The Annals of Statistics
1976, Vol. 4, No. 3, 597-606

ASYMPTOTIC EXPANSION AND A LOCAL LIMIT THEOREM FOR A FUNCTION OF THE KENDALL RANK CORRELATION COEFFICIENT

BY ZUZANA PRÁŠKOVÁ-VÍZKOVÁ

Charles University

In the present paper, an integer-valued version (T_N) of the Kendall rank correlation coefficient is considered. Under the hypothesis of independence, a local limit theorem with the Edgeworth expansion for T_N is proved and an asymptotic expansion of the distribution function of T_N is derived.

1. Introduction. Let $X_1, Y_1, \cdots, X_N, Y_N$ be independent random variables, the X_i's with a continuous distribution function (df) F_0, the Y_i's with a continuous df G_0. Let R_i and S_i be the ranks of X_i and Y_i, respectively.

Consider the Kendall rank correlation coefficient

$$\tau = (N(N-1))^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \operatorname{sign}(R_i - R_j) \operatorname{sign}(S_i - S_j) \tag{1.1}$$

and its transformation T_N defined by

$$\tau = 4(N(N-1))^{-1} T_N - 1 \,. \tag{1.2}$$

It is well known that $ET_N = N(N-1)/4$, $\operatorname{Var} T_N = N(N-1)(2N+5)/72$ and that the distribution of T_N is asymptotically normal (see Kendall (1948), page 69). The cumulants (of the statistic $S^* = 2T_N - N(N-1)/2$) are given explicitly in Moran (1950) and Silverstone (1950). The Edgeworth expansion up to the order N^{-3} for the df of τ was investigated numerically (without studying the rate of convergence) by David, Kendall and Stuart (1951).

In the present paper, the Edgeworth expansion is established both for the probabilities $P(T_N = k)$ (Theorem 2.1) as well as for the df of T_N (Theorem 2.2), in the latter case with additional terms due to the lattice character of the distribution.

2. Basic notation and main results. Let Φ and φ be the df and the density of $\mathscr{N}(0, 1)$, let H_m be the Hermite polynomial of degree m (defined by $d^m\varphi(x)/dx^m = (-1)^m H_m(x)\varphi(x)$). Let κ_j, $j \geqq 1$, be the cumulants of T_N; κ_1, κ_2 being denoted as μ, σ^2, alternatively. Put

$$\tilde{\kappa}_j = \kappa_j/\sigma^j \,, \qquad j \geqq 1 \,. \tag{2.1}$$

For every integer $\nu \geqq 1$ put

$$Q_{2\nu}(x) = \sum{}^* H_{2\nu+2\Sigma k_j}(x) \prod_{j=2}^{\nu+1} (k_j!)^{-1} (\tilde{\kappa}_{2j}/(2j)!)^{k_j} \,, \tag{2.2}$$

Received July 1974; revised November 1975.

AMS 1970 *subject classifications.* Primary 62G10, Secondary 60E05, 60F05.

Key words and phrases. Kendall rank correlation coefficient, characteristic function, Esseen inequality, asymptotic expansion.

and

$$\bar{Q}_{2\nu}(x) = -\sum{}^{*} H_{2\nu+2\Sigma k_j - 1}(x) \prod_{j=2}^{\nu+1} (k_j!)^{-1} (\tilde{\kappa}_{2j}/(2j)!)^{k_j}, \tag{2.3}$$

where the summation $\sum^*$ extends over all nonnegative integers $k_2, \cdots, k_{\nu+1}$ such that $\sum_{j=2}^{\nu+1}(j-1)k_j = \nu$ (and $\sum$ in the subscripts stands for $\sum_{j=2}^{\nu+1}$). Notice that

$$\frac{d}{dx}(\varphi(x)\bar{Q}_{2\nu}(x)) = \varphi(x)Q_{2\nu}(x). \tag{2.4}$$

Further define

$$\Phi_p(x) = \Phi(x) + \varphi(x)\sum_{\nu=1}^{p-1}\bar{Q}_{2\nu}(x), \tag{2.5}$$

$$B_{2\lambda}(x) = \sum_{m=1}^{\infty}\cos(2\pi m x)(2^{2\lambda-1}(\pi m)^{2\lambda})^{-1}, \qquad \lambda \geqq 1, \tag{2.6}$$

$$B_{2\lambda+1}(x) = \sum_{m=1}^{\infty}\sin(2\pi m x)(2^{2\lambda}(\pi m)^{2\lambda+1})^{-1}, \qquad \lambda \geqq 0. \tag{2.7}$$

Throughout the paper, c and C denote positive constants, the values of which are not specified and may differ in different formulas (or even in different places in the same formula). The phrase "$F_N(x) = G_N(x) + O(N^{-p})$, uniformly in x" means the same as "$|F_N(x) - G_N(x)| \leqq cN^{-p}$ for all x."

