The Annals of Statistics
1983, Vol. 11, No. 2, 607–624

AN EDGEWORTH EXPANSION FOR SIMPLE LINEAR RANK STATISTICS UNDER THE NULL-HYPOTHESIS

BY RONALD J. M. M. DOES

Mathematical Centre, Amsterdam and University of Limburg, Maastricht

An Edgeworth expansion with remainder $o(N^{-1})$ is established for simple linear rank statistics under the null-hypothesis. The theorem is proved for a wide class of scores generating functions which includes the normal quantile function.

1. Introduction. Let $X_1, X_2, \cdots, X_N$ be independent and identically distributed random variables with a common continuous distribution function F. If $X_{1:N} < X_{2:N} < \cdots < X_{N:N}$ denotes the sequence $X_1, X_2, \cdots, X_N$ arranged in increasing order, then the rank R_{jN} of X_j is defined by $X_j = X_{R_{jN}:N}$ and the antirank D_{jN} is defined by $X_{D_{jN}} = X_{j:N}$, $j = 1, 2, \cdots, N$. We consider the simple linear rank statistic

$$T_N = \sum_{j=1}^N c_{jN} J\left(\frac{R_{jN}}{N+1}\right) = \sum_{j=1}^N c_{D_{jN}N} J\left(\frac{j}{N+1}\right), \tag{1.1}$$

where $c_{1N}, c_{2N}, \cdots, c_{NN}$, $N = 1, 2, \cdots$, is a triangular array of regression constants and J is a scores generating function defined on (0, 1). The two-sample linear rank statistic is obviously obtained as a special case by setting $c_{jN} = 0$ for $j = 1, 2, \cdots, n$, $c_{jN} = 1$ for $j = n+1, \cdots, N$. If $c_{jN} = j$ for $j = 1, 2, \cdots, N$ and $J(t) = t$ for $t \in (0, 1)$ then the statistic T_N is distributed as Spearman's rank correlation coefficient ρ under the null-hypothesis of independence.

The statistic T_N may be used for testing the null-hypothesis that all observations are independent and identically distributed against classes of alternatives indicated by the choice of regression constants and scores generating function. Both under the hypothesis and under contiguous and fixed alternatives it was shown that T_N is asymptotically normally distributed under very general conditions; cf. Hájek and Šidák (1967, Chapters V and VI), Hájek (1968) and Dupač and Hájek (1969). More recently a number of authors have studied the rate of convergence in these limit theorems. Berry-Esseen type bounds of order $\mathcal{O}(N^{-1/2})$ for simple linear rank statistics were established by Hušková (1977, 1979), Ho and Chen (1978) and Does (1982a). The purpose of this paper is to establish an Edgeworth expansion for simple linear rank statistics under the hypothesis with remainder $o(N^{-1})$ for a wide class of scores generating functions including the normal quantile function. We note that for the special case of the two-sample linear rank statistic, asymptotic expansions both under the hypothesis and under contiguous alternatives were obtained in Bickel and Van Zwet (1978). Asymptotic expansions for the simple linear rank statistics under contiguous alternatives are established in the author's Ph.D. thesis; cf. Does (1982b).

In Section 2 we formulate our theorem. Section 3 contains a number of preliminaries. The proof of the theorem is contained in Section 4. In Section 5 we compare our results with those in Bickel and Van Zwet (1978) for the two-sample linear rank statistic. Finally in the last section we discuss briefly the numerical aspects of our expansions. In the sequel we suppress the index N whenever it is possible.

Received February 1982; revised October 1982.

AMS 1980 *subject classifications.* Primary 62G10, 62G20; secondary 60F05.

Key words and phrases. Simple linear rank statistics, Edgeworth expansions, distributionfree tests.

2. An Edgeworth Expansion. Throughout this paper we make the following assumptions.

ASSUMPTION A. The regression constants $c_{1N}, c_{2N}, \cdots, c_{NN}$ satisfy

$$\sum_{j=1}^{N} c_{jN} = 0, \quad \sum_{j=1}^{N} c_{jN}^2 = 1, \quad \max_{1\le j\le N} |c_{jN}| = \mathcal{O}(N^{-1/2}).$$

This assumption implies that $ET_N = 0$.

ASSUMPTION B. The scores generating function J is three times differentiable on (0, 1) and

$$\limsup_{t\to 0,1} t(1-t)\left|\frac{J''(t)}{J'(t)}\right| < 2; \tag{2.1}$$

there exist positive numbers $\Gamma > 0$ and $\alpha < 3 + 1/14$ such that the third derivative J''' satisfies

$$|J'''(t)| \le \Gamma\{t(1-t)\}^{-\alpha} \quad \text{for } t \in (0, 1). \tag{2.2}$$

Furthermore

$$\int_0^1 J(t)\, dt = 0, \quad \int_0^1 J^2(t)\, dt = 1. \tag{2.3}$$

We note that (2.1) ensures that the function J' does not oscillate too wildly near 0 and 1; see also Appendix 2 of Albers, Bickel and Van Zwet (1976). Condition (2.3) can be assumed without loss of generality.

Taking

$$\bar{J} = \frac{1}{N}\sum_{j=1}^{N} J\left(\frac{j}{N+1}\right), \tag{2.4}$$

we know that the variance σ_N^2 of T_N (cf. (1.1)) is given by

$$\sigma_N^2 = \sigma^2(T_N) = \frac{1}{N-1}\sum_{j=1}^{N}\left(J\left(\frac{j}{N+1}\right) - \bar{J}\right)^2; \tag{2.5}$$

see e.g. Theorem II 3.1.c of Hájek and Šidák (1967). Define for each $N \ge 2$

$$T_N^* = \sigma_N^{-1} T_N \tag{2.6}$$

and

$$F_N^*(x) = P(T_N^* \le x) \quad \text{for} \quad -\infty < x < \infty. \tag{2.7}$$

Furthermore define for each $N \ge 2$ and real x, the function $\tilde{F}_N$ *by*

$$\tilde{F}_N(x) = \Phi(x) - \phi(x)\left\{\frac{\kappa_{3N}}{6}(x^2-1) + \frac{\kappa_{4N}}{24}(x^3-3x) + \frac{\kappa_{3N}^2}{72}(x^5 - 10x^3 + 15x)\right\}, \tag{2.8}$$

where Φ denotes the standard normal distribution function, ϕ its density and where the quantities κ_{3N} and κ_{4N} are given by

$$\kappa_{3N} = \sum_{j=1}^{N} c_{jN}^3\left\{\int_0^1 J^3(t)\, dt\right\} \tag{2.9}$$

and

$$\kappa_{4N} = \sum_{j=1}^{N} c_{jN}^4\left\{\int_0^1 J^4(t)\, dt - 3\right\} - \frac{3}{N}\left\{\int_0^1 J^4(t)\, dt - 1\right\}. \tag{2.10}$$

Our theorem reads as follows.

THEOREM 2.1. *If the Assumptions* A *and* B *are satisfied, then as* $N \to \infty$

$$\sup_{x\in\mathbb{R}} |F_N^*(x) - \tilde{F}_N(x)| = o(N^{-1}). \tag{2.11}$$

We note that κ_{3N} and κ_{4N} (cf. (2.9) and (2.10)) are asymptotic expressions for the third and fourth cumulants of T_N^* where terms of order $o(N^{-1})$ have been neglected. Hence $\tilde{F}_N$ may be said to constitute a genuine Edgeworth expansion for F_N^*. We should also point out that Theorem 2.1 allows scores generating functions tending to infinity in the neighbourhood of 0 and 1 at the rate of $\{t(1-t)\}^{-1/14+\varepsilon}$ for some $\varepsilon > 0$. It is clear that this includes the normal quantile function. Whenever we shall suppose in the remainder of this paper that (2.2) in Assumption B is satisfied, we shall tacitly and without loss of generality assume that $\alpha \in (3, 3 + 1/14)$ and define $\delta = 3 + 1/14 - \alpha$. Hence, from now on we replace (2.2) in Assumption B by

$$|J'''(t)| \le \Gamma\{t(1-t)\}^{-(3+1/14)+\delta} \quad \text{for} \quad t \in (0, 1), \tag{2.12}$$

where

$$0 < \delta < 1/14. \tag{2.13}$$

To conclude this section we define $U_1, U_2, \cdots, U_N$ to be independent and uniformly distributed random variables on (0, 1) and $U_{1:N} < U_{2:N} < \cdots < U_{N:N}$ the corresponding uniform order statistics.

3. Preliminary lemmas. The aim of this section is threefold. In the first place we approximate $(N-1)\sigma_N^2$ (cf. (2.5)) by an integral. Secondly we study the behaviour of the characteristic function of T_N^* (cf. (2.6)) for large values of the argument. Finally we prove two technical lemmas, the purpose of which will become clear in Section 4.

LEMMA 3.1. *If J satisfies Assumption B, then*

$$\sum_{j=1}^N \left\{J\left(\frac{j}{N+1}\right) - \bar{J}\right\}^2 = N + \mathcal{O}(N^{1/7-2\delta}). \tag{3.1}$$

PROOF. Take δ as in (2.12) and (2.13) and let h be a function on (0, 1) with $h'(t) \equiv \Gamma\{t(1-t)\}^{-15/14+\delta}$. With this in mind we can proceed exactly as in Lemma 3.1 of Does (1982a) to obtain (3.1). See also Lemma 3.1 in Does (1981). □

We now consider the behaviour of the characteristic function of T_N^* for large values of the argument. Let

$$\psi_N(t) = Ee^{itT_N^*}. \tag{3.2}$$

LEMMA 3.2. *Suppose that the assumptions of Theorem* 2.1 *are satisfied. Then there exist positive numbers B, β and γ such that*

$$|\psi_N(t)| \le BN^{-\beta\log N} \tag{3.3}$$

for $\log N \le |t| \le \gamma N^{3/2}$ *and* $N = 2, 3, \cdots$.

