

$$\limsup_{N\to\infty} N^{1/2} R_N = 0 \text{ a.s.}$$

The proof follows.

It is also easy to establish the following.

THEOREM 4.2. *Let* $W_N = N^{1/2}(T_N - \mu)$. *If Assumption 2.1 holds, then* $N^{-1/2} W_{[Nt]}/\sigma$, $0 \le t \le 1$, *converges weakly to the standard Wiener process.*

Both Theorems 4.1 and 4.2 have useful applications to sequential tests for independence. Along the lines of Sen and Ghosh (1973), a class of sequential tests having power 1 and arbitrary small Type 1 error can be constructed for testing $H(x, y) = F(x)G(y)$. The invariance principles obtained here are useful for the study of the asymptotic properties of these tests, especially when the null hypothesis does not hold.

References

Bahadur, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37**, 577–580.

Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35**, 138–149.

Chernoff, H., and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29**, 972–944.

Groeneboom, P., Lepage, Y., and Ruymgaart, F. H. (1976). Rank tests for independence with best strong exact Bahadur slope. *Z. Wahrsch. Verw. Gebiete* **36**, 119–127.

Lai, T. L. (1975). On Chernoff-Savage statistics and sequential rank tests. *Ann. Statist.* **3**, 825–845.

Ruymgaart, F. H., Shorack, G. R., and van Zwet, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist.* **43**, 1122–1135.

Sen, P. K. (1972). On the Bahadur representation of sample quantities for sequences of ϕ-mixing random variables. *J. Multivariate Anal.* **2**, 77–95.

Sen, P. K., 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. K., and Ghosh, M. (1973). Some invariance principles for rank statistics for testing independence. *Z. Wahrsch. Verw. Gebiete* **29**, 93–107.

Sen, P. K., and Ghosh, M. (1974). A Chernoff-Savage representation of rank order statistics for stationary ϕ-mixing processes. *Sankhyà, Ser. A* **35**, 153–172.

Department of Mathematics
Indiana University
Bloomington, Indiana

AMS 1970 Subject Classification:
Primary 62E20, 60B10;
Secondary 62G30.

Reprinted from
Contributions to Probability: A Collection of Papers
Dedicated to Eugene Lukacs (1981), pages 267-282,
edited by J. Gani and V.K. Rohatgi,
by permission of the publisher,

ON THE DEGENERATION OF THE VARIANCE IN THE ASYMPTOTIC NORMALITY OF SIGNED RANK STATISTICS

MADAN L. PURI and STEFAN S. RALESCU
Indiana University, Bloomington, IN, and Brown University, Providence, RI, U.S.A.

The purpose of this paper is to establish the asymptotic normality of simple linear signed rank statistics S_N^+ considered by Hušková (1970), Koul and Staudte (1972), and Puri and Ralescu (1980) for the case when the score-generating function is discontinuous and $\mathrm{Var}(S_N^+)$ compared with the variance of S_N^+ under the hypothesis of symmetry is allowed some degree of degeneracy.

The results obtained are extensions of those by Hájek (1968), Dupač and Hájek (1969), Dupač (1970), Koul and Staudte (1972) and Puri and Ralescu (1980).

1. Preliminaries

Let $X_{N1},\ldots, X_{NN}$, $N\geqslant 1$ be independent random variables, with continuous distribution function $F_{N1},\ldots, F_{NN}$ respectively, and let R_{Ni}^+ be the rank of $|X_{Ni}|$ among $|X_{N1}|,\ldots,|X_{NN}|$. Consider the statistic

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

where $c_{N1},\ldots,c_{NN}$ are known regression constants, $a_N(1),\ldots,a_N(N)$ are scores and $\operatorname{sgn} x=1$ if $x\geqslant 0$, $\operatorname{sgn} x=-1$ if $x<0$.

For simplicity of notation, we shall drop the subscript N in X_{Ni}, c_{Ni} and R_{Ni}^+ in the sequel.

In order to study the asymptotic behavior of S_N^+, the ratio $\mathrm{Var}(S_N^+)/\sum_{i=1}^N c_i^2$ plays an important role [see Hájek (1968), Dupač and Hájek (1969), Dupač (1970) and Koul and Staudte (1972)]. For the case of the unit step score-generating function $\psi(t)=1$ for $t\geqslant v$, $\psi(t)=0$ for $t<v$ $(0<v<1)$, under suitable conditions on the distribution functions and regression constants, we shall prove that if the ratio $\mathrm{Var}(S_N^+)/\sum_{i=1}^N c_i^2$ goes to zero at most at the rate $N^{-\alpha}$ for some $0<\alpha<\frac{1}{2}$, then S_N^+ is asymptotically normal with natural parameters $(E(S_N^+),\mathrm{Var}(S_N^+))$ as well as with some other simpler parameters (μ_N^+,σ_N^2).

We assume that the c_i's satisfy the condition

$$\max_{1\leqslant i\leqslant N} c_i^2/\sum_{i=1}^{N} c_i^2 = \mathrm{O}(N^{-1/2}). \tag{1.2}$$

Let $F_i^*(x)$ be the distribution function of $|X_i|$ and define

$$H_N^*(x)=\frac{1}{N}\sum_{i=1}^{N} F_i^*(x),$$

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

$$L_i(t)=F_i\big(H^{*-1}(t)\big),\quad 0<t<1,$$

$$M_i(t)=-F_i\big(-H^{*-1}(t)\big),\quad 0<t<1,$$

$$G_i(t)=F_i^*\big(H^{*-1}(t)\big)=L_i(t)+M_i(t),\quad 0<t<1. \tag{1.3}$$

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

$$a_N(i)=\psi(i/(N+1)),\quad 1\leq i\leq N \tag{1.4}$$

or by a procedure satisfying

$$\sum_{i=1}^{N} |a_N(i)-\psi(i/(N+1))|=O(1). \tag{1.5}$$

If $v\in(0,1)$ represents a jump point of the score-generating function ψ, then for every $K>0$ we assume the existence of the derivatives $L_i'(t)$ and $M_i'(t)$ in the interval $|t-v|\leq KN^{-1/2}Lg^{1/2}N$ and the satisfaction of the following conditions.

$$\max_{1\leq i\leq N}|L_i'(t)|=O(1), \tag{1.6}$$

$$\max_{1\leq i\leq N}|M_i'(t)|=O(1), \tag{1.7}$$

$$\max_{1\leq i\leq N}\ \operatorname*{Sup}_{|t-v|\leq KN^{-1/2}Lg^{1/2}N}|L_i'(t)-L_i'(v)|=O\big(N^{-1/2}Lg^{1/2}N\big), \tag{1.8}$$

$$\max_{1\leq i\leq N}\ \operatorname*{Sup}_{|t-v|\leq KN^{-1/2}Lg^{1/2}N}|M_i'(t)-M_i'(v)|=O\big(N^{-1/2}Lg^{1/2}N\big). \tag{1.9}$$

Another condition concerning the G_i's that we use is:

$$\liminf_{N\to\infty} N^{-1}\sum_{i=1}^{N} G_i(v)(1-G_i(v))>0. \tag{1.10}$$

Sometimes, mainly for purposes of applications, we replace (1.6)–(1.10) by the following condition which is easier to verify:

Suppose that each F_i has a density f_i. For each $\varepsilon>0$ denote $I_\varepsilon=(H^{*-1}(v)-\varepsilon, H^{*-1}(v)+\varepsilon)$.

Suppose that there exist $\varepsilon_1,\varepsilon_2,\varepsilon_3>0$ such that:

(a)

$$\liminf_{N\to\infty} H_N^{*-1}(v)>0, \tag{1.11}$$

(b) $f_i'(x)$ are uniformly bonded (in x, i, N) on $I_{\varepsilon_1} \cup (-I_{\varepsilon_1})$,

(c) for all $N \geqslant 1$,

$$\frac{1}{N} \operatorname{Card}\left\{ 1 \leqslant i \leqslant N: \inf_{x \in I_{\varepsilon_1}} f_i^*(x) > \varepsilon_2 \right\} > \varepsilon_3$$

where f_i^* is the density of F_i^*.

The last condition that we require concerns some possible degeneration of $\operatorname{Var}(S_N^+)$ in the form

$$\liminf_{N \to \infty} \operatorname{Var}(S_N^+) \Big/ \left(N^{-\alpha} \sum_{i=1}^{N} c_i^2 \right) > 0 \tag{1.12}$$

for some $0 < \alpha < \frac{1}{2}$.

Alternatively we shall assume that (1.12) holds with $\operatorname{Var}(S_N^+)$ replaced by some approximate variance σ_N^2:

$$\liminf_{N \to \infty} \sigma_N^2 \Big/ \left(N^{-\alpha} \sum_{i=1}^{N} c_i^2 \right) > 0 \quad (0 < \alpha < \tfrac{1}{2}). \tag{1.13}$$

2. Main theorems

Let $u(t)$ be 1 or 0 according to $t \geqslant 0$ or $t < 0$.

The main result of this paper is the following theorem:

Theorem 2.1. *Let S_N^+ given by* (1.1) *have scores given by* (1.5) *where $\psi(t) = u(t - v)$, $0 < v < 1$.*

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

$$(\tilde{\mathrm{C}}_1^+): (1.2), (1.6), (1.7), (1.8), (1.9), (1.10), (1.12),$$

$$(\tilde{\mathrm{C}}_2^+): (1.2), (1.11), (1.12).$$

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.

Let us begin by assuming that scores are given by (1.4) and that $(\tilde{\mathrm{C}}_1^+)$ holds.

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

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

This will be accomplished by using the Residual variance inequality [see Hájek (1968) and Koul and Staudte (1972)]:

$$E\left(S_N^+ - \hat{S}_N^+\right)^2 \leqslant \sum_{i=1}^{N} c_i^2 E\left(a(R_i^+) - E\left(a(R_i^+)|X_i\right)\right)^2$$

$$+ \sum_{i\neq j} c_i c_j \Big\{ E\left(\operatorname{sgn} X_i \operatorname{sgn} X_j \operatorname{Cov}\left(a(R_i^+), a(R_j^+)|X_i, X_j\right)\right)$$

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

$$\left.\times\left[E\left(a(R_j^+)|X_i, X_j\right) - E\left(a(R_j^+)|X_j\right)\right]\right\}$$

$$- \sum_{k\neq i,j} \operatorname{Cov}\left\{E\left(\operatorname{sgn} X_i a(R_i^+)|X_k\right), E\left(\operatorname{sgn} X_j a(R_j^+)|X_k\right)\right\}\Big\}.$$

We investigate each term in the above inequality. The proof is divided in several steps:

Lemma 2.1. *Let $x, y \in \mathbb{R}$. Then for each $K_1 > \sqrt{6}$ there exists a $K_2 \geqslant \frac{3}{2}$ such that for all $N > N_0(K_1)$ we have*

(i) $v - H^*(|x|) > K_1 N^{-1/2} Lg^{1/2} N \Rightarrow P(R_i^+ \geqslant V | X_i = x, X_j = y) < N^{-K_2}$,

(ii) $v - H^*(|x|) < -K_1 N^{-1/2} Lg^{1/2} N \Rightarrow P(R_i^+ \leqslant V | X_i = x, X_j = y) < N^{-K_2}$ *where* $V = [(N+1)v]$. ($[\cdot] =$ *integer part*).