Now, we shall formulate the theorems; their proofs will be given in Sections 3 and 4, respectively.

THEOREM 2.1. *Denote* $x = (k-\mu)/\sigma$, $k = 0, 1, \cdots, N(N-1)/2$. *Then*

$$\sigma P(T_N = k) = \varphi(x)(1 + \sum_{\nu=1}^{p-1} Q_{2\nu}(x)) + O(N^{-p}), \tag{2.8}$$

uniformly in k.

THEOREM 2.2. *Define*

$$h_\lambda = 1, \qquad \lambda = 4m+1, 4m+2,$$
$$h_\lambda = -1, \qquad \lambda = 4m+3, 4m,$$

where m *is an integer. Then*

$$P(\sigma^{-1}(T_N - \mu) < x) = \Phi_p(x) + \sum_{\lambda=1}^{p-1} h_\lambda \sigma^{-\lambda} B_\lambda(\sigma x + \mu)\frac{d^\lambda}{dx^\lambda}\Phi_p(x) + O(N^{-p}), \tag{2.9}$$

uniformly in x.

REMARKS. 1. The sum on the right-hand side of (2.9) includes terms, which are of higher order than N^{-p}; this is due to the fact that both the expression and the proof are more easy to handle in the present form (see Esseen (1945)).

2. The functions B_λ defined by (2.6) and (2.7) are all periodic with period 1. The functions B_λ, $\lambda \geqq 2$, are continuous while B_1 has the jump 1 at every integer point x. Explicitly, $B_1(m) = 0$, $B_1(m+t) = -t + \frac{1}{2}$, $m = 0, \pm 1, \cdots$, $0 < t < 1$.

Let us specialize Theorem 2.2 for $p = 2$, summing up the terms of order N^{-2}.

COROLLARY 2.3.

$$P(\sigma^{-1}(T_N - \mu) < x) = \Phi(x) - \tilde{\kappa}_4\varphi(x)H_3(x)(4!)^{-1} + \sigma^{-1}B_1(\sigma x + \mu)\varphi(x) + O(N^{-2}), \tag{2.10}$$

uniformly in x.

3. Proof of Theorem 2.1. Let $f_N, \bar{f}_N, \tilde{f}_N$ be the characteristic functions (cf) of T_N, $T_N - \mu$, $(T_N - \mu)/\sigma$, respectively. First we shall give some lemmas.

LEMMA 3.1. *The following inequalities hold true*:

$$|\bar{f}_N(t)| < ce^{-t^2\sigma^2/4} \quad \text{for} \quad 0 < |t| \leqq 2\pi/N, \tag{3.1}$$

$$|\bar{f}_N(t)| < ce^{-CN} \quad \text{for} \quad 2\pi/N \leqq |t| \leqq \pi. \tag{3.2}$$

PROOF. The cf f_N equals

$$f_N(t) = (N!)^{-1}(1 - e^{it}) \cdots (1 - e^{Nit})(1 - e^{it})^{-N} \tag{3.3}$$

which follows from the generating function of T_N established by Hájek (1955) and Kendall and Stuart (1961), page 479.

From (3.3) we immediately get

$$\bar{f}_N(t) = (N!)^{-1} \sin(t/2) \cdots \sin(Nt/2)(\sin(t/2))^{-N}. \tag{3.4}$$

Making use of the symmetry of $\bar{f}_N$ and of the inequalities

$$|\sin x| \geqq x(1 - x^2/6) \quad \text{valid for} \quad x \geqq 0,$$
$$\sin x \leqq x(1 - x^2/12) \quad \text{valid for} \quad 0 \leqq x \leqq \pi,$$
$$(1 - x)^N \geqq 1 - Nx \quad \text{valid for} \quad 0 \leqq x \leqq 1,$$

we obtain for $0 \leqq t \leqq 2\pi/N$

$$\begin{aligned} |\bar{f}_N(t)| &\leqq (N!)^{-1} \textstyle\prod_{j=1}^N (jt/2)(1 - j^2t^2/48)((t/2)(1 - t^2/24))^{-N} \\ &= (1 - t^2/24)^{-N} \textstyle\prod_{j=1}^N (1 - j^2t^2/48) \\ &\leqq c\exp(-t^2 \textstyle\sum_{j=1}^N j^2/48) < c\exp(-t^2\sigma^2/4). \end{aligned}$$