PROOF. The present lemma is a special case of Theorem 2.1 of Van Zwet (1982). Since we are concerned with independent and identically distributed random variables $X_1, X_2, \cdots, X_N$—which we may assume to be uniformly distributed without loss of generality—Condition (2.7) of this theorem is clearly satisfied. Moreover, the assumptions of our theorem guarantee that there exists a positive fraction of the scores which are at a distance of at least $N^{-3/2}\log N$ apart from each other, so Assumption (2.6) of Theorem 2.1 of Van Zwet (1982) is also fulfilled. Finally, it follows from Section 3 in Van Zwet (1982) that the existence of positive numbers c and C such that

$$\sum_{j=1}^{N} c_j^2 \geq c, \quad \sum_{j=1}^{N} c_j^4 \leq CN^{-1}, \tag{3.4}$$

$$\sum_{j=1}^{N} \left\{J\left(\frac{j}{N+1}\right) - \bar{J}\right\}^2 \geq cN, \quad \sum_{j=1}^{N} \left\{J\left(\frac{j}{N+1}\right) - \bar{J}\right\}^4 \leq CN \tag{3.5}$$

suffices to prove the present lemma. Assumption A guarantees the validity of (3.4), and (3.5) is a consequence of Assumption B (cf. also (3.1)). □

Let $[x]$ denote the largest integer not exceeding x. Define $m = [N^{8/15}]$ and $I = \{1, 2, \cdots, m, N-m+1, \cdots, N-1, N\}$.

LEMMA 3.3. *If Assumptions* A *and* B *are satisfied, then*

$$E\left| \sum_{j\in I} c_{D_j} J\left(\frac{j}{N+1}\right)\right|^5 = \mathcal{O}(N^{-1-7\delta/3}), \tag{3.6}$$

$$\left\{\frac{1}{N-2m} \sum_{j=m+1}^{N-m} J\left(\frac{j}{N+1}\right)\right\}^2 E\left(\sum_{j=m+1}^{N-m} c_{D_j}\right)^2 = \mathcal{O}(N^{-4/3-14\delta/15}). \tag{3.7}$$

PROOF. According to Assumption A $\sum c_j = 0$, $\sum c_j^2 = 1$ and

$$\sum_{j=1}^{N} |c_j|^k \leq \max_{1\leq j\leq N} |c_j|^{k-2} \sum_{j=1}^{N} c_j^2 = \mathcal{O}(N^{1-k/2}),$$

for $k > 2$. It follows that for distinct $i, j, h, g, k, l \in I$

$$Ec_{D_i}^6 = \mathcal{O}(N^{-3}), \qquad Ec_{D_i}^5 c_{D_j} = \mathcal{O}(N^{-4}), \qquad Ec_{D_i}^4 c_{D_j}^2 = \mathcal{O}(N^{-3}),$$

$$Ec_{D_i}^3 c_{D_j}^3 = \mathcal{O}(N^{-3}), \qquad Ec_{D_i}^4 c_{D_j} c_{D_h} = \mathcal{O}(N^{-4}), \qquad Ec_{D_i}^3 c_{D_j}^2 c_{D_h} = \mathcal{O}(N^{-4}),$$

$$Ec_{D_i}^2 c_{D_j}^2 c_{D_h}^2 = \mathcal{O}(N^{-3}), \qquad Ec_{D_i}^3 c_{D_j} c_{D_h} c_{D_g} = \mathcal{O}(N^{-5}), \qquad Ec_{D_i}^2 c_{D_j}^2 c_{D_h} c_{D_g} = \mathcal{O}(N^{-4}),$$

$$Ec_{D_i}^2 c_{D_j} c_{D_h} c_{D_g} c_{D_k} = \mathcal{O}(N^{-5}), \qquad Ec_{D_i} c_{D_j} c_{D_h} c_{D_g} c_{D_k} c_{D_l} = \mathcal{O}(N^{-6}).$$

Furthermore, Hölder's inequality yields

$$E\left| \sum_{j\in I} c_{D_j} J\left(\frac{j}{N+1}\right)\right|^5 \leq \left\{E\left(\sum_{j\in I} c_{D_j} J\left(\frac{j}{N+1}\right)\right)^6\right\}^{5/6}. \tag{3.8}$$

In view of (2.12) and (2.13) we have for $k = 1, 2, \cdots, 6$

$$\frac{1}{N}\sum_{j\in I} \left|J\left(\frac{j}{N+1}\right)\right|^k = \mathcal{O}\left(\int_0^{\frac{m}{N+1}} \{t(1-t)\}^{-k/14+k\delta} dt\right) = \mathcal{O}\left(\left\{\frac{m}{N+1}\right\}^{1-k/14+k\delta}\right). \tag{3.9}$$

Direct computation of the right-hand side of (3.8) produces (3.6). Since $\sum c_j = 0$, $Ec_{D_j}^2 = N^{-1}$ and $Ec_{D_i} c_{D_j} = -\{N(N-1)\}^{-1}$ for $i \neq j$, we have

$$E\left(\sum_{j=m+1}^{N-m} c_{D_j}\right)^2 = E\left(\sum_{j\in I} c_{D_j}\right)^2 = \mathcal{O}\left(\frac{m}{N}\right).$$

Since $\int J = 0$ and $\delta \in (0, 1/14)$ (cf. (2.3) and (2.13)) and J satisfies (2.1), we have in view of (A.2.11) in Albers, Bickel and Van Zwet (1976)

$$\begin{aligned} \left|\frac{1}{N}\sum_{j=m+1}^{N-m} J\left(\frac{j}{N+1}\right)\right| &= \left|\frac{1}{N}\sum_{j\in I} J\left(\frac{j}{N+1}\right)\right| + \left|\frac{1}{N}\sum_{j=1}^{N} \left\{J\left(\frac{j}{N+1}\right) - EJ(U_{j:N})\right\}\right| \\ &= \mathcal{O}\left(\left\{\frac{m}{N}\right\}^{13/14+\delta}\right) + \mathcal{O}(N^{-13/14-\delta}) = \mathcal{O}(N^{-13/30-7\delta/15}) \end{aligned} \tag{3.10}$$

and the lemma follows. □

To conclude this section we prove:

LEMMA 3.4. *If Assumption* A *is satisfied, then for any* $\gamma < 1$ *and* $N \to \infty$

$$P\left(\sum_{j\in I} c_{D_j}^2 \geq 1 - \gamma\right) = \mathcal{O}(N^{-22/15}). \tag{3.11}$$

PROOF. Since $E(\sum_{j\in I} c_{D_j}^2) = 2mN^{-1}$ and

$$E\left(\sum_{j\in I} c_{D_j}^2 - \frac{2m}{N}\right)^2 = \frac{2m(N-2m)}{N(N-1)}\left(\sum_{j=1}^N c_j^4 - \frac{1}{N}\right),$$

the Bienaymé-Chebyshev inequality ensures that for every $\gamma < 1$

$$P\left(\left|\sum_{j\in I} c_{D_j}^2 - \frac{2m}{N}\right| \geq \frac{1-\gamma}{2}\right) \leq \frac{4}{(1-\gamma)^2} E\left(\sum_{j\in I} c_{D_j}^2 - \frac{2m}{N}\right)^2 = \mathcal{O}(N^{-22/15}).$$

The lemma follows because $mN^{-1} \to 0$ as $N \to \infty$. □

4. Proof of the theorem. To prove Theorem 2.1 we start with an application of Esseen's smoothing lemma (see e.g., Feller, 1971, page 538), which implies that for all $\gamma > 0$

$$\sup_{x\in\mathbb{R}} |F_N^*(x) - \tilde{F}_N(x)| \leq \frac{1}{\pi}\int_{-\gamma N^{3/2}}^{\gamma N^{3/2}} \frac{|\psi_N(t) - \lambda_N(t)|}{|t|}\,dt + \mathcal{O}(N^{-3/2}), \tag{4.1}$$

where ψ_N denotes the characteristic function of T_N^* (cf. (3.2)) and λ_N denotes the Fourier-Stieltjes transform of $\tilde{F}_N$, i.e.

$$\lambda_N(t) = \int_{-\infty}^{\infty} e^{itx}\,d\tilde{F}_N(x) = e^{-t^2/2}\left\{1 - \frac{\kappa_{3N}}{6}it^3 + \frac{\kappa_{4N}}{24}t^4 - \frac{\kappa_{3N}^2}{72}t^6\right\}. \tag{4.2}$$

The derivative of λ_N is uniformly bounded and also

$$\left|\frac{d\psi_N(t)}{dt}\right| \leq E|T_N^*| \leq 1.$$

Because $\psi_N(0) = \lambda_N(0) = 1$, we have

$$\int_{|t|\leq N^{-3/2}} \frac{|\psi_N(t) - \lambda_N(t)|}{|t|}\,dt = \mathcal{O}(N^{-3/2}). \tag{4.3}$$

Similarly, Lemma 3.2 and (4.2) ensure that

$$\int_{\log N\leq|t|\leq\gamma N^{3/2}} \frac{|\psi_N(t) - \lambda_N(t)|}{|t|}\,dt = \mathcal{O}(N^{-3/2}). \tag{4.4}$$

From (4.1), (4.3) and (4.4) it follows that, in order to prove Theorem 2.1, it suffices to show that

$$\int_{t\in A} \frac{|\psi_N(t) - \lambda_N(t)|}{|t|}\,dt = \mathcal{O}(N^{-1}), \tag{4.5}$$

where $A = \{t : N^{-3/2} \leq |t| \leq \log N\}$.