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

Let

$$D^2 = N^{-1} \sum_{i=1}^{N} G_i(v)(1 - G_i(v)).$$

Lemma 2.2. *Suppose that $|v - H^*(|x|)| \leqslant K_3 N^{-1/2} Lg^{1/2} N$. Then for sufficiently large N, we have*

(i) $$\left|\sum_{i=1}^{N} F_i^*(|x|)(1 - F_i^*(|x|)) - ND^2\right| \leqslant K_4 N^{1/2} Lg^{1/2} N,$$

(ii) $$\left|\phi\left(V; \sum_{i=1}^{N} F_i^*(|x|), \sum_{i=1}^{N} F_i^*(|x|)(1 - F_i^*(|x|))\right) - \phi\left(Nv; \sum_{i=1}^{N} F_i^*(|x|), ND^2\right)\right| \leqslant K_5 N^{-1} Lg^{1/2} N,$$

(iii)
$$\left|\Phi\left(V;\ \sum_{i=1}^{N} F_i^*(|x|),\ \sum_{i=1}^{N} F_i^*(|x|)(1-F_i^*(|x|))\right)\right.$$
$$\left.-\Phi\left(Nv;\ \sum_{i=1}^{N} F_i^*(|x|),\ ND^2\right)\right| \leqslant K_6 N^{-1/2} Lg^{1/2} N$$

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

The proofs of Lemmas 2.1 and 2.2 are analogous to those of Lemmas 5 and 6 of Dupač and Hájek (1969) and are therefore omitted.

Lemma 2.3. *For $N\to\infty$, we have*

$$E\big(a(R_i^+)-E(a(R_i^+)|X_i)\big)^2=o(N^{-\alpha})$$

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

Proof. Let $\nabla^+(X_i)=E(a(R_i^+)|X_i)-[E(a(R_i^+)|X_i)]^2$. Then, by conditioning, we obtain

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

Now, by definition

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

Thus

$$\nabla^+(X_i=x)=P(R_i^+>V|X_i=x)\cdot P(R_i^+\leqslant V|X_i=x).$$

Let $I=\{x\colon |H^*(|x|)-v|\leqslant K_1 N^{-1/2}Lg^{1/2}N\}$ with $K_1>\sqrt{6}$. By Lemma 2.1, if $x\notin I$ we have:

$$\nabla^+(X_i=x)<N^{-K_2}\quad \text{for every } N>N_0(K_1),\ K_2\geqslant\tfrac{3}{2}.$$

On the other hand, if $x\in I$, then since $P(R_i^+=k\,|\,X_i=x)=B^i(k, F_1^*(|x|),\ldots,F_N^*(|x|))$ (in the notation used by Dupač and Hájek (1969)) we obtain, using Lemmas 2.1 and 2.2, that

$$\nabla^+(X_i=x)=\left\{\sum_{k>V} B^i(k, F_1^*(|x|),\ldots,F_N^*(|x|))\right\}$$
$$\times\left\{\sum_{l\leqslant V} B^i(l, F_1^*(|x|),\ldots,F_N^*(|x|))\right\}$$
$$=\cdots=\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\right\}+\theta_1 N^{-1/2}Lg^{1/2}N$$

for sufficiently large N, $|\theta_1|\leqslant K_7$. Here Φ denotes the standard normal distribution function. We use ϕ for its density function in the sequel.

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

$$I'=\{x\colon |H^*(|x|)-v|\leqslant K_9 DN^{-1/2}Lg^{1/2}N\},$$

where K_9 is such that $K_9={}^K1/2K_8$ with $K_8\leqslant D\leqslant\frac{1}{2}$.

Now, using (1.6)–(1.9) it is easy to show that

$$N^{-1/2}Lg^{1/2}N\int_{I'}\theta_1\,\mathrm{d}F_i(x)=\mathrm{o}(N^{-\alpha})$$

and

$$\int_{I'}\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)=\mathrm{o}(N^{-\alpha})$$

uniformly in $1\leqslant i\leqslant N$. Hence

$$E(\nabla^+(x_i))=\mathrm{o}(N^{-\alpha})\text{ uniformly in }1\leqslant i\leqslant N$$

and the proof follows.

Lemma 2.4. *For $N\to\infty$ we have*

$$E\{\operatorname{sgn}X_i\operatorname{sgn}X_j[E(a(R_i^+)|X_i,X_j)-E(a(R_i^+)|X_i)]$$
$$\times[E(a(R_j^+)|X_i,X_j)-E(a(R_j^+)|X_j)]\}=\mathrm{o}(N^{-1-\alpha}),$$

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

The proof of this lemma is similar to that of Lemma 2.3 and is therefore omitted.

Lemma 2.5. *For $N\to\infty$ we have*

$$E[\operatorname{sgn}X_i\operatorname{sgn}X_j\operatorname{Cov}(a(R_i^+),a(R_j^+)|X_i,X_j)]$$
$$=N^{-1}D^2(L_i'(v)-M_i'(v))(L_j'(v)-M_j'(v))+\mathrm{o}(N^{-1-\alpha})$$

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

Proof. We have

$$\Delta^+=\operatorname{Cov}(a(R_i^+),a(R_j^+)|X_i=x,X_j=y)$$
$$=\begin{cases}P(R_i^+>V|X_i=x,X_j=y)P(R_j^+\leqslant V|X_i=x,X_j=y),\\ \quad\text{if }|x|<|y|,\\ P(R_j^+>V|X_i=x,X_j=y)P(R_i^+\leqslant V|X_i=x,X_j=y),\\ \quad\text{if }|x|\geqslant|y|.\end{cases}$$

Let $K_1>\sqrt{6}$. Denote

$$I=\{(x,y)\colon\max(|H^*(|x|)-v|,|H^*(|y|)-v|)\leqslant K_1N^{-1/2}Lg^{1/2}N\}.$$

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

$$\Delta^+(x,y)\begin{cases} <N^{-K_2} \quad \text{for } (x,y)\notin I,\ N>N_0(K_2) \\ =\Phi\left(\dfrac{H^*(|x|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\dfrac{H^*(|y|)-v}{DN^{-1/2}}\right)\right\}+\theta_2 N^{-1/2}Lg^{1/2}N \\ \quad \text{for } N \text{ sufficiently large, } (x,y)\in I,\ |x|<|y| \\ \quad \text{and } |\theta_2|\leqslant K_{10} \\ =\Phi\left(\dfrac{H^*(|y|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\dfrac{H^*(|x|)-v}{DN^{-1/2}}\right)\right\}+\theta_3 N^{-1/2}Lg^{1/2}N \\ \quad \text{for } N \text{ sufficiently large, } (x,y)\in I,\ |x|\geqslant|y| \\ \quad \text{and } |\theta_3|\leqslant K_{11} \end{cases} \tag{2.1}$$

where $K_2 \geqslant \frac{3}{2}$. We note that the *equality* in (2.1) remains true even if we enlarge I to I' where

$$I'=\left\{(x,y):\ \max(|H^*(|x|)-v|,|H^*(|y|)-v|)\leqslant K_9 DN^{-1/2}Lg^{1/2}N\right\}$$

where $K_9={}^K1/2K_8$, $K_8\leqslant D\leqslant\frac{1}{2}$.

We have, using (2.1) that

$$E\left(\operatorname{sgn} X_i \operatorname{sgn} X_j \operatorname{Cov}\left(a(R_i^+),a(R_j^+)\right)|X_i,X_j\right)=$$

$$=\iint_{I'\cap\{|x|<|y|\}} \operatorname{sgn} x \operatorname{sgn} y\,\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(|y|)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y) \tag{2.2}$$

$$+\iint_{I'\cap\{|x|\geqslant|y|\}} \operatorname{sgn} x \operatorname{sgn} y\,\Phi\left(\frac{H^*(|y|)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(|x|)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)$$

$$+N^{-1/2}Lg^{1/2}N\iint_{I'} \operatorname{sgn} x \operatorname{sgn} y\,\theta_4(x,y)\,\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)+\theta_5 N^{-K_2}$$

with $|\theta_4|\leqslant K_{12}$, $|\theta_5|\leqslant 1$.

The last two terms are $o(N^{-1-\alpha})$ uniformly in i, j as follows by using (1.6)–(1.9) and $K_2\geqslant\frac{3}{2}$. It remains to estimate the first two terms.

Denote the first term by $\mathfrak{T}$. Consider

$$\mathfrak{T}_1=\iint_{A_{xy}} \Phi\left(\frac{H^*(x)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(y)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)$$

where

$$A_{xy}=\left\{(x,y):\ \begin{matrix} x>0 \\ y>0 \\ x<y \end{matrix},\ \max(|H^*(x)-v|,|H^*(y)-v|)\leqslant K_9 DN^{-1/2}Lg^{1/2}N\right\}.$$

Set

$$p=\frac{H^*(x)-v}{DN^{-1/2}}, \qquad q=\frac{H^*(y)-v}{DN^{-1/2}},$$

and

$$I''=\{(p,q)\colon \max(|p|,|q|)\leqslant K_9 Lg^{1/2}N\}.$$

Then

$$\mathfrak{T}_1=\iint_{I''\cap\{p<q\}}\Phi(p)(1-\Phi(q))\,\mathrm{d}L_i(v+DN^{-1/2}p)\,\mathrm{d}L_j(v+DN^{-1/2}q).$$

Let $\Omega=I''\cap\{p<q\}$ and $\Omega^*=\{p<q\}\setminus\Omega$. Then using (1.6)–(1.9), one can easily show that

$$\mathfrak{T}_1=D^2N^{-1}L_i'(v)L_j'(v)\iint_\Omega\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q+\mathrm{o}(N^{-1-\alpha}) \tag{2.3}$$

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

Let $\Omega_1^*=\Omega^*\cap\{p>-q\}$. Then, by Fubini's theorem we have:

$$N^{-1}\iint_{\Omega_1^*}\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q=N^{-1}\int_{K_9Lg^{1/2}N}^{\infty}(1-\Phi(q))\left(\int_{-q}^{q}\Phi(p)\,\mathrm{d}p\right)\mathrm{d}q$$
$$\leqslant 2N^{-1}\int_{K_9Lg^{1/2}N}^{\infty}q\Phi(-q)\,\mathrm{d}q.$$

Using integration by parts it follows that:

$$N^{-1}\iint_{\Omega_1^*}\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q=\mathrm{o}(N^{-1-\alpha}). \tag{2.4}$$

Similarly we can show that:

$$N^{-1}\iint_{\Omega_2^*}\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q=\mathrm{o}(N^{-1-\alpha}) \tag{2.5}$$

where $\Omega_2^*=\Omega^*\setminus\Omega_1^*$.