Now suppose that $2\pi/N \leqq t \leqq \pi$; let K denote the largest integer not exceeding $2\pi/t$. Then, utilizing $\sin x \geqq 2x/\pi$ for $0 \leqq x \leqq \pi/2$, we have

$$\begin{aligned} |\bar{f}_N(t)| &\leqq K! \textstyle\prod_{j=1}^K (1 - j^2t^2/48)/(N!(1 - t^2/24)^K(t/\pi)^{N-K}) \\ &\leqq K! \exp(-t^2 \textstyle\sum_{j=1}^K j^2/48)/(N!(1 - t^2/24)^K(t/\pi)^{N-K}). \end{aligned}$$

Applying Stirling's formula and the definition of K, we get

$$\begin{aligned} |\bar{f}_N(t)| &\leqq c\exp(-K + N)(Kt/\pi)^K N^{-N}(\pi/t)^N \exp(-t^2 \textstyle\sum_{j=1}^K j^2/48) \\ &< c\exp(-K + N + K\log 2 + N\log(\pi/Nt) - \pi^3/144t) \\ &\leqq c\max_{2\pi/N \leqq t \leqq \pi} \exp(-2\pi/t + N + (2\pi/t)\log 2 + N\log(\pi/Nt) - \pi^3/144t) \\ &= ce^{-CN}. \end{aligned}$$

□

LEMMA 3.2. *For* $|t| \leqq \sigma^{-1+\alpha}$, $0 < \alpha < \frac{1}{3}$, *the following expansion holds true*:

$$\log \bar{f}_N(t) = \textstyle\sum_{j=1}^p \kappa_{2j}(it)^{2j}/(2j)! + R(t), \tag{3.5}$$

where the remainder $R(t)$ *satisfies the inequality*

(3.6) $$|R(t)| \leqq cN^{2p+3}t^{2p+2} .$$

PROOF. By Kendall and Stuart (1961), page 479,

$$\log f_N(t) = \kappa_1 it + \sum_{j=1}^{\infty} \kappa_{2j}(it)^{2j}/(2j)! ,$$

where κ_j are the cumulants of T_N, given by formulas

(3.7) $$\kappa_1 = N(N-1)/4 ,$$

(3.8) $$\kappa_{2j+1} = 0 , \qquad j \geqq 1 ,$$

(3.9) $$\kappa_{2j} = \bar{B}_{2j}(\sum_{s=1}^{N} s^{2j} - N)/2j , \qquad j \geqq 1 ,$$

$\bar{B}_{2j}$ are the Bernoulli numbers. It is known that

$$|\bar{B}_{2j}| \leqq 4(2j)!(2\pi)^{-2j} , \qquad j \geqq 1 ,$$

hence,

$$|\kappa_{2j}| \leqq 4(2j)!N^{2j+1}(2\pi)^{-2j} , \qquad j \geqq 1 .$$

Now we can write

$$\log \bar{f}_N(t) = \sum_{j=1}^{\infty} \kappa_{2j}(it)^{2j}/(2j)! = \sum_{j=1}^{p} \kappa_{2j}(it)^{2j}/(2j)! + R(t) ,$$
$$|R(t)| \leqq \sum_{j=p+1}^{\infty} |\kappa_{2j}|t^{2j}/(2j)! \leqq 4N \sum_{j=p+1}^{\infty} (Nt)^{2j}(2\pi)^{-2j} .$$

For $|t| \leqq \sigma^{-1+\alpha}$, $0 < \alpha < \frac{1}{3}$, we have

$$|R(t)| \leqq 4N(Nt/2\pi)^{2p+2}(1-(Nt/2\pi)^2)^{-1} \leqq cN^{2p+3}t^{2p+2} . \qquad \square$$

LEMMA 3.3. *For* $|t| \leqq \sigma^{\alpha}$, $0 < \alpha < \frac{1}{6}$,

(3.10) $$\tilde{f}_N(t) = e^{-t^2/2}(1 + \sum_{\nu=1}^{p-1} P_{2\nu}(it)) + Z(t) ,$$

where $P_{2\nu}$ *are polynomials of degree* 4ν *in it, coefficients of which are of order* $N^{-\nu}$, *or, explicitly, in the notation of Section* 2,

(3.11) $$P_{2\nu}(it) = \sum{}^{*} \prod_{j=2}^{\nu+1} (k_j!)^{-1}(\tilde{\kappa}_{2j}/(2j)!)^{k_j}(it)^{2\nu+2\Sigma k_j}$$

and the remainder $Z(t)$ *satisfies the inequality*

(3.12) $$|Z(t)| \leqq ce^{-t^2/2}N^{-p}|t^p||Z_p(t)| ,$$

where Z_p *is a polynomial in t depending only on p.*

PROOF. Suppose that $|t| \leqq \sigma^{\alpha}$, $0 < \alpha < \frac{1}{3}$. From Lemma 3.2 we get

$$\tilde{f}_N(t) = \bar{f}_N(t/\sigma) = e^{-t^2/2}\exp(\sum_{j=2}^{p} \tilde{\kappa}_{2j}(it)^{2j}/(2j)! + R(t/\sigma)) ,$$
$$|R(t/\sigma)| \leqq cN^{2p+3}(t/\sigma)^{2p+2} .$$