To solve this problem we use a conditioning argument. We take δ as in (2.12) and (2.13) and define $m = [N^{8/15}]$ and $I = \{1, 2, \ldots, m, N-m+1, \ldots, N-1, N\}$ as in Section 3. Let $\Omega = \{D_j : j \in I\}$ be the set of antiranks D_j with indices in I and let $\omega = \{d_j : j \in I\}$ be a possible realization of Ω. Finally define

$$Z_N = \sum_{j \in I} c_{D_j} J\left(\frac{j}{N+1}\right). \tag{4.6}$$

Because $(T_N - Z_N)$ and Z_N are conditionally independent given Ω, we have

$$\begin{aligned} \psi_N(t) &= Ee^{itT_N^*} = E[E(e^{it\sigma_N^{-1}(T_N - Z_N)} \mid \Omega) E(e^{it\sigma_N^{-1}Z_N} \mid \Omega)] \\ &= E[E(e^{it\sigma_N^{-1}\{(T_N - Z_N) - E(T_N - Z_N \mid \Omega)\}} \mid \Omega) e^{it\sigma_N^{-1}E(T_N - Z_N \mid \Omega)} E(e^{it\sigma_N^{-1}Z_N} \mid \Omega)]. \end{aligned} \tag{4.7}$$

We note that conditionally on $\Omega = \omega$, $T_N - Z_N = \sum_{j=m+1}^{N-m} c_{D_j} J(j/(N+1))$ is distributed as a simple linear rank statistic for sample size $N - 2m$ based on a set of regression constants $\{c_1, c_2, \cdots, c_N\} \setminus \{c_{d_j} : j \in I\}$ and having a scores generating function

$$J_N(t) = J\left(\frac{m + (N - 2m + 1)t}{N+1}\right), \qquad t \in (0, 1). \tag{4.8}$$

We write this simple linear rank statistic as

$$T_{\omega N} = \sum_{j=1}^{M} b_j J_N\left(\frac{Qj}{M+1}\right), \tag{4.9}$$

where $M = N - 2m$, $\{b_1, b_2, \cdots, b_M\} = \{c_1, c_2, \cdots, c_N\} \setminus \{c_{d_j} : j \in I\}$, $Q_1, Q_2, \cdots Q_M$ are the ranks of $V_1, V_2, \cdots, V_M$, which are independent and uniformly distributed random variables on (0, 1).

Define for $j = 1, 2, \cdots, M$

$$\hat{V}_j = E\left(\frac{Q_j}{M+1} \,\middle|\, V_j\right) = \frac{1}{M+1} + \frac{M-1}{M+1} V_j \tag{4.10}$$

and let $S_{\omega N}$ be a three-term Taylor expansion of $T_{\omega N}$, viz.

$$S_{\omega N} = \sum_{j=1}^{M} b_j \left\{ J_N(\hat{V}_j) + J_N'(\hat{V}_j)\left(\frac{Q_j}{M+1} - \hat{V}_j\right) + \frac{1}{2} J_N''(\hat{V}_j)\left(\frac{Q_j}{M+1} - \hat{V}_j\right)^2 \right\}. \tag{4.11}$$

Our plan of attack of (4.7) is as follows. We expand $E(\exp\{it\sigma_N^{-1}Z_N\} \mid \Omega)$ and control the remainder term by bounding $E\,|Z_N|^5$. To achieve this, we clearly cannot have m tending to infinity too rapidly (cf. (3.6) in Lemma 3.2). On the other hand we approximate $(T_{\omega N} - ET_{\omega N})$ by $(S_{\omega N} - ES_{\omega N})$ which involves bounding J''' on the interval $(m/(N+1), 1 - m/(N+1))$. By (2.2), J''' may tend to infinity near 0 and 1 at a rate depending on α and hence we can't allow m to tend to infinity too slowly. For $\alpha < 3 + 1/14$ as in Assumption B, both demands on m can be reconciled and the resulting choice for m is $[N^{8/15}]$.

For approximating $(T_{\omega N} - ET_{\omega N})$ by $(S_{\omega N} - ES_{\omega N})$ we need:

LEMMA 4.1. *Under the Assumptions* A *and* B *we have, uniformly in* ω,

$$\sigma^2(T_{\omega N} - S_{\omega N}) = \{1 + (\textstyle\sum_{j \in I} c_{d_j})^2\} \mathcal{O}(N^{-2-14\delta/15}). \tag{4.12}$$

PROOF. Let, for $j = 1, 2, \cdots, M$,

$$Y_j = J_N\left(\frac{Q_j}{M+1}\right) - \left\{ J_N(\hat{V}_j) + J_N'(\hat{V}_j)\left(\frac{Q_j}{M+1} - \hat{V}_j\right) + \frac{1}{2} J_N''(\hat{V}_j)\left(\frac{Q_j}{M+1} - \hat{V}_j\right)^2 \right\}.$$

Because $\sum_{j=1}^{M} b_j^2 \le 1$ and

$$\Big|\sum\sum_{j \ne k} b_j b_k\Big| = \Big|(\textstyle\sum_{j=1}^{M} b_j)^2 - \sum_{j=1}^{M} b_j^2\Big| \le 1 + (\sum_{j \in I} c_{d_j})^2,$$

the Cauchy-Schwarz inequality yields

$$\sigma^2(T_{\omega N} - S_{\omega N}) \le E(T_{\omega N} - S_{\omega N})^2 = E(\sum_{j=1}^{M} b_j Y_j)^2$$
$$= \sum_{j=1}^{M} b_j^2 E Y_1^2 + \sum\sum_{j \ne k} b_j b_k E Y_1 Y_2 \le (2 + (\sum_{j \in I} c_{d_j})^2) E Y_1^2.$$

Here $\sum\sum_{j\ne k}$ denotes summation over all non-negative distinct integers j, k satisfying $1 \le j$, $k \le M$. Define $r(t) = \{t(1-t)\}^{-1}$. By Taylor's theorem, (4.8), (2.12) and the convexity of the function $r(t)$ we see that

$$EY_1^2 \le \frac{1}{36} E\left(\frac{Q_1}{M+1} - \hat{V}_1\right)^6 \sup_{0\le\eta\le1}\left\{J'''_N\left(\eta\frac{Q_1}{M+1} + (1-\eta)\hat{V}_1\right)\right\}^2$$
$$\le \frac{\Gamma^2}{36} E\left(\frac{Q_1}{M+1} - \hat{V}_1\right)^6\left\{r^{6+1/7-2\delta}\left(\frac{m+Q_1}{N+1}\right) + r^{6+1/7-2\delta}\left(\frac{m+(M+1)\hat{V}_1}{N+1}\right)\right\}.$$

The independence of the vector of ranks $(Q_1, Q_2, \cdots, Q_M)$ and the vector of order statistics $(V_{1:M}, V_{2:M}, \cdots, V_{M:M})$ and Lemma A.2.3 of Albers, Bickel and Van Zwet (1976) imply that

$$E\left(\frac{Q_1}{M+1} - \hat{V}_1\right)^6 r^{6+1/7-2\delta}\left(\frac{m+Q_1}{M+1}\right)$$
$$= \left(\frac{M-1}{M+1}\right)^6 E\left(\frac{Q_1-1}{M-1} - V_1\right)^6 r^{6+1/7-2\delta}\left(\frac{m+Q_1}{N+1}\right)$$
$$\le \frac{1}{M}\sum_{j=1}^{M} E\left(V_{j:M} - \frac{j-1}{M-1}\right)^6 r^{6+1/7-2\delta}\left(\frac{m+j}{N+1}\right) \tag{4.13}$$
$$= \mathcal{O}\left(\frac{1}{M^4}\sum_{j=1}^{M} r^{-3}\left(\frac{j}{M+1}\right) r^{6+1/7-2\delta}\left(\frac{m+j}{N+1}\right)\right) = \mathcal{O}(N^{-2-14\delta/15}).$$

Furthermore, the conditional distribution of $Q_1 - 1$ given V_1 is binomial with parameters $M-1$ and V_1 and by application of a recursion formula for the central moments of this distribution (cf. Johnson and Kotz, 1969, page 52) we find

$$E(\{Q_1 - E(Q_1 \mid V_1)\}^6 \mid V_1) = \mathcal{O}(\{MV_1(1-V_1)\}^3 + MV_1(1-V_1)).$$

Hence,

$$E\left(\frac{Q_1}{M+1} - \hat{V}_1\right)^6 r^{6+1/7-2\delta}\left(\frac{m+(M+1)\hat{V}_1}{N+1}\right)$$
$$= \mathcal{O}\left(E\left(\left\{\frac{V_1(1-V_1)}{M}\right\}^3 + \frac{V_1(1-V_1)}{M^5}\right) r^{6+1/7-2\delta}\left(\frac{m+(M+1)\hat{V}_1}{N+1}\right)\right) \tag{4.14}$$
$$= \mathcal{O}(N^{-2-14\delta/15}).$$

Combining (4.13) and (4.14) we find that $EY_1^2 = \mathcal{O}(N^{-2-14\delta/15})$. This proves the lemma. □

It follows from Lemma 4.1, (2.5) and (3.8) that

$$|Ee^{it\sigma_N^{-1}(T_{\omega N} - ET_{\omega N})} - Ee^{it\sigma_N^{-1}(S_{\omega N} - ES_{\omega N})}| \le |t|\,\sigma_N^{-1} E\,|T_{\omega N} - ET_{\omega N} - S_{\omega N} + ES_{\omega N}|$$
$$= \mathcal{O}(|t|\, N^{-1-7\delta/15}\{1 + (\sum_{j\in I} c_{d_j})^2\}^{1/2}), \tag{4.15}$$

uniformly in t and ω.