By (2.4) and (2.5)

$$D^2N^{-1}L_i'(v)L_j'(v)\iint_{\Omega^*}\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q=\mathrm{o}(N^{-1-\alpha}). \tag{2.6}$$

From (2.3), (2.6) and the fact that

$$\iint_{\{p<q\}}\Phi(p)(1-\Phi(q))\,\mathrm{d}p\,\mathrm{d}q=\tfrac{1}{2}$$

we obtain:

$$\mathfrak{T}_1=\tfrac{1}{2}D^2N^{-1}L_i'(v)L_j'(v)+\mathrm{o}(N^{-1-\alpha}) \tag{2.7}$$

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

Let

$$\mathfrak{T}_2=\iint_{B_{xy}}-\Phi\left(\frac{H^*(-x)-v}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(y)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)$$

$$=\iint_{I''\cap\{p<q\}}-\Phi(p)(1-\Phi(q))\,\mathrm{d}M_i(v+DN^{-1/2}p)\,\mathrm{d}L_j(v+DN^{-1/2}q)$$

where

$$B_{xy}=\left\{(x,y):\begin{array}{l}x<0\\ y>0\\ -x<y\end{array},\max(|H^*(-x)-v|,|H^*(y)-v|)\leqslant K_9DN^{-1/2}Lg^{1/2}N\right\},$$

$$\mathfrak{T}_3=\iint_{C_{xy}}\Phi\left(\frac{H^*(-x)-y}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(-y)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)$$

and

$$\mathfrak{T}_4=\iint_{D_{xy}}-\Phi\left(\frac{H^*(x)-y}{DN^{-1/2}}\right)\left\{1-\Phi\left(\frac{H^*(-y)-v}{DN^{-1/2}}\right)\right\}\mathrm{d}F_i(x)\,\mathrm{d}F_j(y)$$

where

$$C_{xy}=\left\{(x,y):\begin{array}{l}x<0\\ y<0\\ -x<-y\end{array},\max(|H^*(-x)-v|,|H^*(-y)-v|)\leqslant K_9DN^{-1/2}Lg^{1/2}N\right\}$$

and

$$D_{xy}=\left\{(x,y):\begin{array}{l}x>0\\ y<0\\ x<-y\end{array},\max(|H^*(x)-v|,|H^*(-y)-v|)\leqslant K_9DN^{-1/2}Lg^{1/2}N\right\}.$$

We now repeat the steps used in the derivation of (2.7), this time applying them to $\mathfrak{T}_2$, $\mathfrak{T}_3$ and $\mathfrak{T}_4$ to obtain

$$\mathfrak{T}_2=-\tfrac{1}{2}D^2N^{-1}M_i'(v)L_j'(v)+o(N^{-1-\alpha})\quad\text{uniformly in } 1\leqslant i,\ j\leqslant N,$$

$$\mathfrak{T}_3=\tfrac{1}{2}D^2N^{-1}M_i'(v)M_j'(v)+o(N^{-1-\alpha})\quad\text{uniformly in } 1\leqslant i,\ j\leqslant N$$

and

$$\mathfrak{T}_4=-\tfrac{1}{2}D^2N^{-1}L_i'(v)M_j'(v)+o(N^{-1-\alpha})\quad\text{uniformly in } 1\leqslant i,\ j\leqslant N.$$

Thus

$$\mathfrak{T}=\mathfrak{T}_1+\mathfrak{T}_2+\mathfrak{T}_3+\mathfrak{T}_4=\tfrac{1}{2}D^2N^{-1}(L_i'(v)-M_i'(v))(L_j'(v)-M_j'(v))+o(N^{-1-\alpha})$$

uniformly in $1\leqslant i,\ j\leqslant N$. (2.8)

Proceeding as above, it can be shown that the second term of (2.2) is the same as (2.8). The proof follows.

Lemma 2.6. *For $N \to \infty$ we have (for $i \neq j$)*

$$\sum_{k \neq i, j} \mathrm{Cov}\{E(\mathrm{sgn}\, X_i a(R_i^+)|X_k), E(\mathrm{sgn}\, X_j a(R_j^+)|X_k)\} =$$
$$= D^2 N^{-1}(L_i'(v) - M_i'(v))(L_j'(v) - M_j'(v)) + \mathrm{o}(N^{-1-\alpha})$$

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

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

$$E(a(R_i^+)\,\mathrm{sgn}\, X_i|X_i = x, X_k = z) - E(a(R_i^+)\,\mathrm{sgn}\, X_i|X_i = x) =$$
$$= \mathrm{sgn}\, x[u(|x| - |z|) - F_k^*(|x|)] \cdot P(R_i^+ = V + 1|X_i = x, |X_k| = |x| - 1).$$

From Lemma 2.1, we have

$$P(R_i^+ = V + 1|X_i = x, |X_k| = |x| - 1) < N^{-K_2} \tag{2.9}$$

for some $K_2 \geqslant \frac{3}{2}$ and all $|H^*(|x|) - v| \geqslant K_1 N^{-1/2} Lg^{1/2} N$.

Furthermore Lemmas 2.1 and 2.2 imply:

$$P(R_i^+ = V + 1|X_i = x, |X_k| = |x| - 1) = \phi\left(Nv;\ \sum_{j=1}^{N} F_j^*(|x|), ND^2\right) + \theta_6 N^{-1} Lg^{1/2} N$$

for some $|\theta_6| \leqslant K_{13}$ and all $|H^*(|x|) - v| \leqslant K_1 N^{-1/2} Lg^{1/2} N$.

As before, the last equality remains true even if $|H^*(|x|) - v| \leqslant K_9 D N^{-1/2} Lg^{1/2} N$. Let

$$I' = \{x\colon |H^*(|x|) - v| \leqslant K_9 D N^{-1/2} Lg^{1/2} N\}.$$

Then

$$E(a(R_i^+)\,\mathrm{sgn}\, X_i|X_k = z) - E(a(R_i^+)\,\mathrm{sgn}\, X_i) =$$
$$= \int \mathrm{sgn}\, x[u(|x| - |z|) - F_k^*(|x|)]$$
$$\times P(R_i^+ = V + 1|X_i = x, |X_k| = |x| - 1)\, \mathrm{d}F_i(x)$$
$$= \int_{I'} (\cdots)\, \mathrm{d}F_i(x) + \int_{\mathbf{R} \setminus I'} (\cdots)\, \mathrm{d}F_i(x).$$

The second integral is $\mathrm{o}(N^{-1-\alpha})$ by (2.9), while the first is equal to

$$\int_{I'} \mathrm{sgn}\, x[u(|x| - |z|) - F_k^*(|x|)]\, \phi\left(Nv;\ \sum_{j=1}^{N} F_j^*(|x|), ND^2\right) \mathrm{d}F_i(x)$$
$$+ \int_{I'} \mathrm{sgn}\, x[u(|x| - |z|) - F_k^*(|x|)]\, \theta_6 N^{-1} Lg^{1/2} N\, \mathrm{d}F_i(x).$$

In the last expression, let us denote by $\mathfrak{I}_5$ the first term and $\mathfrak{I}_6$ the second.

From (1.6)–(1.9) and the Mean Value Theorem, it follows that:

$$\max_{1 \leqslant k \leqslant N}\ \sup_{|p| \leqslant K_9 Lg^{1/2} N} |G_k(v + DN^{-1/2} p) - G_k(v)| = \mathrm{O}(N^{-1/2} Lg^{1/2} N). \tag{2.10}$$

Then it is easy to show that

$$D^{-1}N^{-1/2}\int_{\{|p|\leqslant K_9 Lg^{1/2}N\}} G_k(v+DN^{-1/2}p)\phi(p)\,\mathrm{d}L_i(v+DN^{-1/2}p)=$$

$$=D^{-1}N^{-1/2}\int_{\{|p|\leqslant K_9 Lg^{1/2}N\}} G_k(v)\phi(p)\,\mathrm{d}L_i(v+DN^{-1/2}p)+\mathrm{o}(N^{-1-\alpha})$$

uniformly in i and k.

We write

$$\mathfrak{T}_5=D^{-1}N^{-1/2}\int_{\{x>0:\,|H^*(x)-v|\leqslant K_9 DN^{-1/2}Lg^{1/2}N\}}(\cdots)\,\mathrm{d}F_i(x)$$

$$+D^{-1}N^{-1/2}\int_{\{x<0:\,|H^*(-x)-v|\leqslant K_9 DN^{-1/2}Lg^{1/2}N\}}(\cdots)\,\mathrm{d}F_i(x)$$

$$=\mathfrak{T}_5'+\mathfrak{T}_5''.$$

We have:

$$\mathfrak{T}_5'=N^{-1}L_i'(v)\int_{\{|p|\leqslant K_9 Lg^{1/2}N\}}[u(p-q)-G_k(v)]\phi(p)\,\mathrm{d}p$$

$$+N^{-1}\mathrm{O}(N^{-1/2}Lg^{1/2}N)\int_{\{|p|\leqslant K_9 Lg^{1/2}N\}}[u(p-q)-G_k(v)]\phi(p)\,\mathrm{d}p$$

$$+\mathrm{o}(N^{-1-\alpha}).$$

In the last expression, the second term is $\mathrm{o}(N^{-1-\alpha})$ while the first is equal to

$$N^{-1}L_i'(v)[1-\Phi(q)-G_k(v)]$$

$$-N^{-1}L_i'(v)\int_{\{|p|>K_9 Lg^{1/2}N\}}[u(p-q)-G_k(v)]\phi(p)\,\mathrm{d}p.$$

But

$$\left|N^{\alpha}\int_{K_9 Lg^{1/2}N}^{\infty}[u(p-q)-G_k(v)]\phi(p)\,\mathrm{d}p\right|\leqslant$$

$$\leqslant N^{\alpha}\int_{K_9 Lg^{1/2}N}^{\infty}\phi(p)\,\mathrm{d}p=N^{\alpha}\Phi(-K_9 Lg^{1/2}N)\to 0$$

and we obtain:

$$\mathfrak{T}_5'=N^{-1}L_i'(v)[1-\Phi(q)-G_k(v)]+\mathrm{o}(N^{-1-\alpha})$$

uniformly in z, $1\leqslant i\leqslant N$, where $q=(H^*(|z|)-v)/DN^{-1/2}$.

Similarly

$$\mathfrak{T}_5''=-N^{-1}M_i'(v)[1-\Phi(q)-G_k(v)]+\mathrm{o}(N^{-1-\alpha}).$$

Also, it is easy to check that $\mathfrak{T}_6 = o(N^{-1-\alpha})$. Hence

$$E(a(R_i^+)\operatorname{sgn} X_i | X_k = z) - E(a(R_i^+)\operatorname{sgn} X_i)$$
$$= \mathfrak{T}_5' + \mathfrak{T}_5'' + o(N^{-1-\alpha})$$
$$= N^{-1}(L_i'(v) - M_i'(v))[1 - \Phi(q) - G_k(v)] + o(N^{-1-\alpha}) \tag{2.11}$$

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

We now show that under (1.6)–(1.9)

$$\int_{-v/DN^{-1/2}}^{(1-v)/DN^{-1/2}} (1-\Phi(q))\,\mathrm{d}G_k(v + DN^{-1/2}q) = G_k(v) + o(N^{-\alpha}). \tag{2.12}$$

Indeed, using integration by parts

$$\int_{-v/DN^{-1/2}}^{(1-v)/DN^{-1/2}} (1-\Phi(q))\,\mathrm{d}G_k(v + DN^{-1/2}q) - G_k(v) =$$
$$= \left[1 - \Phi\left(\frac{1-v}{DN^{-1/2}}\right)\right]$$
$$+ \int_{-\infty}^{\infty} \left[I_{(-(v/DN^{-1/2}),((1-v)/DN^{-1/2}))}(q) G_k(v + DN^{-1/2}q) - G_k(v)\right] \phi(q)\,\mathrm{d}q.$$

Let $\mathcal{Q}$ denote the last integral in the above relation. Then:

$$\mathcal{Q} = \int_{\{|q| \geq Lg^{1/2}N\}} (\cdots)\phi(q)\,\mathrm{d}q + \int_{\{|q| \leq Lg^{1/2}N\}} (\cdots)\phi(q)\,\mathrm{d}q = \mathcal{Q}_1 + \mathcal{Q}_2.$$

Since

$$N^\alpha \int_{Lg^{1/2}N}^{\infty} \phi(q)\,\mathrm{d}q = N^\alpha \Phi(-Lg^{1/2}N) \to 0$$

and

$$|\mathcal{Q}_1| \leq \int_{\{|q| \geq Lg^{1/2}N\}} \phi(q)\,\mathrm{d}q$$

it follows that $\mathcal{Q}_1 = o(N^{-\alpha})$. On the other hand, (2.10) entails for sufficiently large N that:

$$|N^\alpha \mathcal{Q}_2| \leq N^\alpha \int_{\{|q| \leq Lg^{1/2}N\}} |G_k(v + DN^{-1/2}q) - G_k(v)| \phi(q)\,\mathrm{d}q$$
$$= O(N^{\alpha - 1/2} Lg^{1/2} N).$$

Thus $\mathcal{Q}_2 = o(N^{-\alpha})$ and we get

$$\mathcal{Q} = \mathcal{Q}_1 + \mathcal{Q}_2 = o(N^{-\alpha}).$$

This, together with the fact that

$$1 - \Phi\left(\frac{1-v}{DN^{-1/2}}\right) = o(N^{-\alpha})$$

proves (2.12).