Denote

$$\xi(t, \sigma) = \sum_{j=2}^{p} \tilde{\kappa}_{2j}(it)^{2j}/(2j)! + R(t/\sigma) .$$

Now, suppose that $|t| \leqq \sigma^{\alpha}$, $0 < \alpha < \frac{1}{6}$. Then $|\xi(t, \sigma)| \leqq C$ and from the Taylor expansion we get for a $\vartheta \in (0, 1)$

(3.13) $$\tilde{f}_N(t) = e^{-t^2/2}(\sum_{k=0}^{p-1} (k!)^{-1}\xi^k(t, \sigma) + (p!)^{-1}\xi^p(t, \sigma)e^{\vartheta\xi(t,\sigma)}) .$$

Making use of the multinomial formula, we have for $k \geqq 1$

$$\xi^k(t, \sigma) = \sum k! \prod_{j=2}^{p+1} (k_j!)^{-1} \prod_{j=2}^{p} (\tilde{\kappa}_{2j}(it)^{2j}/(2j)!)^{k_j} R(t/\sigma)^{k_{p+1}} ,$$

where the sum is extended over all nonnegative integers $k_2, \cdots, k_{p+1}$ such that $\sum_{j=2}^{p+1} k_j = k$. Observe that $\tilde{\kappa}_{2j} = O(N^{-j+1})$ and $R(t/\sigma) = t^{2p+2}S(t/\sigma)$, where $S(t/\sigma) = O(N^{-p})$. Hence, for given $k_2, \cdots, k_{p+1}, \sum_{j=2}^{p+1} k_j = k$,

$$\prod_{j=2}^{p} (\tilde{\kappa}_{2j}(it)^{2j}/(2j)!)^{k_j} R(t/\sigma)^{k_{p+1}} = t^{2\Sigma jk_j} D(N) ,$$

where

$$D(N) = O(N^{-\Sigma jk_j + k})$$

and $\sum$ in the superscripts stands for $\sum_{j=2}^{p+1}$. Consequently, the terms of (3.13) can be ordered in increasing powers of N. We shall write

$$\tilde{f}_N(t) = e^{-t^2/2}(1 + P_2(it) + \cdots + P_{2(p-1)}(it)) + Z(t) .$$

In this notation,

$$P_{2\nu}(it) = \sum{}^* \prod_{j=2}^{p+1} (k_j!)^{-1} \prod_{j=2}^{p} (\tilde{\kappa}_{2j}(it)^{2j}/(2j)!)^{k_j} R(t/\sigma)^{k_{p+1}} ,$$

where the sum $\sum^*$ is extended over all nonnegative integers $k_2, \cdots, k_{p+1}$, satisfying (for given ν, $1 \leqq \nu \leqq p - 1$) the conditions

$$k_2 + \cdots + k_{p+1} = k ,$$
$$\sum_{j=2}^{p+1} jk_j - k = \nu$$

with k varying from 1 to $p - 1$. These conditions can be written as

$$\sum_{j=2}^{p+1} (j-1)k_j = \nu . \tag{3.14}$$

On the other hand, $k_j \geqq 0$ for all $2 \leqq j \leqq p + 1$. Therefore $k_j = 0$ for $\nu + 2 \leqq j \leqq p + 1$ and thus (3.11) holds true. Maximizing $2 \sum_{j=2}^{\nu+1} jk_j$ under conditions $\sum_{j=2}^{\nu+1} (j-1)k_j = \nu$ for nonnegative integers k_j, $2 \leqq j \leqq \nu$, we obtain the degree of the polynomial $P_{2\nu}$. Especially, we have

$$P_2(it) = \tilde{\kappa}_4(it)^4/4! ,$$
$$P_4(it) = \tilde{\kappa}_6(it)^6/6! + \tilde{\kappa}_4^2(it)^8/(2(4!)^2) ,$$
$$P_6(it) = \tilde{\kappa}_8(it)^8/8! + \tilde{\kappa}_4\tilde{\kappa}_6(it)^{10}/(4!6!) + \tilde{\kappa}_4^3(it)^{12}/(3!(4!)^3) .$$