Our next task is to evaluate $E\exp\{it\sigma_N^{-1}(S_{\omega N} - ES_{\omega N})\}$. The technique for doing this resembles that in Helmers (1980). Let χ be the indicator function of $(0, \infty)$ and define

$$(4.16)\quad \begin{aligned} S_1 &= \textstyle\sum_{j=1}^M b_j\{J_N(\hat V_j) - EJ_N(\hat V_j)\} = \sum_{j=1}^M b_j\tilde J_N(\hat V_j),\\ S_2 &= \frac{1}{M+1}\sum\sum_{j\neq k} b_j J_N'(\hat V_j)\{\chi(V_j - V_k) - V_j\},\\ S_3 &= \frac{1}{2(M+1)^2}\sum\sum_{j\neq k} b_j[J_N''(\hat V_j)\{\chi(V_j - V_k) - V_j\}^2 - EJ_N''(\hat V_j)\{\chi(V_j - V_k) - V_j\}^2],\\ S_4 &= \frac{1}{2(M+1)^2}\sum\sum\sum_{j\neq k\neq l} b_j J_N''(\hat V_j)\{\chi(V_j - V_k) - V_j\}\{\chi(V_j - V_l) - V_j\}. \end{aligned}$$

It is easy to see that $S_{\omega N} - ES_{\omega N} = \sum_{\nu=1}^4 S_\nu$ and $ES_\nu = 0$ for $\nu = 1, \cdots, 4$. First of all we compute a number of moments.

LEMMA 4.2. *Under the Assumptions* A *and* B *we have, uniformly in* ω,

$$(4.17)\quad E\,|\,S_2\,|^3 = \mathcal{O}(N^{-13/10-7\delta/5}),\quad ES_3^2 = \mathcal{O}(N^{-22/15-14\delta/15}),\quad ES_4^2 = \mathcal{O}(N^{-7/5-14\delta/15}).$$

PROOF. By applying Hölder's inequality we obtain $E\,|\,S_2\,|^3 \le (ES_2^4)^{3/4}$. Let, for distinct j and k, $h(V_j, V_k) = J_N'(\hat V_j)\{\chi(V_j - V_k) - V_j\}$. Define $h(x, x) = 0$ for all $0 < x < 1$. Direct computation of ES_2^4 shows that

$$(4.18)\quad \begin{aligned} ES_2^4 = \frac{1}{(M+1)^4}\Big[&\textstyle\sum_{j=1}^M b_j^4\{\sum_{r=1}^M\sum_{s=1}^M\sum_{t=1}^M\sum_{u=1}^M Eh(V_1, V_r)h(V_1, V_s)h(V_1, V_t)\\ &\quad\cdot h(V_1, V_u)\}\\ &+ 4\sum\sum_{j\neq k} b_j^3 b_k\{\textstyle\sum_{r=1}^M\sum_{s=1}^M\sum_{t=1}^M\sum_{u=1}^M Eh(V_1, V_r)h(V_1, V_s)h(V_1, V_t)\\ &\quad\cdot h(V_2, V_u)\}\\ &+ 3\sum\sum_{j\neq k} b_j^2 b_k^2\{\textstyle\sum_{r=1}^M\sum_{s=1}^M\sum_{t=1}^M\sum_{u=1}^M Eh(V_1, V_r)h(V_1, V_s)h(V_2, V_t)\\ &\quad\cdot h(V_2, V_u)\}\\ &+ 6\sum\sum\sum_{j\neq k\neq l} b_j^2 b_k b_l\{\textstyle\sum_{r=1}^M\sum_{s=1}^M\sum_{t=1}^M\sum_{u=1}^M Eh(V_1, V_r)h(V_1, V_s)h(V_2, V_t)\\ &\quad\cdot h(V_3, V_u)\}\\ &+ \sum\sum\sum\sum_{j\neq k\neq l\neq n} b_j b_k b_l b_n\{\textstyle\sum_{r=1}^M\sum_{s=1}^M\sum_{t=1}^M\sum_{u=1}^M Eh(V_1, V_r)h(V_2, V_s)h(V_3, V_t)\\ &\quad\cdot h(V_4, V_u)\}\Big]. \end{aligned}$$

To bound the right-hand side of (4.18) we note that an expectation in (4.18) equals zero if at least one of the indices (r, s, t, u) occurs only once. With the aid of the Cauchy-Schwarz inequality, the non-zero expectations may be bounded by either $Eh^4(V_1, V_2)$ or $\{Eh^4(V_1, V_2)\}^{1/2}Eh^2(V_1, V_2)$ or $\{Eh^2(V_1, V_2)\}^2$. In view of (4.8) and Assumption B we can prove that for $1 \le k \le 4$,

$$(4.19)\quad E\,|\,h^k(V_1, V_2)| = \mathcal{O}(N^{k/2-14/15-7k\delta/15}).$$

According to Assumption A and the fact that $\{b_1, b_2, \cdots, b_M\} = \{c_1, c_2, \cdots, c_N\}\backslash\{c_{d_j}: j \in I\}$, we have

$$|\,\textstyle\sum_{j=1}^M b_j\,| = |\,\sum_{j\in I} c_{d_j}\,| = \mathcal{O}(mN^{-1/2}) = \mathcal{O}(N^{1/30})$$

and similarly

$$\sum_{j=1}^{M} b_j^4 = \mathcal{O}(N^{-1}), \qquad \Big|\sum_{j\neq k}\sum b_j^3 b_k\Big| = \mathcal{O}(N^{-7/15}),$$

$$\sum_{j\neq k}\sum b_j^2 b_k^2 = 1 + \mathcal{O}(N^{-7/15}), \quad \Big|\sum_{j\neq k\neq l}\sum\sum b_j^2 b_k b_l\Big| = \mathcal{O}(N^{1/15}), \tag{4.20}$$

$$\Big|\sum_{j\neq k\neq l\neq n}\sum\sum\sum b_j b_k b_l b_n\Big| = \mathcal{O}(N^{2/15}).$$

Combining these results we find that $ES_2^4 = \mathcal{O}(N^{-26/15-28\delta/15})$ and hence $E|S_2|^3 = \mathcal{O}(N^{-13/10-7\delta/5})$. In the same way one can obtain the other two assertions in (4.17). □

Define, for real t and $N \geq 2$,

$$\rho_N(t) = Ee^{it\sigma_N^{-1}(S_{\omega N}-ES_{\omega N})} \tag{4.21}$$

and

$$\rho_{1N}(t) = Ee^{it\sigma_N^{-1}S_1}\left\{1 + \frac{it}{\sigma_N}(S_2 + S_3 + S_4) + \frac{(it)^2}{2\sigma_N^2}S_2^2\right\}. \tag{4.22}$$

The next lemma shows that ρ_N can be approximated by ρ_{1N}.

LEMMA 4.3. *If the Assumptions* A *and* B *are satisfied, then*

$$|\rho_N(t) - \rho_{1N}(t)| = \mathcal{O}(t^2 N^{-17/15-14\delta/15}) \tag{4.23}$$

uniformly for $|t| \leq \log N$ *and* ω.

PROOF. Repeated use of Lemma XV 4.1 of Feller (1971) yields

$$|\rho_N(t) - \rho_{1N}(t)| = \mathcal{O}(t^2\sigma_N^{-2}E|S_2||S_3 + S_4| + t^2\sigma_N^{-2}E(S_3^2 + S_4^2) + |t|^3\sigma_N^{-3}E|S_2|^3).$$

From (2.5) and (3.5) it follows that for all sufficiently large N there exist positive numbers $\varepsilon_1 \leq \varepsilon_2$ such that $\varepsilon_1 \leq \sigma_N^2 \leq \varepsilon_2$. Lemma 4.2 produces the desired result. □

Clearly our next task is to evaluate the right-hand side of (4.22) and we start with the leading term. According to (4.16) $S_1 = \sum_{j=1}^{M} b_j\tilde{J}_N(\hat{V}_j)$. We have $E\tilde{J}_N(\hat{V}_1) = 0$ and for all sufficiently large N, there exist positive numbers $\gamma_1 \leq \gamma_2$ such that $\gamma_1 \leq E\tilde{J}_N^2(\hat{V}_1) \leq \gamma_2$ (cf. (2.3)). In the sequel we shall assume

$$\sum_{j\in I} c_{d_j}^2 < 1 - \gamma \tag{4.24}$$

for some $\gamma \in (0, 1)$, to guarantee that

$$\gamma\gamma_1 \leq \sigma^2(S_1) \leq \gamma_2. \tag{4.25}$$

Finally we note that Assumptions A and B imply that $\sum_{j=1}^{M} b_j^4 = \mathcal{O}(N^{-1})$ and that the random variable $\tilde{J}_N(\hat{V}_1)$ has a finite 14th absolute moment. It follows from the classical theory of Edgeworth expansions for sums of independent and non-identically distributed random variables (see, e.g., Lemma VI 4.11 in Petrov, 1975) that

$$\begin{aligned}\Bigg| E\exp\{itS_1/\sigma(S_1)\} - e^{-t^2/2}\Bigg\{1 - \frac{it^3}{6\sigma^3(S_1)}\sum_{j=1}^{M} b_j^3 E\tilde{J}_N^3(\hat{V}_1) + \frac{t^4}{24\sigma^4(S_1)}\\ \cdot\sum_{j=1}^{M} b_j^4\{E\tilde{J}_N^4(\hat{V}_1) - 3[E\tilde{J}_N^2(\hat{V}_1)]^2\} - \frac{t^6}{72\sigma^6(S_1)}\{\sum_{j=1}^{M} b_j^3 E\tilde{J}_N^3(\hat{V}_1)\}^2\Bigg\}\Bigg|\\ = o(N^{-1}(t^4 + |t|^9)e^{-t^2/2})\end{aligned} \tag{4.26}$$

uniformly for $|t| \leq \log N$ and ω for which (4.24) is satisfied. Replacing t by $t_N =$