Similarly we can show that

$$\int_{-v/DN^{-1/2}}^{(1-v)/DN^{-1/2}} (1-\Phi(q))^2 \mathrm{d}G_k(v+DN^{-1/2}q)=G_k(v)+\mathrm{o}(N^{-\alpha}). \quad (2.13)$$

Finally, using (2.11), (2.12) and (2.13) and proceeding as in Dupač and Hájek (1969), the proof follows.

By Lemmas 2.3–3.6 and the Residual variance inequality

$$E\left(S_N^+ - \hat{S}_N^+\right)^2 = \mathrm{o}\left(N^{-\alpha} \sum_{i=1}^N c_i^2\right). \quad (2.14)$$

Let us show now that

$$E\left\{\Phi\left(\frac{v-H^*(|X_i|)}{DN^{-1/2}}\right) - u(v-H^*(|X_i|))\right\}^2 = \mathrm{o}(N^{-\alpha}). \quad (2.15)$$

The left-hand side of (2.15) equals $I_1 + I_2$ where

$$\begin{aligned} I_1 &= \int_{-v/DN^{-1/2}}^{0} \Phi^2(p)\,\mathrm{d}G_i(v+DN^{-1/2}p) \\ &= -\int_{-\infty}^{0} \left[I_{(-v/DN^{-1/2},0)}(p)G_i(v+DN^{-1/2}p) - G_i(v)\right] 2\Phi(p)\phi(p)\,\mathrm{d}p \end{aligned}$$

and

$$\begin{aligned} I_2 &= \int_0^{(1-v)/DN^{-1/2}} [1-\Phi(p)]^2 \mathrm{d}G_i(v+DN^{-1/2}p) \\ &= \Phi^2\left(-\frac{(1-v)}{DN^{-1/2}}\right) + \int_0^{\infty} \left[I_{(0,((1-v')/DN^{-1/2}))}(p)G_i(v+DN^{-1/2}p) - G_i(v)\right] \\ &\qquad \times 2\Phi(-p)\phi(p)\,\mathrm{d}p. \end{aligned}$$

Then, proceeding as in the derivation of (2.12) and (2.13) it follows that $I_1 = \mathrm{o}(N^{-\alpha})$ and $I_2 = \mathrm{o}(N^{-\alpha})$ and hence (2.15) holds true.

Now, using (1.6)–(1.9) and (2.15) we obtain

$$E(Y_i - Z_i)^2 = \mathrm{o}\left(N^{-1-\alpha} \sum_{i=1}^N c_i^2\right) \quad (2.16)$$

where

$$Y_i = \sum_{j=1}^N c_j\left\{E\left(a(R_j^+)\operatorname{sgn} X_j | X_i\right) - E\left(a(R_j^+)\operatorname{sgn} X_j\right)\right\}, \quad 1 \leqslant i \leqslant N \quad (2.17)$$

and

$$\begin{aligned} Z_i = N^{-1}&\left[\sum_{\substack{j=1\\ j\neq i}}^N c_j\left(L_j'(v) - M_j'(v)\right)\right]\left[u(v-H^*(|X_i|)) - G_i(v)\right] \\ &+ c_i\left[E\left(\operatorname{sgn} X_i a(R_i^+)|X_i\right) - E\left(\operatorname{sgn} X_i a(R_i^+)\right)\right], \quad 1 \leqslant i \leqslant N. \quad (2.18) \end{aligned}$$

Then, since $\hat{S}_N^+ - E(\hat{S}_N^+) = \sum_{i=1}^N Y_i$, (2.14) and (2.16) entail

$$\mathrm{Var}\left(S_N^+ - \sum_{i=1}^N Z_i \right) = \mathrm{o}\left(N^{-\alpha} \sum_{i=1}^N c_i^2 \right). \tag{2.19}$$

Proceeding as in Lemma 13 of Dupač and Hájek (1969), it is easy to show that (1.12) holds if and only if (1.13) holds with $\sigma_N^2 = \sum_{i=1}^N \mathrm{Var}(Z_i)$ and in this case

$$\lim_{N\to\infty} \mathrm{Var}(S_N^+)/\sigma_N^2 = 1.$$

Finally, the asymptotic normality of $\sum_{i=1}^N Z_i$ with parameters $(0, \sigma_N^2)$ follows as in Lemma 14 of Dupač and Hájek (1969) with the help of (1.2), (1.6), (1.7), (1.13) and the Lindeberg central limit theorem.

Hence, since we have proved that

$$\sum_{i=1}^N Z_i/\sigma_N \xrightarrow{\mathfrak{D}} N(0,1), \qquad \left(S_N^+ - E(S_N^+) - \sum_{i=1}^N Z_i \right) \Big/ \sigma_N \xrightarrow{\mathfrak{L}^2} 0$$

and

$$\mathrm{Var}(S_N^+)/\sigma_N^2 \to 1,$$

we obtain

$$(S_N^+ - E(S_N^+))/(\mathrm{Var}\, S_N^+)^{1/2} \xrightarrow{\mathfrak{D}} N(0,1).$$

Suppose we want to relax condition (1.4) to (1.5). Let us denote the statistic corresponding to (1.4) by S_N^+ and the statistic corresponding to (1.5) by S_N^{+*}. Then, using (1.2) and (1.5)

$$\mathrm{Var}(S_N^+ - S_N^{+*}) = \mathrm{o}\left(N^{-\alpha} \sum_{i=1}^N c_i^2 \right).$$

Consequently, the asymptotic normality of S_N^{+*} easily follows from the last relation and the asymptotic normality of S_N^+.

We have proved Theorem 2.1 under condition $(\tilde{\mathrm{C}}_1^+)$. It remains to show that this set of conditions is implied by the conditions $(\tilde{\mathrm{C}}_2^+)$. The proof of this fact is similar to the implications $(\mathrm{C}_3) \Rightarrow (\mathrm{C}_1)$ and $(\mathrm{C}_2) \Rightarrow (\mathrm{C}_1)$ in Dupač and Hájek (1969, Section 5) and is therefore omitted.

The following theorem shows that under the same conditions $(\tilde{\mathrm{C}}_1^+)$ or $(\tilde{\mathrm{C}}_2^+)$, S_N^+ is asymptotically normal with (simpler) parameters (μ_N^+, σ_N^2). This problem is of practical interest since μ_N^+ and σ_N^2 are easier to evaluate:

Let us define:

$$\mu_N^+ = \sum_{i=1}^N c_i E[\mathrm{sgn}\, X_i \psi(H^*(|X_i|))] \tag{2.20}$$

and

$$\sigma_N^2 = \sum_{i=1}^N \mathrm{Var}\, Z_i' \tag{2.21}$$

where

$$Z_i' = N^{-1}\left[\sum_{\substack{j=1\\ j\neq i}}^{N} c_j\left(L_j'(v) - M_j'(v)\right)\right]\left[u\left(v - H^*(|X_i|)\right) - G_i(v)\right]$$

$$+ c_i\left[\operatorname{sgn} X_i \psi\left(H^*(|X_i|)\right) - E\left(\operatorname{sgn} X_i \psi\left(H^*(|X_i|)\right)\right)\right]. \tag{2.22}$$

Theorem 2.1. *Let S_N^+ be given by* (1.1) *with scores satisfying* (1.5) *where $\psi(t) = u(t - v)$.*

Assume that $(\tilde{\mathrm{C}}_1^+)$ *or* $(\tilde{\mathrm{C}}_2^+)$ *holds, with* (1.12) *replaced by* (1.13), *where σ_N^2 is given by* (2.21).

Then S_N^+ is asymptotically normal with parameters (μ_N^+, σ_N^2) defined in (2.20) *and* (2.21).

Proof. We shall follow the proof of Theorem 2.1 (where it is first assumed that $a(i) = \psi(i/(N+1))$ and that $(\tilde{\mathrm{C}}_i^+)$ holds).

With Y_i and Z_i defined by (2.17) and (2.18) respectively, we have:

$$E(Y_i - Z_i)^2 = \mathrm{o}\left(N^{-1-\alpha}\sum_{j=1}^{N} c_j^2\right) \tag{2.23}$$

uniformly in $1 \leqslant i \leqslant N$, as follows from (2.11) and (2.15).

Define

$$\Delta_i(X_i) = \left\{E\left(\operatorname{sgn} X_i a(R_i^+)|X_i\right) - E\left(\operatorname{sgn} X_i a(R_i^+)\right)\right\}$$

$$-\left\{\operatorname{sgn} X_i u\left(H^*(|X_i|) - v\right) - E\left(\operatorname{sgn} X_i u\left(H^*(|X_i|) - v\right)\right)\right\}.$$

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

$$\operatorname{Var}(\Delta_i) = E\left(\Delta_i^2\right) = \mathrm{O}(N^{-1/2})$$

where $\Delta_i = \Delta_i(X_i)$.

Then, since $Z_i = Z_i' + c_i\Delta_i$, we have

$$E(Z_i - Z_i')^2 = c_i^2\,\mathrm{O}(N^{-1/2}) \tag{2.24}$$

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

But $\hat{S}_N^+ - E(\hat{S}_N^+) = \sum_{i=1}^{N} Y_i$ and from (2.23) and (2.24) we obtain:

$$\operatorname{Var}\left(\hat{S}_N^+ - \sum_{i=1}^{N} Z_i'\right) = \mathrm{o}\left(N^{-\alpha}\sum_{i=1}^{N} c_i^2\right).$$

This together with (2.14) entails

$$\operatorname{Var}\left(S_N^+ - \sum_{i=1}^{N} Z_i'\right) = \mathrm{o}\left(N^{-\alpha}\sum_{i=1}^{N} c_i^2\right).$$

Then, proceeding precisely as in the proof of Theorem 2.1, it follows that:

$$\sum_{i=1}^{N} Z_i'/\sigma_N \xrightarrow{\mathfrak{D}} N(0,1),$$

$$\left(S_N^+ - E(S_N^+) - \sum_{i=1}^{N} Z_i'\right)/\sigma_N \xrightarrow{\mathfrak{L}^2} 0$$

and

$$\operatorname{Var}(S_N^+)/\sigma_N^2 \to 1. \tag{2.25}$$

Further set

$$\rho_i = E(\operatorname{sgn} X_i a(R_i^+) - E(\operatorname{sgn} X_i u(H^*(|X_i|) - v)), \quad 1 \leqslant i \leqslant N.$$

It can be shown (again omitting the details of computation) that:

$$\rho_i = o(N^{-\alpha/2-1/2}) \quad \text{uniformly in } 1 \leqslant i \leqslant N. \tag{2.26}$$

Now, using the inequality:

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

together with (2.26) we get

$$(E(S_N^+) - \mu_N^+)^2 = o\left(N^{-\alpha} \sum_{i=1}^{N} c_i^2\right). \tag{2.27}$$

Finally, making use of (1.13), (2.25) and (2.27), the proof follows.