For the remainder $Z(t)$ we get

$$Z(t) = e^{-t^2/2}\Big(\sum_{k=1}^{p-1} \sum{}^1 \prod_{j=2}^{p+1} (k_j!)^{-1}(it)^{2\Sigma jk_j} D(N) + \exp(\vartheta\xi(t, \sigma)) \sum{}^2 \prod_{j=2}^{p+1} (k_j!)^{-1}(it)^{2\Sigma jk_j} D(N)\Big) ,$$

where $\sum^1$ denotes the summation over all nonnegative integers $k_2, \cdots, k_{p+1}$ such that $\sum_{j=2}^{p+1} k_j = k$ and $\sum_{j=2}^{p+1} jk_j - k \geqq p$ for $1 \leqq k \leqq p - 1$, and $\sum^2$ denotes the summation over all $k_2, \cdots, k_{p+1}$ such that $\sum_{j=2}^{p+1} k_j = p$. Hence,

$$|Z(t)| \leqq ce^{-t^2/2}\Big(\Big|\sum_{k=1}^{p-1} \sum{}^1 \prod_{j=2}^{p+1} (k_j!)^{-1} t^{2\Sigma jk_j} N^{-\Sigma jk_j + k} + \sum{}^2 \prod_{j=2}^{p+1} (k_j!)^{-1} t^{2\Sigma jk_j} N^{-\Sigma jk_j + p}\Big|\Big) \leqq cN^{-p} e^{-t^2/2} |t^p| \, |Z_p(t)| . \qquad \square$$

REMARK. The Fourier transform implies

$$(3.15) \qquad (2\pi)^{-1} \int_{-\infty}^{\infty} \exp(-itx - t^2/2) P_{2\nu}(it)\, dt = \varphi(x) Q_{2\nu}(x) \,,$$

$Q_{2\nu}$ being defined by (2.4).

PROOF OF THEOREM 2.1. Making use of the formula

$$P(T_N = k) = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-ivk} f_N(v)\, dv$$

and putting $v = t/\sigma$ and $x = (k - \mu)/\sigma$, we have

$$\sigma P(T_N = k) = (2\pi)^{-1} \int_{-\pi\sigma}^{\pi\sigma} e^{-itx} \tilde{f}_N(t)\, dt = J_1 + J_2 \,,$$

where

$$J_1 = (2\pi)^{-1} \int_{|t| \leq \sigma^\alpha} e^{-itx} \tilde{f}_N(t)\, dt \,,$$
$$J_2 = (2\pi)^{-1} \int_{\sigma^\alpha < |t| \leq \pi\sigma} e^{-itx} \tilde{f}_N(t)\, dt$$

and $0 < \alpha < \frac{1}{6}$.

We obtain from Lemma 3.1 that $J_2 = O(N^{-j})$ for every positive integer j. Utilizing (3.10) we have

$$\begin{aligned} J_1 = {} & (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itx - t^2/2} \Big(1 + \sum_{\nu=1}^{p-1} P_{2\nu}(it)\Big)\, dt \\ & - (2\pi)^{-1} \int_{|t| > \sigma^\alpha} e^{-itx - t^2/2} \Big(1 + \sum_{\nu=1}^{p-1} P_{2\nu}(it)\Big)\, dt \\ & + (2\pi)^{-1} \int_{|t| \leq \sigma^\alpha} e^{-itx} Z(t)\, dt = I_1 + I_2 + I_3 \,. \end{aligned}$$

From (3.15) we obtain

$$I_1 = \varphi(x)\Big(1 + \sum_{\nu=1}^{p-1} Q_{2\nu}(x)\Big) \,.$$

As we can easily show, $I_2 = O(N^{-p})$, $I_3 = O(N^{-p})$ and all order symbols are independent of k. □

4. Proof of Theorem 2.2.

LEMMA 4.1. *Let A, T and ε be arbitrary positive constants, let F be a nondecreasing saltus function and G a real function of bounded variation on the whole real line, f and g the corresponding Fourier–Stieltjes transforms such that*

(i) $F(-\infty) = G(-\infty) = 0$, $F(+\infty) = G(+\infty)$,

(ii) *F and G may be discontinuous only at $x = x_n$, $x_n < x_{n+1}$, $n = 0, \pm 1, \pm 2, \cdots$, and there exists a constant $L > 0$ such that* $\min (x_{n+1} - x_n) \geqq L$,

(iii) $|G'(x)| \leqq A$ *everywhere except when $x = x_n$, $n = 0, \pm 1, \cdots$,*

(iv) $\int_{-T}^{T} (|f(t) - g(t)|/|t|)\, dt = \varepsilon$.

Then to every number $k > 1$ there correspond two finite positive constants $c_1(k)$ and $c_2(k)$ depending only on k such that

$$(4.1) \qquad |F(x) - G(x)| \leqq k\varepsilon(2\pi)^{-1} + c_1(k) A T^{-1},$$

provided that $TL \geqq c_2(k)$.

For proof see Esseen (1945).