$t\sigma(S_1)/\sigma_N$ and expanding $\exp\{-\frac{1}{2}t_N^2\}$ we find that uniformly for $|t| \le \log N$ and ω for which (4.24) is satisfied

$$\Big| Ee^{it\sigma_N^{-1}S_1} - e^{-t^2/2}\Big\{1 - \frac{it^3}{6\sigma_N^3}\sum_{j=1}^M b_j^3 E\tilde{J}_N^3(\hat{V}_1) + \frac{t^4}{24\sigma_N^4}\sum_{j=1}^M b_j^4\{E\tilde{J}_N^4(\hat{V}_1) - 3[E\tilde{J}_N^2(\hat{V}_1)]^2\}$$

$$- \frac{t^6}{72\sigma_N^6}\{\sum_{j=1}^M b_j^3 E\tilde{J}_N^3(\hat{V}_1)\}^2 + \frac{t^2}{2\sigma_N^2}(\sigma_N^2 - \sigma^2(S_1))$$

$$+ \frac{t^4}{8\sigma_N^4}(\sigma_N^2 - \sigma^2(S_1))^2 - \frac{it^5}{12\sigma_N^5}(\sigma_N^2 - \sigma^2(S_1))\sum_{j=1}^M b_j^3 E\tilde{J}_N^3(\hat{V}_1)\Big\}\Big| \tag{4.27}$$

$$= o(N^{-1}(t^4 + |t|^9)e^{-t^2/2}) + \mathcal{O}(|\sigma_N^2 - \sigma^2(S_1)|^3 |t| P_1(t)e^{-\theta t^2})$$

$$+ \mathcal{O}(N^{-1}|\sigma_N^2 - \sigma^2(S_1)||t| P_2(t)e^{-\theta t^2}),$$

where $0 < \theta < \frac{1}{2}$ and P_1 and P_2 are fixed polynomials.

We now turn to the remaining terms on the right in (4.22). Let

$$\mu_N(t) = Ee^{it\tilde{J}_N(\hat{V}_1)} \tag{4.28}$$

denote the characteristic function of $\tilde{J}_N(\hat{V}_1)$, so that

$$Ee^{it\sigma_N^{-1}S_1} = \prod_{j=1}^M \mu_N\left(\frac{b_j t}{\sigma_N}\right). \tag{4.29}$$

From the Assumptions A and B it follows by Taylor expansion that for distinct integers $l_1, \cdots, l_n$ where $1 \le n \le 4$

$$\prod_{\nu=1}^n \mu_N\left(\frac{b_{l_\nu} t}{\sigma_N}\right) = 1 - \frac{t^2}{2\sigma_N^2}\{\sum_{\nu=1}^n b_{l_\nu}^2\} E\tilde{J}_N^2(\hat{V}_1) + \mathcal{O}(N^{-3/2}|t|^3), \tag{4.30}$$

uniformly for $|t| \le \log N$ and ω for which (4.24) is satisfied.

In the last lemma we summarize the results we need.

LEMMA 4.4. *If the Assumptions* A *and* B *are satisfied then, uniformly for* $|t| \le \log N$ *and* ω *for which* (4.24) *is satisfied*

$$\Big| E(e^{it\sigma_N^{-1}S_1}S_2) - Ee^{it\sigma_N^{-1}S_1}\Big\{\frac{it}{\sigma_N}ES_1S_2 + \frac{(it)^2}{2\sigma_N^2}ES_1^2S_2 - \frac{(it)^3}{4N\sigma_N^3}[E\tilde{J}_N^4(\hat{V}_1) - \{E\tilde{J}_N^2(\hat{V}_1)\}^2]\Big\}\Big| = \mathcal{O}(N^{-1-\varepsilon}|t|P(t)e^{-\theta t^2}), \tag{4.31}$$

$$\Big| E(e^{it\sigma_N^{-1}S_1}S_3) - Ee^{it\sigma_N^{-1}S_1}\Big\{\frac{it}{\sigma_N}ES_1S_3\Big\}\Big| = \mathcal{O}(N^{-1-\varepsilon}|t|P(t)e^{-\theta t^2}), \tag{4.32}$$

$$|E(e^{it\sigma_N^{-1}S_1}S_4)| = \mathcal{O}(N^{-1-\varepsilon}|t|P(t)e^{-\theta t^2}), \tag{4.33}$$

$$\Big| E(e^{it\sigma_N^{-1}S_1}S_2^2) - Ee^{it\sigma_N^{-1}S_1}\Big\{ES_2^2 + \frac{it}{\sigma_N}ES_1S_2^2 + \frac{(it)^2}{4N\sigma_N^2}[E\tilde{J}_N^4(\hat{V}_1) - \{E\tilde{J}_N^2(\hat{V}_1)\}^2]\Big\}\Big| = \mathcal{O}(N^{-1-\varepsilon}|t|P(t)e^{-\theta t^2}), \tag{4.34}$$

where $0 < \theta < \frac{1}{2}$, $\varepsilon > 0$ *and* P *is a fixed polynomial.*

PROOF. The proofs of the statements (4.30) through (4.34) are highly technical and laborious. As they all proceed in essentially the same manner, we shall only give the basic ideas of the proof of the first statement; the interested reader is referred to Does

(1981) for more details. Applying Lemma XV 4.1 of Feller (1971) to $\exp\{it\sigma_N^{-1}(b_j J_N(V_j) + b_k \tilde{J}_N(\hat{V}_k))\}$, it follows after adding some zero expectations and some computations that

$$E\exp\{it\sigma_N^{-1}(b_j\tilde{J}_N(\hat{V}_j) + b_k\tilde{J}_N(\hat{V}_k))\}J_N'(\hat{V}_j)(\chi(V_j - V_k) - V_j)$$
$$= E[J_N'(\hat{V}_j)(\chi(V_j - V_k) - V_j)]\bigg[\frac{it}{\sigma_N}S_1 + \frac{(it)^2}{2\sigma_N^2}S_1^2 + \frac{(it)^3}{6\sigma_N^3}\{3b_j^2 b_k \tilde{J}_N^2(\hat{V}_j)\tilde{J}_N(\hat{V}_k)$$
$$+ 3b_j b_k^2 \tilde{J}_N(\hat{V}_j)\tilde{J}_N^2(\hat{V}_k) + b_k^3\tilde{J}_N^3(\hat{V}_k)\}\bigg] + \mathcal{O}(t^4\, E\,|\,J_N'(\hat{V}_j)(\chi(V_j - V_k) - V_j)|$$
$$\cdot\, \{b_j^4\tilde{J}_N^4(\hat{V}_j) + b_k^4\tilde{J}_N^4(\hat{V}_k)\}).$$

From (4.30) it follows that for distinct integers $1 \le j,\, k \le M$ and $|\,t\,| \le \log N$

$$\prod_{\ell\neq j,k}\mu_N\left(\frac{b_\ell t}{\sigma_N}\right) = Ee^{it\sigma_N^{-1}S_1}\left\{1 + \frac{t^2}{2\sigma_N^2}(b_j^2 + b_k^2)E\tilde{J}_N^2(\hat{V}_1) + \mathcal{O}(N^{-3/2}\,|\,t\,|^3)\right\},$$

uniformly for $|\,t\,| \le \log N$ and ω for which (4.24) is satisfied. Hence, combining these results with Assumption A, we find after some algebra

$$(4.35)\quad \begin{aligned} E(e^{it\sigma_N^{-1}S_1}S_2) &= [Ee^{it\sigma_N^{-1}S_1}]\bigg[\frac{it}{\sigma_N}ES_1S_2 + \frac{(it)^2}{2\sigma_N^2}ES_1^2\,S_2 \\ &+ \frac{(it)^3}{6\sigma_N^3}\sum\sum_{j\neq k}E[J_N'(\hat{V}_j)(\chi(V_j - V_k) - V_j)]\bigg[\frac{3b_j^3 b_k}{M+1}\tilde{J}_N^2(\hat{V}_j)\tilde{J}_N(\hat{V}_k) \\ &+ \frac{3b_j^2 b_k^2}{M+1}\tilde{J}_N(\hat{V}_j)\tilde{J}_N^2(\hat{V}_k) + \frac{b_j b_k^3}{M+1}\tilde{J}_N^3(\hat{V}_k)\bigg] \\ &+ \frac{(it)^3}{2\sigma_N^3}\sum\sum_{j\neq k}\frac{b_j b_k(b_j^2 + b_k^2)}{M+1}\{E\tilde{J}_N(\hat{V}_k)J_N'(\hat{V}_j)(\chi(V_j - V_k) - V_j)\}\{E\tilde{J}_N^2(\hat{V}_1)\}\bigg] \\ &+ \mathcal{O}(N^{-3/2}t^4e^{-\theta t^2}E\{|\,\tilde{J}_N(\hat{V}_1)| + \tilde{J}_N^4(\hat{V}_1)\}\{|\,J_N'(\hat{V}_1)| + |\,J_N'(\hat{V}_2)|\}), \end{aligned}$$

uniformly for $|\,t\,| \le \log N$ and ω for which (4.24) is satisfied. From Assumption B and (4.8) it follows that (see also (4.19))

$$(4.36)\quad \begin{aligned} E\,|\,\tilde{J}_N^2(\hat{V}_1)\tilde{J}_N(\hat{V}_2)J_N'(\hat{V}_1)\,| &= \mathcal{O}(N^{1/10-7\delta/5}); \\ E\,|\,\tilde{J}_N(\hat{V}_1)J_N'(\hat{V}_1)\,| &= \mathcal{O}(N^{1/15-14\delta/15}); \\ E\,|\,\tilde{J}_N^4(\hat{V}_1)J_N'(\hat{V}_1)\,| &= \mathcal{O}(N^{1/6-7\delta/3});\quad E\tilde{J}_N^2(\hat{V}_1) = \mathcal{O}(1); \\ E(\tilde{J}_N^4(\hat{V}_1) + |\,\tilde{J}_N^3(\hat{V}_1)| + |\,\tilde{J}_N(\hat{V}_1)|)|\,J_N'(\hat{V}_2)\,| &= \mathcal{O}(N^{1/30-7\delta/15}). \end{aligned}$$