3. An example

Assume that $X_1, \ldots, X_N$ are i.i.d. with common density function f, and consider the problem of testing the hypothesis of symmetry with normal underlying density H_0: $f(x) = \phi(x)$ against the sequence of shift alternatives

$$H_1: f(x) = \phi(x - \Delta) \quad (\Delta = \Delta_N > 0).$$

Assume that $\Delta \to \infty$ sufficiently slow such that:

$$\limsup_{N \to \infty} \Delta \cdot Lg^{-1/2} N < \tfrac{1}{2}. \tag{3.1}$$

We shall prove that under H_1, (3.1) implies the asymptotic normality of S_N^+ where

$$S_N^+ = \sum_{i=1}^{N} u\left(\frac{R_i^+}{N+1} - v\right) \operatorname{sgn} X_i.$$

First we note that since $\Delta \to \infty$, $H^{*-1}(v) \to \infty$, in such a way that:

$$\lim_{N \to \infty} \left(H^{*-1}(v) - \Delta\right) = 0. \tag{3.2}$$

Using (3.2) it is easy to see that condition (1.11) is satisfied. Also, since $c_1 = \cdots = c_N = 1$, it is easy to check that $\sigma_N^2/\Sigma_{i=1}^N c_i^2 \to 0$, where σ_N^2 is defined by (2.21).

Now, (3.1) implies the existence of a constant $0 < C < \frac{1}{2}$ such that, for sufficiently large N

$$\Delta^2 \leqslant C^2 LgN. \tag{3.3}$$

It can be shown (omitting the details of computation) that for sufficiently large N,

$$\sigma_N^2 / \sum_{i=1}^{N} c_i^2 \geqslant C'\Phi(-2\Delta - 1) \tag{3.4}$$

for some constant $C' > 0$.

Let α and C'' satisfy $\frac{1}{2} < C'' < (1/8C^2)$ and $4C''C^2 < \alpha < \frac{1}{2}$. Thus from (3.3) and (3.4) we have:

$$\sigma_N^2 \Bigg/ \left(N^{-\alpha} \sum_{i=1}^{N} c_i^2 \right) > C' N^{\alpha - 4C''C^2}.$$

The last relation clearly implies the satisfaction of (1.13). The result follows by an application of Theorem 2.2.

References

Dupač, V. (1970). *A Contribution to the Asymptotic Normality of Simple Linear Rank Statistics. Nonparametric Techniques in Statistical Inference*. Cambridge Univ. Press, London, pp. 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. Wahrscheinlichkeitstheorie verw. Geb.* **14**, 308–322.

Koul, H.L. and Staudte, R.G. (1972). Asymptotic normality of signed rank statistics. *Z Wahrscheinlichkeitstheorie verw. Geb.* **22**, 295–300.

Puri, M.L. and Ralescu, S.S. (1980). Asymptotic normality of signed rank statistic with discontinuous score generating function. Preprint.

Reprinted from
Statistics and Probability: Essays in Honor of C.R. Rao,
dited by G. Kallianpur, P.R. Krishnaiah and J.K. Ghosh,
by permission of M.L. Puri's co-author

ON THE ORDER OF MAGNITUDE OF CUMULANTS OF VON MISES FUNCTIONALS AND RELATED STATISTICS

R. N. BHATTACHARYA[1] and M. L. PURI[2]

Indiana University

It is shown that under appropriate conditions the sth cumulant of a von Mises statistic or a U (or V) statistic is $O(n^{-s+1})$, $s \geq 2$, as the sample size n goes to infinity. A possible route toward the derivation of an asymptotic expansion of the characteristic function is indicated.

1. Introduction. The Edgeworth expansion of the characteristic function of a normalized sum of n independent and identically distributed (i.i.d) random variables derives from the order of magnitude $O(n^{-(s-2)/2})$ of the sth *cumulant* ($s \geq 2$) (See, e.g., Bhattacharya, 1977). For statistics which may be expressed as or approximated by polynomials in several average sample characteristics (e.g., (i) polynomials in sample moments and (ii) maximum likelihood estimators in the regular case), the validity of the so-called "formal Edgeworth expansion" depends crucially on the above order of magnitude of the sth cumulant ($s \geq 2$) of the normalized statistic (see Bhattacharya and Ghosh, 1978). In this note it is shown that cumulants of normalized *U-statistics* and *von Mises functionals* have the above order of magnitude, if certain conditions are satisfied. For general background on these statistics we refer to von Mises (1947) and Serfling (1980). Assuming the validity of (a) the above order of magnitude of the cumulants and (b) the Edgeworth expansion of the distribution function of a von Mises functional, Withers (1980) has given an algorithm for computing the coefficients in the asymptotic expansion. Some of the moment computations in Section 2 are similar to those in Withers (loc. cit). In Section 3 a new method of derivation of Cramér-Edgeworth expansions of characteristic functions of a class of statistics is provided.

2. Moments and cumulants. Let χ be a separable metric space (e.g., a subset of R^d), $\mathscr{B}_\chi$ its Borel sigma field, and P a given probability measure on $\mathscr{B}_\chi$, whose *support* is S. Let $\mathscr{P}_f$ denote the set of all probability measures on $\mathscr{B}_\chi \cap S$ having *finite supports.* Endow $\mathscr{P}_f \cup \{P\}$ with the weak-star topology. Consider for each n the product space $(\chi^n, \mathscr{B}_{\chi^n})$, and let $X_1, \cdots, X_n$ be the n coordinate random variables. Let $G^{\otimes n} = G \times G \times \cdots \times G$ denote the product probability measure on $\mathscr{B}_{\chi^n}$, where G is a probability measure on $\mathscr{B}_\chi$. Under $G^{\otimes n}$ the random variables $X_1, \cdots, X_n$ are i.i.d. with common distribution G. We shall write E_G to denote expectation under $G^{\otimes n}$. Denote the *empirical distribution* of the "observations" $X_1, \cdots, X_n$ by F_n, i.e., $F_n = n^{-1}\sum_{i=1}^n \delta_{X_i}$, where δ_x is the *Dirac measure* with point mass at x.

Let $h(x_1, x_2, \cdots, x_r)$ be a real-valued, Borel measurable, symmetric function on χ^r, for some $r \geq 2$. Define the *V-statistic* (with *kernel h*)

$$V_n = n^{-r} \textstyle\sum_{i_1=1}^n \cdots \sum_{i_r=1}^n h(X_{i_1}, X_{i_2}, \cdots, X_{i_r}), \tag{2.1}$$

and the *U-statistic* (with *kernel h*)

$$U_n = \binom{n}{r}^{-1} \sum h(X_{i_1}, X_{i_2}, \cdots, X_{i_r}) \tag{2.2}$$

where the summation is over $1 \leq i_1 < i_2 < \cdots < i_r \leq n$.

Received February 1982; revised August 1982.

[1] Research supported by NSF Grant MCS 8201628.

[2] Research supported by Office of Naval Research under Contract N00014-77-C-0659.

AMS 1970 *subject classifications.* Primary 62E20; secondary 62G05, 62G10.

Key words and phrases. V-Statistics, U-statistics, Edgeworth expansion.

Theorem 2.1. (a) *If for some integer $s \geq 3$ one has*

$$E_P|h(X_{j_1}, X_{j_2}, \cdots, X_{j_r})|^s < \infty \tag{2.3}$$

for all choices of $j_1, j_2, \cdots, j_r (1 \leq j_1, j_2, \cdots, j_r \leq r)$, then the pth *cumulant $k_{p,n}(P)$ of V_n under P is of the form*

$$k_{p,n}(P) = \textstyle\sum_{m=p-1}^{s-1} n^{-m}\lambda_{m,p}(P) + o(n^{-s+1}), \qquad (2 \leq p \leq s). \tag{2.4}$$

The quantities $\lambda_{m,p}(P)$ are independent of n.

(b) *Suppose that, for some integer $s \geq 3$ one has*

$$E_P|h(X_1, X_2, \cdots, X_r)|^s < \infty. \tag{2.5}$$

Then the cumulants of the statistic U_n also are of the form (2.4).

Proof. (a) Write

$$V_n = \int \cdots \int h(x_1, \cdots, x_r) F_n(dx_1) \cdots F_n(dx_r). \tag{2.6}$$

For $G = \sum_{i=1}^q \alpha_i \delta_{y_i}$ in $\mathscr{P}_f$, F_n may be expressed as $\sum_{i=1}^q \hat{\alpha}_i \delta_{y_i}$ (with $G^{\otimes n}$-probability one), where $\hat{\alpha}_i$ is the proportion of y_i's in the "sample" $\{X_1, \cdots, X_n\}$. Thus V_n becomes a polynomial in the q-variables $\hat{\alpha}_i$, $1 \leq i \leq q$. Hence by a result of James and Mayne (1962) (this may also be derived from the results of Leonov and Shiryaev, 1959), the pth cumulant of V_n under G is of the form

$$k_{p,n}(G) = \textstyle\sum_{m=p-1}^{rp-1} n^{-m}\lambda_{m,p}(G), \qquad (p \geq 2). \tag{2.7}$$

On the other hand, for all G such that $E_G|V_n|^p < \infty$, one has (for all $n > rp$)

$$\begin{aligned} E_G V_n^p &= n^{-rp} E_G \int \cdots \int \left(\textstyle\prod_{t=1}^p h(x_{r(t-1)+1}, \cdots, x_{rt})\right) \\ &\quad \cdot (\delta_{X_1} + \delta_{X_2} + \cdots + \delta_{X_n})(dx_1) \cdots (\delta_{X_1} + \delta_{X_2} + \cdots + \delta_{X_n})(dx_{rp}) \\ &= n^{-rp} \int \cdots \int \left(\textstyle\prod_{t=1}^p h(x_{r(t-1)+1}, \cdots, x_{rt})\right) \\ &\quad \Bigg[\textstyle\sum_{m=1}^{rp} \dfrac{n!}{(n-m)!} \sum_2 \sum_1 E_G\{\delta_{X_1}(dx_{j_{11}})\delta_{X_1}(dx_{j_{12}}) \cdots \delta_{X_1}(dx_{j_{1s_1}}) \\ &\quad \delta_{X_2}(dx_{j_{21}}) \cdots \delta_{X_2}(dx_{j_{2s_2}}) \cdots \delta_{X_m}(dx_{j_{m1}}) \cdots \delta_{X_m}(dx_{j_{ms_m}})\}\Bigg]. \end{aligned} \tag{2.8}$$