PROOF OF THEOREM 2.2. We shall apply Lemma 4.1. Put

$$F(x) = P((T_N - \mu)/\sigma < x) , \tag{4.2}$$

$$G(x) = \Phi_p(x) + \sum_{\lambda=1}^{p-1} h_\lambda \sigma^{-\lambda} B_\lambda(\sigma x + \mu) \frac{d^\lambda}{dx^\lambda} \Phi_p(x) , \tag{4.3}$$

where Φ_p, B_λ and h_λ were defined in Section 2. The corresponding Fourier–Stieltjes transforms are

$$f(t) = \tilde{f}_N(t) \tag{4.4}$$

and

$$g(t) = \int_{-\infty}^{\infty} e^{itx}\, dG(x) .$$

Let us denote

$$g_p(t) = e^{-t^2/2}(1 + \sum_{\nu=1}^{p-1} P_{2\nu}(it)) , \tag{4.5}$$

where $P_{2\nu}$ is defined in (3.11). Then we get

$$g(t) = g_p(t) + \psi(t) , \tag{4.6}$$

where

$$\psi(t) = -it \sum_{\lambda=1}^{p-1} h_\nu \sigma^{-\lambda} \int_{-\infty}^{\infty} e^{itx} B_\lambda(\sigma x + \mu) \frac{d^\lambda}{dx^\lambda} \Phi_p(x)\, dx .$$

Making use of the identities

$$\frac{d^r}{dx^r} \Phi(x) = (-1)^{r-1} \varphi(x) H_{r-1}(x) , \tag{4.7}$$

$$\frac{d^r}{dx^r} (\varphi(x) H_s(x)) = (-1)^r \varphi(x) H_{r+s}(x) , \tag{4.8}$$

$$\int_{-\infty}^{\infty} e^{itx} \varphi(x) H_r(x)\, dx = (it)^r e^{-t^2/2} , \tag{4.9}$$

we obtain after rather lengthy but easy calculations

$$\psi(t) = -t \sum_{\nu=1}^{p-1} \sum_{m=-\infty}^{\prime\infty} e^{2\pi i m\mu} (2\pi m\sigma)^{-\nu} (t + 2\pi m\sigma)^{\nu-1} g_p(t + 2\pi m\sigma) , \tag{4.10}$$

where $\sum'$ means that the summation extends over all integers except zero. We can see that G is discontinuous only at points $x_n = (-\mu + n)\sigma^{-1}$, $n = 0, \pm 1, \pm 2, \cdots$. For $x \neq x_n$ we have

$$G'(x) = \varphi(x)(1 + \sum_{\nu=1}^{p-1} Q_{2\nu}(x)) + \sum_{\lambda=1}^{p-1} h_\lambda \sigma^{-\lambda} \left\{ \frac{d}{dx} B_\lambda(\sigma x + \mu) \frac{d^\lambda}{dx^\lambda} \Phi_p(x) + B_\lambda(\sigma x + \mu) \frac{d^{\lambda+1}}{dx^{\lambda+1}} \Phi_p(x) \right\} .$$

Observe that

$$\frac{d}{dx} B_1(x) = -1 , \qquad x \neq 0, \pm 1, \pm 2, \cdots,$$

$$\frac{d}{dx} B_{2\lambda}(x) = -B_{2\lambda-1}(x) \qquad \text{for all } x ,$$

$$\frac{d}{dx} B_{2\lambda+1}(x) = B_{2\lambda}(x) \qquad \text{for all } x .$$

Hence,

$$G'(x) = h_{p-1}\sigma^{-(p-1)}B_{p-1}(\sigma x + \mu)\frac{d^p}{dx^p}\Phi_p(x)$$
$$= (-1)^{p-1}h_{p-1}\sigma^{-(p-1)}B_{p-1}(\sigma x + \mu)\varphi(x)(H_{p-1}(x) + \textstyle\sum_{\nu=1}^{p-1}\tilde{Q}_{2\nu}(x)) ,$$

where $\tilde{Q}_{2\nu}$ is the polynomial $Q_{2\nu}$ defined by (2.2) with $H_{2\nu+2\Sigma k_j+p-1}$ instead of $H_{2\nu+2\Sigma k_j}$. Hence, for $x \neq x_n$,

$$|G'(x)| \leqq c\sigma^{-(p-1)} = A . \tag{4.11}$$

Further put $L = \sigma^{-1}$, $T = \pi\sigma^2$ and suppose that N is so large that $TL = \pi\sigma \geqq c_2(k)$, $c_2(k)$ being the constant in Lemma 4.1. It remains to estimate

$$\varepsilon = \textstyle\int_{-T}^{T} (|f(t) - g(t)|/|t|)\, dt = \int_{-\pi\sigma^2}^{-\pi\sigma} + \int_{-\pi\sigma}^{\pi\sigma} + \int_{\pi\sigma}^{\pi\sigma^2} \tag{4.12}$$
$$= \varepsilon_1 + \varepsilon_2 + \varepsilon_3 .$$