Finally we obtain by partial integration

$$(4.37)\quad \begin{aligned} &E\tilde{J}_N(\hat{V}_1)\tilde{J}_N^2(\hat{V}_2)J_N'(\hat{V}_1)(\chi(V_1 - V_2) - V_1) \\ &\qquad = -\frac{1}{2}\left(\frac{M+1}{M-1}\right)^2 E\tilde{J}_N^4(\hat{V}_1) + \frac{1}{2}\left(\frac{M+1}{M-1}\right)^3\{E\tilde{J}_N^2(\hat{V}_1)\}^2. \end{aligned}$$

Combining (4.35) through (4.37) and (4.20) we arrive at (4.31). □

From Lemma 4.4 it follows that uniformly for $|\,t\,| \le \log N$ and ω for which (4.24) is satisfied (cf. (4.22)),

$$\rho_{1N}(t) = \{Ee^{it\sigma_N^{-1}S_1}\}\bigg\{1 + \frac{(it)^2}{2\sigma_N^2}[2ES_1S_2 + 2ES_1S_3 + ES_2^2] + \frac{(it)^3}{2\sigma_N^3}[ES_1^2S_2 + ES_1S_2^2]$$
$$- \frac{(it)^4}{8N\sigma_N^4}[E\tilde{J}_N^4(\hat{V}_1) - \{E\tilde{J}_N^2(\hat{V}_1)\}^2]\bigg\} + \mathcal{O}(N^{-1-\varepsilon}\,|\,t\,|\,P(t)e^{-\theta t^2}),$$

where $\varepsilon > 0$, $0 < \theta < ½$ and P is a fixed polynomial. Using (4.25), Lemmas 4.1 and 4.2, as well as the fact that $ES_1S_4 = 0$, we obtain

$$(4.38)\quad \begin{aligned} 2ES_1S_2 + 2ES_1S_3 + ES_2^2 &= \sigma^2(S_{\omega N}) - \sigma^2(S_1) + \mathcal{O}(N^{-17/15-14\delta/15}) \\ &= \sigma^2(T_{\omega N}) - \sigma^2(S_1) + (1 + (\textstyle\sum_{j\in I} c_{d_j})^2)^{1/2}\mathcal{O}(N^{-1-7\delta/15}), \end{aligned}$$

uniformly for ω satisfying (4.24). Writing $h(V_1, V_2) = J_N'(\hat{V}_1)\{\chi(V_1 - V_2) - V_1\}$ as before, we find by repeated use of Assumptions A and B (cf. (4.19), (4.20) and (4.36)) that, uniformly for ω satisfying (4.24),

$$ES_1^2S_2 + ES_1S_2^2 = \frac{A_{1N}}{N}\sum_{j\neq k}\sum b_j^2 b_k + \mathcal{O}(N^{-1-\varepsilon}),$$

where $\varepsilon > 0$ and

$$(4.39)\quad \begin{aligned} A_{1N} = E\tilde{J}_N^2(\hat{V}_1)h(V_2, V_1) + 2E\tilde{J}_N(\hat{V}_1)\tilde{J}_N(\hat{V}_2)h(V_1, V_2) \\ + 2E\tilde{J}_N(\hat{V}_1)h(V_1, V_3)h(V_2, V_3). \end{aligned}$$

It follows that uniformly for $|t| \le \log N$ and ω satisfying (4.24),

$$(4.40)\quad \begin{aligned} \rho_{1N}(t) = \{Ee^{it\sigma_N^{-1}S_1}\}\Bigg\{1 &+ \frac{(it)^2}{2\sigma_N^2}[\sigma^2(T_{\omega N}) - \sigma^2(S_1)] \\ &+ \frac{(it)^3}{2\sigma_N^3}\frac{A_{1N}}{N}\sum_{(j\neq k)}\sum b_j^2 b_k - \frac{(it)^4}{8N\sigma_N^4}[E\tilde{J}_N^4(\hat{V}_1) - \{E\tilde{J}_N^2(\hat{V}_1)\}^2]\Bigg\} \\ &+ \mathcal{O}(N^{-1-\varepsilon}\,|t|\,P(t)e^{-\theta t^2}(1 + (\textstyle\sum_{j\in I} c_{d_j})^2)^{1/2}), \end{aligned}$$

where $\varepsilon > 0$, $0 < \theta < \frac{1}{2}$ and P is a fixed polynomial.

Let us turn back to our starting point (4.7). Choose $\gamma \in (0, 1)$ and define the event $B = \{\sum_{j\in I} c_{D_j}^2 < 1 - \gamma\}$ (cf. (4.24)). According to Lemma 3.4, $P(B^c) = \mathcal{O}(N^{-22/15})$, so

$$\begin{aligned} \psi_N(t) &= Ee^{itT_N^*} \\ &= E[\chi(B)E(e^{it\sigma_N^{-1}\{T_N - Z_N - E(T_N - Z_N|\Omega)\}}\mid\Omega)e^{it\sigma_N^{-1}E(T_N - Z_N|\Omega)}E(e^{it\sigma_N^{-1}Z_N}\mid\Omega)] + \mathcal{O}(N^{-22/15}). \end{aligned}$$

From Lemma 3.3 it follows that $E|Z_N|^5 = \mathcal{O}(N^{-1-7\delta/3})$ and $E(E(T_N - Z_N|\Omega))^2 = \mathcal{O}(N^{-4/3-14\delta/15})$. Hence by Taylor expansion we obtain

$$(4.41)\quad \begin{aligned} \psi_N(t) = Ee^{itT_N^*} = E\Bigg[\chi(B)E(e^{it\sigma_N^{-1}\{T_N - Z_N - E(T_N - Z_N|\Omega)\}}\mid\Omega) \\ \cdot\Bigg\{1 + \frac{it}{\sigma}\{E(Z_N\mid\Omega) + E(T_N - Z_N\mid\Omega)\} + \frac{(it)^2}{2\sigma_N^2}\{E(Z_N^2\mid\Omega) \\ + 2E(Z_N\mid\Omega)E(T_N - Z_N\mid\Omega)\} + \frac{(it)^3}{6\sigma_N^3}E(Z_N^3\mid\Omega) + \frac{(it)^4}{24\sigma_N^4}E(Z_N^4\mid\Omega)\Bigg\}\Bigg] \\ + \mathcal{O}(N^{-22/15}) + \mathcal{O}([t^2 + |t|^5]N^{-1-7\delta/3}), \end{aligned}$$

uniformly for $|t| \le \log N$. In view of (4.15), (4.21) and (4.23) we have, uniformly for $|t| \le \log N$ and ω satisfying (4.24)

$$(4.42)\quad \begin{aligned} E(e^{it\sigma_N^{-1}\{T_N - Z_N - E(T_N - Z_N|\Omega=\omega)\}}\mid\Omega = \omega) &= Ee^{it\sigma_N^{-1}\{T_{\omega N} - E(T_{\omega N})\}} \\ &= \rho_N(t) + \mathcal{O}(|t|\,N^{-1-7\delta/15}(1 + (\textstyle\sum_{j\in I} c_{d_j})^2)^{1/2}) \\ &= \rho_{1N}(t) + \mathcal{O}(N^{-1-\varepsilon}\,|t|\,P(t)(1 + (\textstyle\sum_{j\in I} c_{d_j})^2)^{1/2}), \end{aligned}$$

where $\varepsilon > 0$ and P is a fixed polynomial.

Before substituting this in (4.41) we shall provide uniform bounds for the quantities $\sigma_N^2 - \sigma^2(T_{\omega N})$ and $\sigma^2(T_{\omega N}) - \sigma^2(S_1)$. Theorem II 3.1.c of Hájek and Šidák (1967) and

Assumption A imply that

$$\sigma^2(T_{\omega N}) = \frac{1}{M-1}\left(1 - \sum_{j\in I} c_{d_j}^2 - \frac{1}{M}\left(\sum_{j\in I} c_{d_j}\right)^2\right) \sum_{j=1}^{M} \left(J_N\left(\frac{j}{M+1}\right) - \bar{J}_N\right)^2,$$

where (cf. (4.8))

$$\bar{J}_N = \frac{1}{M}\sum_{j=1}^{M} J_N\left(\frac{j}{M+1}\right) = \frac{1}{M}\sum_{j=m+1}^{N-m} J\left(\frac{j}{N+1}\right).$$

It follows from (3.10) that $|\bar{J}_N| = \mathcal{O}(N^{-13/30-7\delta/15})$ and from Assumption A that $|\sum_{j\in I} c_{d_j}| = \mathcal{O}(N^{1/30})$, hence

$$\sigma^2(T_{\omega N}) = \frac{1}{M-1}\left(1 - \sum_{j\in I} c_{d_j}^2\right) \sum_{j=1}^{M} J_N^2\left(\frac{j}{M+1}\right) + \mathcal{O}(N^{-13/15-14\delta/15}), \tag{4.43}$$

uniformly in ω. Furthermore we know from (3.10) that $|\bar{J}| = \mathcal{O}(N^{-13/14-\delta})$, so in view of (2.5) and Assumption B we have

$$\begin{aligned} &|\sigma_N^2 - \sigma^2(T_{\omega N})| \\ &= \left|\frac{1}{N-1}\sum_{j=1}^{N} J^2\left(\frac{j}{N+1}\right) - \frac{1}{M-1}\left(1 - \sum_{j\in I} c_{d_j}^2\right)\sum_{j=1}^{M} J_N^2\left(\frac{j}{M+1}\right)\right| \\ &\quad + \mathcal{O}(N^{-13/15-14\delta/15}) \\ &= \left|\frac{1}{N-1}\sum_{j\in I} J^2\left(\frac{j}{N+1}\right) + \frac{1}{M-1}\left(\sum_{j\in I} c_{d_j}^2 - \frac{2m}{N}\right)\sum_{j=1}^{M} J_N^2\left(\frac{j}{M+1}\right)\right| \\ &\quad + \mathcal{O}(N^{-13/15-14\delta/15}) = \mathcal{O}(N^{-2/5-14\delta/15}), \end{aligned} \tag{4.44}$$

uniformly in ω.