Here, for a given m, $\sum_2$ denotes summation over all collections of m positive integers $\{s_1, s_2, \cdots, s_m\}$ satisfying $\sum s_i = rp$; for a given collection $\{s_1, s_2, \cdots, s_m\}$, $\sum_1$ denotes summation over all partitions of $\{1, 2, \cdots, rp\}$ into m groups of $s_1, s_2, \cdots, s_m$ elements, a typical partition being $(\{j_{11}, j_{12}, \cdots, j_{1s_1}\}, \{j_{21}, j_{22}, \cdots, j_{2s_2}\}, \{j_{m1}, j_{m2}, \cdots, j_{ms_m}\})$. Denote by $H_s(G)$ the distribution of the s-dimensional random vector $(X_1, X_1, \cdots, X_1)$ under G, and let $H_{s_1,s_2,\cdots,s_m}(G)$ stand for the measure

$$\begin{aligned} H_{s_1,s_2,\cdots,s_m}(G)(dx_1 dx_2 \cdots dx_{rp}) &= \textstyle\sum_1 H_{s_1}(G)(dx_{j_{11}} dx_{j_{12}} \cdots dx_{j_{1s_1}}) H_{s_2}(G) \\ &\quad \cdot (dx_{j_{21}} dx_{j_{22}} \cdots dx_{j_{2s_2}}) \cdots H_{s_m}(dx_{j_{m1}} dx_{j_{m2}} \cdots dx_{j_{ms_m}}). \end{aligned} \tag{2.9}$$

Also note that

$$\begin{aligned} n^{-rp} \frac{n!}{(n-m)!} &= n^{-rp} n(n-1) \cdots (n-m+1) \\ &= \textstyle\sum_{m'=1}^{m} (-1)^{m-m'} n^{-rp+m'} \theta(m-m'; m-1), \end{aligned} \tag{2.10}$$

where $\theta(i; N)$ is the sum of all products of i distinct integers taken from $\{1, 2, \cdots, N\}$,

$\theta(0; N) = 1$. From (2.8)–(2.10) one obtains

$$E_G V_n^p = \sum_{m=1}^{rp} \sum_{m'=1}^{m} (-1)^{m-m'} n^{-rp+m'} \theta(m - m'; m - 1)$$
$$\cdot \int \cdots \int (\prod_{t=1}^{p} h(x_{r(t-1)+1}, \cdots, x_{rt})) \sum_2 H_{s_1,s_2,\cdots,s_m}(G)(dx_1 \cdots dx_{rp})$$

$$= \sum_{j=0}^{rp-1} n^{-j} \{ \sum_{m=rp-j}^{rp} (-1)^{m-rp+j} \theta(m - rp + j; m - 1) \tag{2.11}$$
$$\cdot \int \cdots \int (\prod_{t=1}^{p} h(x_{r(t-1)+1}, \cdots, x_{rt})) \sum_2 H_{s_1,s_2,\cdots,s_m}(G)(dx_1 \cdots dx_{rp}) \}$$
$$= \sum_{j=0}^{rp-1} n^{-j} \mu_{j,p}(G), \qquad (1 \le p \le s),$$

say. Here $\mu_{j,p}(G)$ is a linear combination (with coefficients not depending on n, G or h) of terms like

$$\int \cdots \int (\prod_{t=1}^{p} h(x_{r(t-1)+1}, \cdots, x_{rt})) H_{s_1,s_2,\cdots,s_m}(G)(dx_1 \cdots dx_{rp}). \tag{2.12}$$

Using the familiar relations between moments and cumulants one has

$$k_{p,n}(G) = \sum_{j=0}^{rp-1} n^{-j} \bar{\lambda}_{j,p}(G), \tag{2.13}$$

where $\bar{\lambda}_{j,p}(G)$ is a polynomial in $\mu_{j,p'}(1 \le p' \le p)$, whose coefficients are absolute constants. Since the map $G \to H_{s_1}(G)$ is continuous in the weak-star topology, so is the map $G \to H_{s_1,s_2,\cdots,s_m}(G)$. It follows that for a bounded continuous h the integral (2.12) is a weak-star continuous function of G; this implies that the maps $G \to \mu_{j,p}$ and, therefore, $G \to \bar{\lambda}_{m,p}(G)$ are continuous. If $p \ge 2$, then $\bar{\lambda}_{m,p}(G) = 0$ for $2 \le m < p - 1$ and $G \in \mathscr{P}_f$. Also there exists $G_N \in \mathscr{P}_f (N = 1, 2, \cdots)$ such that G_N converges to P (This is where the separability of χ is made use of; see, e.g., Parthasarathy (1967), Theorem 6.3). Therefore, one must have $\bar{\lambda}_{m,p}(P) = 0$ for $1 \le m < p - 1$. This completes the proof of (a) for bounded continuous h. Since functions of the form $\prod_{t=1}^{p} h(x_{r(t-1)+1}, \cdots, x_{rt})$ belonging to $L^1(\chi^{rp}, H_{s_1,s_2,\cdots,s_m}(P))$ may be approximated (in L^1) by continuous bounded functions of the same form, the proof is complete. Note that for this last argument (2.3) is needed.

(b) First assume (i) $h(x_1, x_2, \cdots, x_r) = 0$ *if* $x_i = x_j$ *for some* $i, j (i \ne j)$. Then the cumulants of U_n satisfy (2.4), since

$$U_n = \left(\frac{n!}{(n-r)!}\right)^{-1} n^r V_n = \left(\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{r-1}{n}\right)\right)^{-1} V_n = (1 + o(1)) V_n.$$

Next, instead of (i) assume (ii) *P has no atoms*. Then modify h so as to satisfy (i); this does not change U_n, except on a set of probability zero. Finally, consider an arbitrary P. Let D be its set of atoms. Let D' be a subset of reals in one-one correspondence with D. Consider the space $\chi' = (\chi \backslash D) \cup R$, with $\chi \backslash D$ and R each carrying its own topology but their union is topologically disconnected. Then P lifted to this space χ' (by placing the discrete mass on D') is a weak-star limit of nonatomic probability measures. Extend h to $(\chi')^r$ by setting it zero if any coordinate is in $R \backslash D'$. Now apply an argument entirely analogous to that in the preceding paragraph.

REMARK 2.1.1. The U_n and V_n defined above are not centered around their expectations (under P). Centering has been avoided deliberately to ensure that h does not depend on P. For the general von Mises functional considered below centering seems unavoidable; this causes some technical problems.

REMARK 2.1.2. Under the hypotheses of Theorem 2.1 the pth cumulants of the normalized statistics $\sqrt{n}(V_n - EV_n)$, $\sqrt{n}(U_n - EU_n)$ are of the order $O(n^{-(p-2)/2})$, $2 \le p \le s$.

Let T be a *von Mises functional* defined on $\mathscr{P}_f \cup \{P\}$, and let the statistic $T(F_n)$ have the expansion

$$T(F_n) - T(G) = \sum_{i=1}^{r} \int \cdots \int T^{(i)}(G; x_1, x_2, \cdots, x_i) \prod_{j=1}^{i} (F_n - G)(dx_j) + R_n$$
$$(2.14) \qquad = \sum_{i=1}^{r} V_{n,i}(G) + R_n \qquad (G \in \mathscr{P}_f \cup \{P\}),$$

where $T^{(i)}$ is a real-valued, symmetric (in the arguments $x_1, \cdots, x_i$), Borel measurable function on $\mathscr{P}_f \cup \{P\} \times \chi^i$ satisfying

$$(2.15) \qquad E_P|T^{(i)}(P; X_{j_1}, X_{j_2}, \cdots, X_{j_i})|^s < \infty, \qquad (1 \le i \le r),$$

for all $1 \le j_1, j_2, \cdots, j_i \le r$, and the "remainder term" R_n satisfies

$$(2.16) \qquad E_P|R_n|^p = o(n^{-s+1}), \qquad (1 \le p \le s).$$

Write

$$(2.17) \qquad V_n(G) = \sum_{i=1}^{r} V_{n,i}(G).$$

Then

$$(2.18) \qquad E_G V_n^p(G) = \sum\nolimits_3 V_{n,i_1}(G) V_{n,i_2}(G) \cdots V_{n,i_p}(G),$$

where $\sum_3$ denotes summation over all p-tuples $(i_1, i_2, \cdots, i_p)$ such that $1 \le i_1, \cdots, i_p \le r$. Now let $I_p = i_1 + \cdots i_p$ and write, as in (2.8),

$$E_G V_n^p(G) = \sum\nolimits_3 E_G \int \cdots \int \left(\prod_{\ell=1}^{p} T^{(i_\ell)}(G; x_{I_{\ell-1}+1}, \cdots, x_{I_\ell})\right)$$
$$(2.19) \qquad \cdot \Bigg[n^{-I_p} \sum_{m=1}^{[I_p/2]} \frac{n!}{(n-m)!} \sum\nolimits_2' \sum\nolimits_1' (\delta_{X_1} - G)(dx_{j_{11}})(\delta_{X_1} - G)(dx_{j_{12}})$$
$$\cdots (\delta_{X_1} - G)(dx_{j_{1s_1}}) \cdots (\delta_{X_m} - G)(dx_{j_{m1}}) \cdots (\delta_{X_m} - G)(dx_{j_{ms_m}}) \Bigg].$$

Here, for a given m, $\sum_2'$ denotes summation over all collections of m integers $\{s_1, s_2, \cdots, s_m\}$ satisfying $s_i \ge 2$ and $\sum s_i = I_p$; and $\sum_1'$ denotes, for each collection $\{s_1, s_2, \cdots, s_m\}$, summation over all partitions of $\{1, 2, \cdots, I_p\}$ into m subgroups of $s_1, s_2, \cdots, s_m$ elements such as $(\{j_{11}, j_{12}, \cdots, j_{1s_1}\}, \cdots, \{j_{m1}, j_{m2}, \cdots, j_{ms_m}\})$. Note that expectations of terms involving $s_i = 1$ for some i vanish. Next let $H_q(i_1, i_2, \cdots, i_\ell; G)$ denote the distribution of a q-dimensional random vector whose i_1th, $\cdots$, i_ℓth coordinates are X_1, while the remaining coordinates are i.i.d with distribution G and independent of X_1. Write $\tilde{H}_q(G)$ for the *signed* measure

$$(2.20) \qquad \tilde{H}_q(G) = \sum_{\ell=0}^{q} (-1)^{q-\ell} \sum\nolimits_4 H_q(i_1, \cdots, i_\ell; G),$$

where $\sum_4$ denotes summation over all choices $\{i_1, i_2, \cdots, i_\ell\}$ of ℓ distinct integers from $\{1, 2, \cdots, q\}$. Now define

$$(2.21) \qquad \tilde{H}_{s_1, s_2, \cdots, s_m}(G; dx_1 dx_2 \cdots dx_{I_p})$$
$$= \sum\nolimits_1' \tilde{H}_{s_1}(G)(dx_{j_{11}} dx_{j_{12}} \cdots dx_{j_{1s_1}}) \cdots \tilde{H}_{s_m}(G)(dx_{j_{m1}} dx_{j_{m2}} \cdots dx_{j_{ms_m}}).$$

Then, as in (2.10),

$$E_G V_n^p(G) = \sum\nolimits_3 \Bigg[\sum_{j=\left[\frac{I_p+1}{2}\right]}^{I_p-1} n^{-j} \Bigg\{ \sum_{m=I_p-j}^{[I_p/2]} (-1)^{m-I_p+j} \theta(m - I_p + j; m-1)$$
$$(2.22) \qquad \cdot \int \cdots \int \left(\prod_{\ell=1}^{p} T^{(i_\ell)}(G; x_{I_{\ell-1}+1}, \cdots, x_{I_\ell})\right) \sum\nolimits_2' \tilde{H}_{s_1, \cdots, s_m}(G; dx_1 \cdots dx_{I_p}) \Bigg\} \Bigg].$$

For $G = \sum_{i=1}^{q} \alpha_i \delta_{y_i}$, $V_n(G)$ is a polynomial in $\hat{\alpha}_i - \alpha_i$, so that the pth cumulant of $V_n(G)$ is of the order $O(n^{-p+1})$ under $G(2 \le p \le s)$. In view of (2.16) and (2.22), the proof of the following theorem is now complete.