First, we shall consider ε_2. Choosing $0 < \alpha < \frac{1}{6}$, we obtain from (4.4), (3.10), (4.5) and (4.6)

$$\varepsilon_2 \leqq \textstyle\int_{|t|\leqq\sigma^\alpha} (|Z(t)|/|t|)\, dt + \int_{|t|\leqq\sigma^\alpha} (|\psi(t)|/|t|)\, dt$$
$$+ \textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|f(t) - g(t)|/|t|)\, dt .$$

From (3.12) we get

$$\textstyle\int_{|t|\leqq\sigma^\alpha} (|Z(t)|/|t|)\, dt = O(N^{-p}) . \tag{4.13}$$

Now we shall consider the function $\psi(t)/t$ for $|t| \leqq \pi\sigma$. Obviously,

$$|\psi(t)|/|t| \leqq \textstyle\sum_{\nu=1}^{p-1} \sum_{m=-\infty}^{\prime\infty} |2\pi m\sigma|^{-\nu}|t + 2\pi m\sigma|^{\nu-1}|g_p(t + 2\pi m\sigma)|$$

and (4.5), (3.11) and the assumption $|t| \leqq \pi\sigma$ imply

$$|\psi(t)|/|t| \leqq c \textstyle\sum_{m=1}^{\infty} e^{-m^2\pi^2\sigma^2/2}(m\pi\sigma)^{4p-5} .$$

Consequently, for every integer $j \geqq 2$, there exists a positive constant C such that

$$|\psi(t)|/|t| \leqq C\sigma^{-j} , \qquad |t| \leqq \pi\sigma . \tag{4.14}$$

Hence,

$$\textstyle\int_{|t|\leqq\sigma^\alpha} (|\psi(t)|/|t|)\, dt = O(N^{-p}) . \tag{4.15}$$

It remains to estimate $\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|f(t) - g(t)|/|t|)\, dt$. We have

$$\textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|f(t) - g(t)|/|t|)\, dt \leqq \int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|f(t)|/|t|)\, dt$$
$$+ \textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|g(t)|/|t|)\, dt .$$

Making use of (3.1) and (3.2) we obtain

$$\textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|f(t)|/|t|)\, dt = O(N^{-p}) . \tag{4.16}$$

Finally, it can be easily shown that

$$\textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|g(t)|/|t|)\, dt \leqq \int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|g_p(t)|/|t|)\, dt \tag{4.17}$$
$$+ \textstyle\int_{\sigma^\alpha<|t|\leqq\pi\sigma} (|\psi(t)|/|t|)\, dt = O(N^{-p}) ,$$

which together with (4.13), (4.15), (4.16) implies that $\varepsilon_2 = O(N^{-p})$.

To investigate ε_3, we shall use the method given by Esseen (1945), Theorem 4.3. Obviously, we have

$$\varepsilon_3 = \int_{\pi\sigma}^{\pi\sigma^2} (|f(t) - g(t)|/|t|)\, dt \leqq \int_{\pi\sigma}^{\pi\sigma^2} (|g_p(t)|/|t|)\, dt + \int_{\pi\sigma}^{\pi\sigma^2} (|f(t) - \psi(t)|/|t|)\, dt = \int_{\pi\sigma}^{\pi\sigma^2} (|f(t) - \psi(t)|/|t|)\, dt + O(N^{-p}) .$$

After the substitution $t = v\sigma$, we have

$$I = \int_{\pi\sigma}^{\pi\sigma^2} (|f(t) - \psi(t)|/|t|)\, dt = \int_{\pi}^{\pi\sigma} (|f(v\sigma) - \psi(v\sigma)|/|v|)\, dv = \sum_{k=1}^{r} \int_{(2k-1)\pi}^{(2k+1)\pi} + \int_{(2r+1)\pi}^{\pi\sigma} ,$$

where r is the integral part of $(\sigma - 1)/2$. Let us consider

$$I_k = \int_{(2k-1)\pi}^{(2k+1)\pi} (|f(v\sigma) - \psi(v\sigma)|/|v|)\, dv \tag{4.18}$$

and put $v = t + 2\pi k$; then

$$I_k = \int_{-\pi}^{\pi} (|f((t + 2\pi k)\sigma) - \psi((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt = \int_{-\pi}^{\pi} (|e^{-2\pi i k\mu} f(t\sigma) - \psi((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt$$

which follows from (4.4) and from the periodicity of f_N. By (4.10)

$$\psi((t + 2\pi k)\sigma) = -(t + 2\pi k) \sum_{\nu=1}^{p-1} (-1)^\nu (2\pi k)^{-\nu} t^{\nu-1} g_p(t\sigma) e^{-2\pi i k\mu} + \psi_1((t + 2\pi k)\sigma) = \psi_0((t + 2\pi k)\sigma) + \psi_1((t + 2\pi k)\sigma) ,$$

where $\psi_1((t + 2\pi k)\sigma)$ means $\psi((t + 2\pi k)\sigma)$ with the $(-k)$th term of the sum $\sum'$ left out. Thus we can write

$$I_k \leqq \int_{-\pi}^{\pi} (|e^{-2\pi i k\mu} f(t\sigma) - \psi_0((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt + \int_{-\pi}^{\pi} (|\psi_1((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt = I_k' + I_k'' .$$