To obtain the second bound, we argue as in Lemma 3.1 with J and $h(t)$ replaced by J_N and $h_N(t) = h((N+1)^{-1}(m + (M+1)t))$ to conclude that

$$\frac{1}{M}\sum_{j=1}^{M} J_N^2\left(\frac{j}{M+1}\right) = EJ_N^2(V_1) + \mathcal{O}(N^{-14/15-14\delta/15}).$$

One easily verifies that $|EJ_N^2(V_1) - E\tilde{J}_N^2(\hat{V}_1)| = \mathcal{O}(N^{-13/15-14\delta/15})$ and together with (4.43) and (4.16) this yields

$$|\sigma^2(T_{\omega N}) - \sigma^2(S_1)| = \mathcal{O}(N^{-13/15-14\delta/15}), \tag{4.45}$$

uniformly in ω.

A few more facts are needed to complete our calculation of $\psi_N(t)$. First we note that for $a = (m+1)(N+1)^{-1} = \mathcal{O}(N^{-7/15})$, Assumption B and (4.8) imply that

$$\int_0^a \{|J(t)|^k + |J(1-t)|^k\}\, dt = \mathcal{O}(N^{-7/15+k/30-7k\delta/15}),$$

for $k = 1, \cdots, 4$ and hence

$$\begin{aligned} &|EJ_N(\hat{V}_1)| = \mathcal{O}(N^{-13/30-7\delta/15}), \\ &E\tilde{J}_N^2(\hat{V}_1) = \frac{N+1}{M-1}\int_a^{1-a} J^2(t)\, dt + \mathcal{O}(N^{-13/15-14\delta/15}), \\ &E\tilde{J}_N^3(\hat{V}_1) = \frac{N+1}{M-1}\int_a^{1-a} J^3(t)\, dt - 3\left(\frac{N+1}{M-1}\right)^2 \left\{\int_a^{1-a} J^2(t)\, dt\right\}\left\{\int_a^{1-a} J(t)\, dt\right\} \\ &\qquad + \mathcal{O}(N^{-13/10-7\delta/3}), \\ &E\tilde{J}_N^4(\hat{V}_1) = \frac{N+1}{M-1}\int_a^{1-a} J^4(t)\, dt + \mathcal{O}(N^{-13/30-7\delta/15}). \end{aligned} \tag{4.46}$$

Furthermore, Lemma 3.3 yields

$$(4.47)\quad E(\sigma_N^2 - \sigma^2(T_N - Z_N|\Omega)) = E(E(Z_N^2|\Omega)) + 2E(E(Z_N|\Omega)E(T_N - Z_N|\Omega)) + \mathcal{O}(N^{-4/3-14\delta/15}).$$

Substituting the random versions of (4.42), (4.40) and (4.27) in (4.41) and combining (4.46) and (4.47) with (4.44) and (4.45), it follows after some computations and repeated use of Assumptions A and B that, uniformly for $N^{-3/2} \le |t| \le \log N$,

$$(4.48)\quad \psi_N(t) = e^{-t^2/2}\left\{1 - \frac{it^3}{6\sigma_N^3}\kappa_{3N} + \frac{t^4}{24\sigma_N^4}\kappa_{4N} - \frac{t^6}{72\sigma_N^6}\kappa_{3N}^2\right\} + o(N^{-1}|t|P(t)e^{-\theta t^2}) + \mathcal{O}(N^{-1-\varepsilon}|t|P(t)),$$

where $\varepsilon > 0$, $0 < \theta < ½$, P is a fixed polynomial and κ_{3N} and κ_{4N} are given by (2.9) and (2.10).

To conclude the proof of Theorem 2.1 we note that (3.1) implies

$$\sigma_N^2 = 1 + \mathcal{O}(N^{-6/7-2\delta}).$$

Substituting this in (4.48) we obtain (4.5) with λ_N as in (4.2) and the proof of the theorem is complete. □

5. Two-sample linear rank statistics. In this section we compare our results with the expansions for the two-sample linear rank statistics in Bickel and Van Zwet (1978). Let $1 \le n \le N$, $\lambda = nN^{-1}$ and assume that $\varepsilon \le \lambda \le 1 - \varepsilon$ for some fixed $\varepsilon \in (0, ½)$ and all N. Define $c_j = (1-\lambda)/\{N\lambda(1-\lambda)\}^{1/2}, j = 1, 2, \cdots, n$ and $c_j = -\lambda/\{N\lambda(1-\lambda)\}^{1/2}, j = n+1, \cdots, N$. It is easy to check that in this case the c_j's satisfy Assumption A.

Taking a scores generating function J which satisfies Assumption B, we define the two-sample linear rank statistic as in (1.1). For the distribution function F_N^* of the standardized version of this statistic, Theorem 2.1 provides an Edgeworth expansion with remainder $o(N^{-1})$:

if

$$(5.1)\quad \begin{aligned} \tilde{F}_N(x) = \Phi(x) - \phi(x)\Bigg\{&\frac{1-2\lambda}{6\{N\lambda(1-\lambda)\}^{1/2}}\left(\int_0^1 J^3(t)\,dt\right)(x^2-1) \\ &+ \frac{1}{24N\lambda(1-\lambda)}\left[(1-6\lambda+6\lambda^2)\int_0^1 J^4(t)\,dt - 3(1-2\lambda)^2\right](x^3-3x) \\ &+ \frac{(1-2\lambda)^2}{72N\lambda(1-\lambda)}\left(\int_0^1 J^3(t)\,dt\right)^2(x^5-10x^3+15x)\Bigg\}, \end{aligned}$$

then

$$\sup_{x\in\mathbb{R}}|F_N^*(x) - \tilde{F}_N(x)| = o(N^{-1}), \quad \text{as} \quad N \to \infty.$$

Bickel and Van Zwet (1978) consider the two-sample linear rank statistic T_N' for an arbitrary vector of scores $a = (a_1, a_2, \cdots, a_N)$, i.e.

$$(5.2)\quad T_N' = \textstyle\sum_{j=1}^N a_j V_j,$$

where

$$V_j = \begin{cases} 1, & 1 \le D_j \le n, \\ 0, & \text{otherwise}, \end{cases}$$

for $j = 1, 2, \cdots, N$ and where $D_1, D_2, \cdots, D_N$ denote the antiranks. In their paper they establish asymptotic expansions for the distribution function of T_N' under the null-hypothesis as well as under contiguous alternatives. A related paper is that of Robinson (1978) which deals only with the null-hypothesis.

In order to compare the results in Bickel and Van Zwet (1978) with Theorem 2.1 in the present paper, we introduce the following assumption on the scores a_j.

ASSUMPTION C. Let $a_j = J(j/(N+1))$ for $j = 1, 2, \cdots, N$. This scores generating function J is twice continuously differentiable on (0, 1) and satisfies (2.1) and (2.3); there exist positive numbers $K > 0$ and $0 < \beta < 1/6$ such that its first derivative J' satisfies

$$|J'(t)| \le K\{t(1-t)\}^{-7/6+\beta} \quad \text{for} \quad t \in (0, 1). \tag{5.3}$$

LEMMA 5.1. *If $\varepsilon \le \lambda \le 1-\varepsilon$ for some fixed $\varepsilon \in (0, ½)$ and Assumption* C *are satisfied, then as $N \to \infty$*

$$\sup_{x\in R} \left| P\left(\frac{T'_N - ET'_N}{\sigma(T'_N)} \le x\right) - \tilde{F}_N(x) \right| = o(N^{-1}), \tag{5.4}$$

where $\tilde{F}_N$ is defined in (5.1).

PROOF. The present lemma is almost an immediate consequence of Corollary 2.1 of Bickel and Van Zwet (1978). Assumption C guarantees that there exists a positive fraction of the scores which are at a distance of at least $N^{-3/2} \log N$ apart from each other. Furthermore, in view of Lemma 3.1 and Appendix 2 of Albers, Bickel and Van Zwet (1976), Assumption C yields that

$$\sum_{j=1}^N a_j = \mathcal{O}(N^{1/6-\beta}), \quad \sum_{j=1}^N a_j^2 = N + \mathcal{O}(N^{1/3-2\beta}),$$

$$\sum_{j=1}^N a_j^3 = N\int_0^1 J^3(t)\,dt + \mathcal{O}(N^{1/2-3\beta}), \quad \sum_{j=1}^N a_j^4 = N\int_0^1 J^4(t)\,dt + \mathcal{O}(N^{2/3-4\beta}).$$

Substituting this in the expansion $\tilde{R}(x, \bar{\lambda})$ (cf. (2.56) in Bickel and Van Zwet, 1978) and standardizing T'_N with the exact variance $\sigma^2(T'_N)$, the result follows. □

For the two-sample case, Lemma 5.1 is clearly a better result than Theorem 2.1, as was to be expected. Roughly speaking, Assumption B in Theorem 2.1 requires a bit more smoothness than Assumption C in Lemma 5.1; it also requires $\int |J|^{14+\varepsilon} < \infty$ instead of $\int |J|^{6+\varepsilon} < \infty$. For practical purposes, however, Assumption B is already quite satisfactory. It is gratifying to find that the expansions in the two results coincide. We note that some numerical examples are contained in Bickel and Van Zwet (1978).