THEOREM 2.2 *Suppose that* (2.14)-(2.16) *hold. Assume, in addition, that there exists a sequence* $\{G_N: N \geq 1\}$ *having finite support such that*

$$(2.23)\qquad \lim_{N\to\infty} E_{G_N}\left(\prod_{t=1}^{p} T^{(i_t)}(G_N; X_{t_1}, \cdots, X_{t_{i_t}})\right) = E_P\left(\prod_{t=1}^{p} T^{(i_t)}(P; X_{t_1}, \cdots, X_{t_{i_t}})\right)$$

for all $1 \leq i_1, i_2, \cdots, i_p \leq r$, *and all* $1 \leq t_1, t_2, \cdots, t_{i_t} \leq rp$ $(1 \leq t \leq p)$. *Then the* pth *cumulant of* $T(F_n)$ *under* P *is of the order* $O(n^{-p+1})$ *for* $2 \leq p \leq s$.

REMARK 2.2.1. Notice that the statement "condition (2.23) holds for *some* $\{G_N; N \geq 1\} \subset \mathscr{P}_f$" is much weaker than the statement "condition (2.23) holds for all sequences $\{G_N: N \geq 1\}$ converging to P (weak-star)", the latter being equivalent to saying that the integral is weak-star continuous at P (on $\mathscr{P}_f \cup \{P\}$). To illustrate this point, note that even such functionals as $T(G) = \int x^k G(dx)$, $k \geq 1$, are *not* weak-star continuous on $\mathscr{P}_f \cup \{P\}$, where P is a probability measure on the line having a finite kth moment. The difficulty is that one may place a mass $O(N^{-k/2})$ at $x = N$ which goes to zero to ensure weak-star convergence, but is large enough to blow up the integral as $N \to \infty$. On the other hand, one may integrate (with respect to P) a step-function approximation, $f_N(x)$ to x^k, which amounts to integrating x^k with respect to an appropriate $G_N \in \mathscr{P}_f$; and the latter integral $\int x^k G_N(dx)$ will converge to $\int x^k P(dx)$, as the intervals of constancy decrease to zero in width. These considerations apply to more general functions (see, Serfling (1980), pages 214–216, for examples).

REMARK 2.2.2. The fact that the sth cumulant of V_n (or T_n) is $O(n^{-s+1})$ when G has finite support means the vanishing of a number of polynomials in the variables $\mu_p(G)$. One should be able to prove that these polynomials are identically zero by showing that the $\mu_p(G)$'s assume a broad enough spectrum of values as G ranges over the set of all probability measures having finite support. This would enable one to dispense with the condition (2.23) in Theorem 2.2. However, we are unable to make this algebraic argument firm.

Finally, the method used here should be more widely applicable in deriving orders of magnitudes of cumulants.

3. A method of derivation of Edgeworth expansions of characteristic functions, and an unsolved problem. In the present section we provide a method (which appears to be new) for the derivation of Cramér-Edgeworth expansions of characteristic functions of a class of statistics T_n having zero means, finite moment generating functions (m.g.f.'s), and cumulants $\chi_{p,n}$ satisfying

$$(3.1)\qquad \chi_{p,n} = n^{-(p-2)/2}\lambda_p + o(n^{-(p-2)/2}), \qquad (p \geq 2), \qquad \lambda_2 > 0.$$

Let

$$(3.2)\qquad f_n(\xi) = E\exp\{i\xi T_n\}$$

denote the characteristic function of T_n. One may write

$$(3.3)\qquad f_n(\xi) = f(i\xi, \varepsilon),$$

with $\varepsilon = n^{-1/2}$. Under the additional assumption that $f(i\xi, \varepsilon)$ has an absolutely convergent power series expansion in ξ and ε in a neighborhood of the origin (0, 0), it is shown in Theorem 3.1 that $f_n(\xi)$ and its derivatives have a proper asymptotic expansion of the Cramér-Edgeworth type. The *unsolved problem* is to identify a large enough class of von Mises functionals for which this analyticity holds. In particular, we do not know if the analyticity property holds for U-statistics (see (2.2)) with kernels h satisfying:

$$(3.4)\qquad E\exp\{th(X_1, X_2, \cdots, X_r)\} < \infty, \qquad (-\infty < t < \infty).$$

In remarks following the corollaries to Theorem 3.1 it is shown that the assumption of analyticity does hold for some special classes. We expect the moment computations of Section 2 to be crucial in resolving the problem of analyticity in the general case.

THEOREM 3.1. *Let $T_n (n = 1, 2, \cdots)$ be a sequence of random variables having zero means. Assume that* (i) $E\exp\{tT_n\} < \infty$ *for all* $t(-\infty < t < \infty)$ *and* n, (ii) $f(i\xi, \varepsilon)$ *can be extended as an analytic function* $f(z, \eta)$ *of the complex variables* z *and* η *in a neighborhood of the origin* $(0, 0)$ *in* C^2, *and* (iii) *the cumulants* $\chi_{p,n}$ *of* T_n *satisfy* (3.1). *Then the following results hold*:

(a) *There exist a positive constant δ_0 and polynomials P_j, whose coefficients do not depend on n, such that for all ξ, $-\delta_0\sqrt{n} < \xi < \delta_0\sqrt{n}$, one has*

$$f_n(\xi) = \exp\left\{-\frac{\lambda_2}{2}\xi^2\right\}(1 + \textstyle\sum_{j=1}^{\infty} n^{-j/2}P_j(i\xi)).$$

(b) *For every pair of integers m and p satisfying $p \geq 2$, $0 \leq m \leq p$, there exist positive constants δ_0, c_1, c_2 such that*

$$\left|\frac{d^m}{d\xi^m}\left[f_n(\xi) - \exp\left\{-\frac{\lambda_2}{2}\xi^2\right\}(1 + \textstyle\sum_{j=1}^{p-2} n^{-j/2}P_j(i\xi))\right]\right|$$
$$\leq \frac{c_1}{n^{(p-1)/2}}[|\xi|^{p+1-m} + |\xi|^{3(p-1)+m}]\exp\{-c_2\xi^2\}, \qquad (|\xi| < \delta_0\sqrt{n}).$$

PROOF. Since $f(z, \eta)$ is analytic in a neighborhood of $(0, 0)$, and $f(0, 0) = 1$, $\phi(z, \eta) = \log f(z, \eta)$ (we take the principal branch of the logarithm) is defined and analytic in a neighborhood of $(0, 0)$. In view of (3.1) and the fact that $ET_n = 0$, one may express $\phi(z, \eta)$ as

$$\begin{aligned}\phi(z, \eta) &= \frac{z^2}{2!}(\lambda_2 + \textstyle\sum_{j=1}^{\infty}\lambda_{2,j}\eta^j) + \cdots + \frac{z^k}{k!}\eta^{k-2}(\sum_{j=0}^{\infty}\lambda_{k,j}\eta^j) + \cdots \\ &= z^2\left[\textstyle\sum_{k=2}^{\infty}\frac{(\eta z)^{k-2}}{k!}(\sum_{j=0}^{\infty}\lambda_{k,j}\eta^j)\right], \qquad (\lambda_{k,0} = \lambda_k).\end{aligned} \tag{3.5}$$

Since this last series is absolutely convergent in a neighborhood of $(0, 0)$, so is the series within square brackets. Let δ_1, δ_2 be two positive numbers such that this last series is absolutely convergent for $|z| = \delta_1$, $|\eta| = \delta_2$. Then

$$\textstyle\sum_{k=2}^{\infty}\frac{(\delta_1\delta_2)^{k-2}}{k!}\sum_{j=0}^{\infty}|\lambda_{k,j}|\delta_2^j < \infty. \tag{3.6}$$

It follows that (3.5) is absolutely convergent for $|z\eta| \leq \delta_1\delta_2$ and $|\eta| \leq \delta_2$. Therefore, the last expression in (3.5) defines an analytic function in the region $D = \{(z, \eta) \in C^2\colon |z| < \delta_1\delta_2/|\eta|, |\eta| < \delta_2\}$, and over this region $\exp\{\phi(z, \eta)\}$ defines an analytic continuation of $f(z, \eta)$. We shall refer to this extension also by $f(z, \eta)$. Since the characteristic function $\xi \to f_n(\xi)$ is entire (by assumption (i)) and since analytic continuations are unique, $f_n(\xi) = f(i\xi, n^{-1/2})$ for $-\infty < \xi < \infty$ (note that one could not assume a priori that this equality holds between f_n and the analytically extended f). In addition, on D one has

$$|f(z, \eta) - 1| < c' < 1, \tag{3.7}$$

for some constant c', and $\phi(z, \eta)$ is the principal branch of the logarithm of $f(z, \eta)$ on D. The relations (3.5) now hold on D and one may rewrite the first relation in (3.5) as

$$\log f(z, \eta) - \frac{\lambda_2}{2!}z^2 = \textstyle\sum_{j=1}^{\infty}\eta^jQ_j(z), \qquad (z, \eta) \in D, \tag{3.8}$$

where Q_j is a polynomial of degree $j + 2$. Thus

$$(3.9)\qquad f(z, \eta) \exp\left\{-\frac{\lambda_2}{2} z^2\right\} = \exp\left\{\sum_{j=1}^{\infty} \eta^j Q_j(z)\right\} = 1 + \sum_{j=1}^{\infty} \eta^j P_j(z), \qquad (z, \eta) \in D,$$

where P_j's are appropriate polynomials. From (3.9) one gets

$$(3.10)\qquad f(z, \eta) = \exp\left\{\frac{\lambda_2}{2} z^2\right\} (1 + \sum_{j=1}^{\infty} \eta^j P_j(z)), \qquad (z, \eta) \in D,$$

and, in particular (with $z = i\xi$, $\eta = n^{-1/2}$),

$$(3.11)\qquad f_n(\xi) = \exp\left\{-\frac{\lambda_2}{2} \xi^2\right\} (1 + \sum_{j=1}^{\infty} n^{-j/2} P_j(i\xi)), \qquad (-\delta_1\delta_2 \sqrt{n} < \xi < \delta_1\delta_2 \sqrt{n}).$$

This proves part (a). To prove part (b) one may first approximate $\log f(z, \eta)$ by

$$(3.12)\qquad \phi_p(z, \eta) = z^2 \sum_{k=2}^{p+2} \frac{(\eta z)^{k-2}}{k!} (\sum_{j=0}^{\infty} \lambda_{k,j} \eta^j).$$