Similarly as in (4.14) we obtain that for every integer $j \geqq 2$

$$|\psi_1((t + 2\pi k)\sigma)|/|t + 2\pi k| \leqq C\sigma^{-j} \qquad \text{for } |t| \leqq \pi ,$$

hence,

$$I_k'' = O(N^{-p}\sigma^{-1}k^{-1}) . \tag{4.19}$$

Let $0 < \alpha < \frac{1}{6}$. Then for $|t| \leqq \sigma^{-1+\alpha}$,

$$e^{-2\pi i k\mu} f(t\sigma) - \psi_0((t + 2\pi k)\sigma) = e^{-2\pi i k\mu}((-1)^{p-1}(t/2\pi k)^{p-1} g_p(t\sigma) + Z(t\sigma))$$

which follows from (4.4), (3.10) and (4.5). Hence,

$$J_k' = \int_{|t| \leqq \sigma^{-1+\alpha}} (|e^{-2\pi i k\mu} f(t\sigma) - \psi_0((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt \leqq \int_{|t| \leqq \sigma^{-1+\alpha}} (|(t/2\pi k)^{p-1} g_p(t\sigma)|/|t + 2\pi k|)\, dt + \int_{|t| \leqq \sigma^{-1+\alpha}} (|Z(t\sigma)|/|t + 2\pi k|)\, dt = O(N^{-p}k^{-1}\sigma^{-\frac{1}{2}}) .$$

From (3.1), (3.2) and from the properties of g_p we obtain

$$J_k'' = \int_{\sigma^{-1+\alpha} < |t| \leqq \pi} (|e^{-2\pi i k\mu} f(t\sigma) - \psi_0((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt \leqq \int_{\sigma^{-1+\alpha} < |t| \leqq \pi} (|f(t\sigma)|/|t + 2\pi k|)\, dt + \int_{\sigma^{-1+\alpha} < |t| \leqq \pi} (|\psi_0((t + 2\pi k)\sigma)|/|t + 2\pi k|)\, dt = O(N^{-p}k^{-1}\sigma^{-\frac{1}{2}}) .$$

We may conclude that $I_k' = J_k' + J_k'' = O(N^{-p}k^{-1}\sigma^{-\frac{1}{2}})$ and thus

$$I_k = O(N^{-p}k^{-1}\sigma^{-\frac{1}{2}}) . \tag{4.20}$$

Finally,

$$I = \sum_{k=1}^{O(\sigma)} I_k = O(\log \sigma / N^p \sigma^{\frac{1}{2}}) = O(N^{-p}) ,$$

hence,

$$\varepsilon_3 = O(N^{-p}) .$$

Similarly we obtain that $\varepsilon_1 = O(N^{-p})$ and we may conclude that

$$\varepsilon = O(N^{-p}) .$$

Now we can apply inequality (4.1) of Lemma 4.1. We get

$$|F(x) - G(x)| \leqq \varepsilon k(2\pi)^{-1} + c_1(k)\sigma^{-(p+1)} \leqq cN^{-p} ,$$

or,

$$F(x) = G(x) + O(N^{-p}) . \qquad \square$$

Acknowledgment. I am grateful to Professor J. Hájek for suggesting this problem, and to Professor V. Dupač for his kind and essential help throughout the preparation of this work. I am also grateful to the referee for his critical comments.

REFERENCES

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

[2] ESSEEN, C. G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Math.* **77** 1–125.

[3] HÁJEK, J. (1955). Some rank distributions and their use. (In Czech.) *Časopis Pěst. Mat.* **80** 17–31.

[4] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academia, Prague.

[5] KENDALL, M. G. (1948). *Rank Correlation Methods.* Griffin, London. 2nd ed. 1955.

[6] KENDALL, M. G. and STUART, A. (1961). *The Advanced Theory of Statistics* **2**. Griffin, London. 2nd ed. 1967.

[7] MORAN, P. A. P. (1950). Recent developments in ranking theory. *J. Roy. Statist. Soc. Ser B* **12** 153–162.

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

[9] SILVERSTONE, H. (1950). A note on the cumulants of Kendall's *S*-distribution. *Biometrika* **37** 231–235.

DEPARTMENT OF STATISTICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
18600 PRAGUE 8
CZECHOSLOVAKIA