6. Finite sample computations. In this section we investigate the performance of the Edgeworth expansions as approximations for the finite sample distributions of one special statistic, namely Spearman's rank correlation coefficient ρ_N. In particular we compare our expansions with the usual normal approximation. As noted in Section 1 we know that, under the null-hypothesis of independence, Spearman's rank correlation coefficient ρ_N is distributed as

$$T_N^* = \frac{12}{N(N+1)(N-1)^{1/2}} \sum_{j=1}^N jR_j - \frac{3(N+1)}{(N-1)^{1/2}}. \tag{6.1}$$

From Theorem 2.1 it follows that, as $N \to \infty$

$$F_N^*(x) = P(T_N^* \le x) = \tilde{F}_N(x) + o(N^{-1}), \tag{6.2}$$

where

$$\tilde{F}_N(x) = \Phi(x) + \phi(x)\left(\frac{9N^2 - 21}{100N(N^2-1)} + \frac{1}{10N}\right)(x^3 - 3x). \tag{6.3}$$

We note that the third cumulant is zero because the scores generating function is symmetric.

In Olds (1938) the exact distribution of T_N^* under the null-hypothesis was given for $N = 2$ through 7. The same results, together with the exact distribution for $N = 8$, were obtained by Kendall, Kendall and Babington Smith (1939). Further extensions of the exact distribution of Spearman's rank correlation coefficient under the hypothesis of independence were given in David, Kendall and Stuart (1951). They established the exact distribution for $N = 9$ and 10 and showed that the formal Edgeworth expansions including the N^{-3} term would be quite satisfactory in practice for $N \geq 10$.

In Table 6.1 a comparison of the Edgeworth expansion $\tilde{F}_N$ and the normal approximation Φ with the exact distribution F_N^* is made for sample sizes $N = 5$, 10 and 20 and various values of the argument. We note that $\tilde{F}_N$ is truncated at 1. Furthermore, we note that for

TABLE 6.1
Comparison of the exact distribution function with the Edgeworth expansion and normal approximation for $N = 5$, 10, and 20

x	F_5^*	$\tilde{F}_5$	F_{10}^*	$\tilde{F}_{10}$	F_{20}^*	$\tilde{F}_{20}$	Φ
0.0	0.5250	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
0.2	0.6083	0.5707	0.5810	0.5749	0.5759	0.5771	0.5793
0.4	0.6583	0.6399	0.6460	0.6475	0.6506	0.6515	0.6554
0.6	0.7417	0.7062	0.7200	0.7158	0.7190	0.7207	0.7257
0.8	0.7750	0.7679	0.7760	0.7778	0.7821	0.7830	0.7881
1.0	0.8250	0.8234	0.8350	0.8322	0.8363	0.8368	0.8413
1.2	0.8833	0.8715	0.8760	0.8781	0.8821	0.8815	0.8849
1.4	0.9333	0.9112	0.9169	0.9151	0.9174	0.9172	0.9192
1.6	0.9583	0.9423	0.9431	0.9437	0.9450	0.9445	0.9452
1.8	0.9917	0.9653	0.9666	0.9647	0.9647	0.9644	0.9641
2.0	1.0000	0.9812	0.9805	0.9793	0.9791	0.9783	0.9772
2.2	1.0000	0.9914	0.9913	0.9888	0.9879	0.9875	0.9861
2.4	1.0000	0.9973	0.9964	0.9946	0.9935	0.9932	0.9918
2.6	1.0000	1.0000	0.9992	0.9978	0.9966	0.9966	0.9953
2.8	1.0000	1.0000	0.9999	0.9995	0.9984	0.9985	0.9974
3.0	1.0000	1.0000	1.0000	1.0000	0.9994	0.9994	0.9987

TABLE 6.2
Comparison of the exact distribution function with the Edgeworth expansion and normal distribution after a continuity correction, for $N = 5$ and 10

x	F_5^*	$\tilde{F}_5$	Φ	F_{10}^*	$\tilde{F}_{10}$	Φ
0.0	0.5250	0.5354	0.5398	0.5000	0.5000	0.5000
0.2	0.6083	0.6056	0.6179	0.5810	0.5816	0.5864
0.4	0.6583	0.6736	0.6915	0.6460	0.6475	0.6554
0.6	0.7417	0.7377	0.7580	0.7200	0.7217	0.7318
0.8	0.7750	0.7965	0.8159	0.7760	0.7778	0.7881
1.0	0.8250	0.8485	0.8643	0.8350	0.8367	0.8457
1.2	0.8833	0.8924	0.9032	0.8760	0.8781	0.8849
1.4	0.9333	0.9278	0.9332	0.9169	0.9181	0.9219
1.6	0.9583	0.9548	0.9554	0.9431	0.9437	0.9452
1.8	0.9917	0.9741	0.9713	0.9666	0.9663	0.9655
2.0	1.0000	0.9870	0.9821	0.9805	0.9793	0.9772
2.2	1.0000	0.9948	0.9893	0.9913	0.9895	0.9867
2.4	1.0000	0.9991	0.9938	0.9964	0.9946	0.9918
2.6	1.0000	1.0000	0.9965	0.9992	0.9980	0.9956
2.8	1.0000	1.0000	0.9981	0.9999	0.9995	0.9974
3.0	1.0000	1.0000	0.9995	1.0000	1.0000	0.9987

$N = 20$ we have employed a Monte-Carlo simulation based on 90,000 samples to estimate the exact distribution function F_N^*.

Inspection of Table 6.1 shows that the agreement between the estimated exact distribution function F_{20}^* and the expansion $\tilde{F}_{20}$ is almost perfect. It also shows that the expansion performs much better than the normal approximation. For $N = 5$ and 10 the agreement between F_N^* and $\tilde{F}_N$ is reasonable but not nearly as good as for $N = 20$. This is due to the fact that the probabilities of single values are still rather large for such small values of N; one can't expect to approximate a distribution function with large jumps by a continuous one in a satisfactory manner. To overcome this problem, we have employed a continuity correction. In Table 6.2 we summarize the results with this continuity correction for $N = 5$ and 10. Inspection of this table shows that the approximations $\tilde{F}_N$ are much improved; for sample size $N = 10$ the expansion $\tilde{F}_{10}$ performs quite well. It also shows that the expansions provide much better approximations than the usual normal approximation.

Acknowledgments. It is a pleasure to acknowledge the many helpful discussions with Professor W. R. van Zwet during the preparation of this paper; I greatly appreciated his encouragement and constructive criticism. The author also thanks the editor and the referee for their suggestions.

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R. (1976). Asymptotic expansions for the power of distributionfree tests in the one-sample problem. *Ann. Statist.* **4** 108–156.

BICKEL, P. J. and VAN ZWET, W. R. (1978). Asymptotic expansions for the power of distributionfree tests in the two-sample problem. *Ann. Statist.* **6** 937–1004.

DAVID, S. T., KENDALL, M. G., and STUART, A. (1951). Some questions of distribution in the theory of rank correlation. *Biometrika* **38** 131–140.

DOES, R. J. M. M. (1981). An Edgeworth expansion for simple linear rank statistics under the null-hypothesis. Mathematical Centre Report SW 78, Mathematisch Centrum, Amsterdam.

DOES, R. J. M. M. (1982a). Berry-Esseen theorems for simple linear rank statistics under the null-hypothesis. *Ann. Probability* **10** 982–991.

DOES, R. J. M. M. (1982b). *Higher Order Asymptotics for Simple Linear Rank Statistics.* Ph.D. Thesis, University of Leiden; Mathematical Centre Tracts 151, Mathematisch Centrum, Amsterdam.

DUPAČ, V. and HÁJEK, J. (1969). Asymptotic normality of simple linear rank statistics under alternatives II. *Ann. Math. Statist.* **40** 1992–2017.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd edition. Wiley, New York.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic, New York.

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

HELMERS, R. (1980). Edgeworth expansions for linear combinations of order statistics with smooth weight functions. *Ann. Statist.* **8** 1361–1374.

HO, S. T. and CHEN, L. H. Y. (1978). An L_P bound for the remainder in a combinatorial central limit theorem. *Ann. Probability* **6** 231–249.

HUŠKOVÁ, M. (1977). The rate of convergence of simple linear rank statistics under hypothesis and alternatives. *Ann. Statist.* **5** 658–670.

HUŠKOVÁ, M. (1979). The Berry-Esseen theorem for rank statistics. *Comment. Math. Univ. Carolin.* **20** 399–415.

JOHNSON, N. L. and KOTZ, S. (1969). *Distributions in Statistics: Discrete Distributions.* Wiley, New York.

KENDALL, M. G., KENDALL, S. F. H. and BABINGTON SMITH, B. (1939). The distribution of Spearman's coefficient of rank correlation in a universe in which all rankings occur an equal number of times. *Biometrika* **30** 251–273.

OLDS, E. G. (1938). Distributions of sums of squares of rank differences for small numbers of individuals. *Ann. Math. Statist.* **9** 133–148.

PETROV, V. V. (1975). *Sums of Independent Random Variables.* Springer Verlag, Berlin.

ROBINSON, J. (1978). An asymptotic expansion for samples from a finite population. *Ann. Statist.* **6** 1005–1011.

VAN ZWET, W. R. (1982). On the Edgeworth expansion for the simple linear rank statistic. *Coll. Math. Soc. J. Bolyai* **32** 889–909. B.V. Gredenko, M.L. Puri, I. Vincze (eds.). North-Holland, Amsterdam.

DEPARTMENT OF MEDICAL INFORMATICS
AND STATISTICS
UNIVERSITY OF LIMBURG
P.O. Box 616
6200 MD MAASTRICHT
THE NETHERLANDS.