Writing

$$(3.13)\qquad \psi(z, \eta) = \phi(z, \eta) - \frac{\lambda_2}{2} z^2, \qquad \psi_p(z, \eta) = \phi_p(z, \eta) - \frac{\lambda_2}{2} z^2,$$

one has (using (3.6), or analyticity on D)

$$(3.14)\qquad |\phi(z, \eta) - \phi_p(z, \eta)| = |\psi(z, \eta) - \psi_p(z, \eta)| \le c_3 |\eta|^{p-1} |z|^{p+1}, \qquad (z, \eta) \in D,$$

for an appropriate constant c_3. By (3.6) and (3.14), if δ_1 is small, then

$$(3.15)\qquad |\exp\{\psi(z, \eta)\} - \exp\{\psi_p(z, \eta)\}| \le c_4 |\eta|^{p-1} |z|^{p+1} \exp\left\{\frac{\lambda_2 |z|^2}{4}\right\},$$

for some constant c_4; this may be written as

$$(3.16)\qquad \left| e^{-\frac{\lambda_2}{2} z^2} [f(z, \eta) - \exp\{\phi_p(z, \eta)\}] \right| \le c_4 |\eta|^{p-1} |z|^{p+1} \exp\left\{\frac{\lambda_2 |z|^2}{4}\right\}.$$

Letting $z = i\xi$, $\eta = n^{-1/2}$, (3.16) becomes

$$(3.17)\qquad |f_n(\xi) - \exp\{\phi_p(i\xi, n^{-1/2})\}| \le c_4 n^{-(p-1)/2} |\xi|^{p+1} \exp\left\{-\frac{\lambda_2}{4} \xi^2\right\}, \qquad (|\xi| < \delta_1\delta_2 \sqrt{n}).$$

The comparison of $\exp\{\phi_p(i\xi, n^{-1/2})\}$ with $\exp\left\{-\frac{\lambda_2}{2} \xi^2\right\} (1 + \sum_{j=1}^{p-2} n^{-j/2} P_j(i\xi))$ is carried out exactly as in Lemmas 9.7, 9.8 in Bhattacharya and Ranga Rao (1976). □

COROLLARY 3.1.1. *Under the hypothesis of Theorem* 3.1 *one has the Berry-Esseen bound*

$$(3.18)\qquad \sup_x |P(T_n \le x) - \Phi_{\lambda_2}(x)| \le c n^{-1/2},$$

for some constant $c > 0$. *Here* Φ_{λ_2} *is the normal distribution function with mean zero and variance* λ_2.

PROOF. Use Theorem 3.1 (b) and Esseen's inequality (see Lemmas 12.1, 12.2 in Bhattacharya and Ranga Rao (1976)). □

COROLLARY 3.1.2. *Assume the hypothesis of Theorem* 3.1. *If, for some* $p \ge 2$, g *is a* p*-times continuously differentiable function on* $\mathbb{R}^1$ *such that* $\sup\{(1 + |x|^p)|g^{(m)}(x)|$:

$x \in I\!R^1\} < \infty$ *for* $0 \le m \le p$, *then*

$$(3.19) \qquad |Eg(T_n) - \int_{I\!R^1} g(x)\left[1 + \textstyle\sum_{j=1}^{p-2} n^{-j/2} P_j\left(-\frac{d}{dx}\right)\right]\phi_{\lambda_2}(x)dx| \le dn^{-(p-1)/2}$$

for some positive constant d.

PROOF. One may apply the method of Götze and Hipp (1978) to the estimate in Theorem 3.1 (b) to derive (3.19) directly. Alternatively, first establish (3.19) for the class of all Schwartz functions as in Bhattacharya and Ranga Rao (1976), Theorem 20.7, expressing the error estimate in terms of a Sobolev norm; then extend the result to a wider class by completion in the Sobolev norm. □

REMARK 3.1.3. Let $X_1, X_2, \cdots$ be an i.i.d. sequence having mean zero and a positive variance. The hypothesis of Theorem 3.1 is satisfied for the statistics $T_n = n^{-1/2}(X_1 + \cdots + X_n)$ if the m.g.f. of X_1 is finite everywhere. Of course, in this classical case Theorem 3.1 (b) holds under less stringent assumptions (see, e.g., Bhattacharya and Ranga Rao (1976), Chapter 2). Note, however, the conclusion of part (a) of Theorem 3.1 requires stronger assumptions than finiteness of moments.

REMARK 3.1.4 Let U_n be a U-statistic with kernel h (see (2.2)). Assume, without loss of generality, that $Eh(X_1, X_2, \cdots, X_r) = 0$. If $E\exp\{th(X_1, \cdots, X_r)\} < \infty$ for all t, $-\infty < t < \infty$, then hypothesis (i) of Theorem 3.1 is satisfied for the statistic $T_n = \sqrt{n}\ U_n$ (see Serfling (1980), Lemma C, page 200). In addition, assume $E\phi^2(X_1) = \lambda_2 > 0$, where $\phi(x) = Eh(x, X_2, \cdots, X_r)$. Then T_n is asymptotically normal (see Serfling (1980), Theorem A, page 192) and, by Theorem 2.1 (b), hypothesis (iii) of Theorem 3.1 also holds. It would be of great interest to see if hypothesis (ii) of Theorem 3.1 is a consequence of the above assumptions. We emphasize that this is the main unresolved problem in the context of the present article. For kernels h which are sums of products of functions of single variables, analyticity of $f(z, \eta)$ in a neighborhood of the origin in C^2 has been proved by methods of statistical mechanics (see, e.g., Ruelle, 1969). However, for these special kernels an adequate theory of Edgeworth expansions has been derived in Bhattacharya and Ghosh (1978) under less stringent assumptions.

REMARK 3.1.5. Some partial expansions of characteristic functions of U-statistics have been obtained by Callaert, Janssen and Veraverbeke (1980).

Acknowledgement. The authors wish to thank the referee for suggesting that some terse comments in an earlier draft be expanded; the present Section 3 constitutes the resulting expansion. Thanks are also due to the Associate Editor for indicating that asymptotic expansions for von Mises functionals are of potential use in robust estimation. Finally, we are indebted to Charles M. Newman for pointing out that the analyticity hypothesis in Theorem 3.1 has been verified in the context of statistical mechanics for the special class of kernels mentioned at the end of Remark 3.1.4.

REFERENCES

BHATTACHARYA, R. N. (1977). Refinements of the multidimensional central limit theorem and applications. *Ann. Probab.* **5** 1–27.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

CALLAERT, H., JANSSEN, P. and VERAVERBEKE, N. (1980). An Edgeworth expansion for U-statistics. *Ann. Probab.* **8** 299–312.

GÖTZE, F. and HIPP, C. (1978). Asymptotic expansions in the central limit theorem under moment conditions. *Z. Wahrsch. verw. Gebiete* **42** 67–87.

JAMES, G. S. and MAYNE ALAN, J. (1962). Cumulants of functions of random variables. *Sankhya Ser. A* **24** 47–54.

LEONOV, V. P. and SHIRYAEV, A. N. (1959). On a method of calculation of semi-invariants. *Theor. Probab. Appl.*, 4 319–329.

PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic, New York.

RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results.* Benjamin, Reading.

SERFLING, ROBERT J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

WITHERS, C. S. (1980). The distribution and quantiles of a regular functional of the empirical distribution, Report No. 96, D.S.I.R., Wellington, New Zealand.

DEPARTMENT OF MATHEMATICS
SWAIN HALL EAST
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47405

ON BERRY–ESSÉEN RATES, A LAW OF THE ITERATED LOGARITHM AND AN INVARIANCE PRINCIPLE FOR THE PROPORTION OF THE SAMPLE BELOW THE SAMPLE MEAN*

STEFAN RALESCU and MADAN L. PURI

Brown University and Indiana University Communicates by M. Rosenblatt

Let $F_n(x)$ be the empirical distribution function based on n independent random variables $X_1, \ldots, X_n$ from a common distribution function $F(x)$, and let $\bar{X} = \sum_{i=1}^n X_i/n$ be the sample mean. We derive the rate of convergence of $F_n(\bar{X})$ to normality (for the regular as well as nonregular cases), a law of iterated logarithm, and an invariance principle for $F_n(\bar{X})$.

1. Introduction

Let $X_1, \ldots, X_n$ be independent real valued *rv*'s with common distribution function (*df*) $F(x)$, and let $F_n(x)$ be the corresponding empirical *df*, i.e., $nF_n(x) =$ number of $X_i \leqslant x$, $1 \leqslant i \leqslant n$. Let $\bar{X} = \sum_{i=1}^n X_i/n$, and consider the statistic

$$T_n = F_n(\bar{X}) \tag{1.1}$$

which represents the proportion of the sample below the sample mean. Such a statistic is often used in estimating a functional $\theta = F(\mu)$, where $\mu = EX_1$ if both F and μ are unknown or in testing the hypothesis that F is symmetric about an unknown location μ against certain classes of alternatives (see Gastwirth (1971)). The asymptotic normality of T_n was first derived by David (1962) under the assumption that F is normal. Later Ghosh (1971) derived this result under weaker assumptions that $0 < \operatorname{Var} X_1 < \infty$ and F is differentiable at μ with $0 < F'(\mu) < \infty$. (See also Sarkadi, Schnell and Vincze

Received July 26, 1982; revised November 10, 1982.

AMS 1980 subject classification: Primary 60F05, 60F15; secondary 62E20.

Key words and phrases: Berry–Esséen rates, law of iterated logarithm, invariance principle, Gâteaux-differential.

* Research supported by the National Science Foundation Grant MCS-8301409.

(1962) for the connection between the limit law of T_n and the occupancy problem).

Quite often, however, one needs more precise information than the asymptotic normality can provide. On the one hand, in view of applications, one may need bounds on the rate of convergence of T_n to normality, and, on the other hand, one may be interested in deriving the rates of strong convergence of T_n to $F(\mu)$ or certain invariance principles for T_n and incorporate them in the study of the asymptotic properties of the procedures (testing and estimation etc.) based on this statistic. The present note addresses these problems. Under different assumptions on F we derive (i) the Berry–Esséen rate $O(n^{-1/2})$ for the convergence of T_n to normality in the nonregular cases (i.e., when $F'(\mu)=0$), (ii) a law of the iterated logarithm, and (iii) an invariance principle for T_n. We also obtain a bound on the rate at which T_n converges to normality in the general case when $F'(\mu)$ does not necessarily vanish.

2. The Berry–Esséen Theorem for T_n in a Nonregular Case

Here we consider the question: What happens to the asymptotic law of T_n when $F'(\mu)=0$? Note that in such a case the asymptotic variance of the modified sign test equals that of the regular sign test (Gastwirth (1971)). Ghosh's (1971) method fails when $F'(\mu)=0$, while Gastwirth (1971) provides a heuristic argument. However, it will become clear from our Lemma 4.1 that if one is just interested in the asymptotic law of T_n under the assumptions that $F'(\mu)=0$, $0<\operatorname{Var} X_1<\infty$ and $0<F(\mu)<1$, then one may derive the representation

$$F_n(\bar{X})=F_n(\mu)+R_n \tag{2.1}$$

and show that $n^{1/2}R_n \to^P 0$. The asymptotic normality of T_n then follows immediately.

To motivate our study, consider the following example (cf. Chandra (1975)). Let

$$\begin{aligned} F(x) &= 0 && \text{if} \quad x\leqslant \mu-1 \\ &= \frac{(x-\mu)^{2m+1}+1}{2} && \text{if} \quad |x-\mu|\leqslant 1 \\ &= 1 && \text{if} \quad x\geqslant \mu+1 \end{aligned} \tag{2.2}$$