The Annals of Statistics
1976, Vol. 4, No. 1, 175-186

ASYMPTOTIC LINEARITY OF WILCOXON SIGNED-RANK STATISTICS[1]

BY ANDRÉ ANTILLE

University of California, Berkeley

As a "robust" alternative to the least squares estimates for a regression parameter, Koul (1969) proposed new estimates based on signed-rank statistics. To find out their asymptotic distribution Koul proved that under quite general assumptions, the signed-rank statistics of Wilcoxon type are asymptotically linear in the sense that they are uniformly approximable by linear forms in the regression parameter. More general results have been obtained by Van Eeden (1972) in a paper which is an analog to Jurečková's paper (1969) dealing with linear rank statistics.

In 1972 the author proved that the statistic used to define the Hodges–Lehmann estimate for a location parameter is asymptotically linear in a stronger sense, the result being to Koul's theorem what the central limit theorem is to the weak law of large numbers.

For the general linear regression model with one parameter the signed-rank statistics are proved to be linear in a strong sense, that is, the differences between the statistics and the linear forms mentioned above, properly normalized, converge weakly to linear processes. Results in this direction for linear rank statistics have been obtained by Jurečková (1973). As an application of the theorems presented here, one can construct new estimates for the squared L_2-norm of the underlying density, and this in much the same way as in Antille (1974). It is also possible to get more information about the asymptotic behavior of the linearized versions, proposed by Kraft and Van Eeden (1972).

0. Introduction. Let $x_1, x_2, \cdots, x_n$ be independent real random variables with continuous distribution functions $F(y - b_{in}t)$, $i = 1, 2, \cdots, n$, where the numbers b_{in}, $i = 1, 2, \cdots, n$ are known and t is unknown. Consider the problem of estimating t in the case where $F(y)$ has a density $f(y)$, symmetric around the origin. The usual least square estimates are very sensitive to large deviations. A "robust" alternative to this method was proposed by Koul (1969). His estimates are derived from linear signed-rank statistics. To study the asymptotic behavior of such estimates Koul proved that under quite general assumptions the signed-rank statistics of Wilcoxon type are asymptotically linear in the parameter, that is, the statistics differ from a linear function by an amount which tends to zero in probability as the number of observations increases. A similar result about linearity in the two-sample situation is due to Jurečková (1969). In 1972 the author proved that in the case of a location parameter the

Received August 1974; revised May 1975.

[1] This research was supported by the Swiss National Science Foundation.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 60F05.

Key words and phrases. Weak convergence, projection of a statistic, symmetric distribution, "Skorohod" topology, signed-rank statistics.

rank statistic used to define the Hodges–Lehmann estimate is not only linear in the sense mentioned above, but if the difference between the statistic and the linear form is multiplied by a suitable constant (usually $n^{\frac{1}{2}}$) one obtains a stochastic process which converges weakly (in the "Skorohod" topology) to a linear process.

Such a property is shown to hold for more general processes in the case of linear regression with one parameter. However, this paper deals with Wilcoxon signed-rank statistics only, i.e. statistics with linear score generating function. Similar results were proved by Jurečková (1973) for linear rank statistics of Wilcoxon type.

The results presented here enable us to get more precise information about the asymptotic behavior of the linearized versions proposed by Kraft and Van Eeden (1972), as well as new estimates of the squared L_2-norm of the underlying density. This can be done in much the same way as in Antille (1974).

The paper is divided in three sections and an appendix. In Section 1 we state the general assumptions and the main theorems. We also set up some notation which will be used throughout the paper. Section 2 contains the proof of the weak convergence of the finite dimensional distributions, while the "tightness" property is considered in Part 1 of Section 3. A proof of Theorem 2 is given in Part 2. Two propositions and two lemmas constitute the Appendix.

1. Notation, assumptions, theorems.

1.1. *Notation.* If an integral extends over $(-\infty, +\infty)$ then write $\int$. The indicator function of a set A is denoted by $I(A)$. Symbols of the form $\sum_{i,j\neq i}^{n}$ mean that the summation extends over all $1 \leqq i, j \leqq n$ with $j \neq i$. Denote by $E(X\mid Y)$ the conditional expectation of X given Y. Let $||f||_2$ be the L_2-norm of f. For every real, bounded function x, defined on $[0, 1]$, let

$$w(x, \delta) = \sup\{|x(t) - x(s)| : t, s \in [0, 1], |t - s| \leqq \delta\}.$$

1.2. *Assumptions and definitions.* Let $X_1, X_2, \cdots$ be i.i.d. real random variables with symmetric density $f(x)$. Let b_{in} and c_{in}, $1 \leqq i \leqq n$, $n = 1, 2, \cdots$, be constants satisfying:

(a) $\sup_{1\leqq i\leqq n} |b_{in}| + \sup_{1\leqq i\leqq n} |c_{in}| \leqq M$ for all n,
(b) $[\sup_{1\leqq i\leqq n} (b_{in}^2 c_{in}^2)][\sum_i^n b_{in}^2 c_{in}^2]^{-1} \to 0$ as $n \to \infty$,
(c) $(\sum_i^n b_{in} c_{in}) n^{-1} \to k \neq 0$ as $n \to \infty$,
(d) for all n, either:

$$1.\quad c_{in} b_{in} \geqq 0 \qquad \text{for all} \quad i,$$
$$2.\quad (|c_{in}| - |c_{jn}|)(|b_{in}| - |b_{jn}|) \geqq 0 \qquad \text{for all} \quad i, j,$$

or

$$1.\quad c_{in} b_{in} \leqq 0 \qquad \text{for all} \quad i,$$
$$2.\quad (|c_{in}| - |c_{jn}|)(|b_{in}| - |b_{jn}|) \geqq 0 \qquad \text{for all} \quad i, j.$$

Set $d_{in} = n^{-\frac{1}{2}} b_{in}$. Let R_{in}^t be the rank of $|X_i - td_{in}|$ among $|X_k - td_{kn}|$,

$k = 1, 2, \cdots, n$. Let $\operatorname{sign}(X_i) = 1$ for $X_i > 0$, $= -1$ for $X_i \leqq 0$. Without confusion R_i^t, c_i, d_i will be used instead of R_{in}^t, c_{in}, d_{in}.

Define:

$$S_n(t) = (n+1)^{-1} \textstyle\sum_1^n c_i \operatorname{sign}(X_i - d_i t) R_i^t \,,$$
$$T_n(t) = S_n(t) - S_n(0) \,,$$
$$Y_n(t) = T_n(t) - E[T_n(t)] \,.$$

1.3 *Theorems.*

THEOREM 1. *Assume that f is continuous except for a finite number of jump discontinuities. Suppose further that $\int f^3(x)\,dx < \infty$. Let $c^2 = 4[\int f^3(x)\,dx - (\int f^2(x)\,dx)^2]$ and*

$$A_n = [\textstyle\sum_1^n c_i^2 d_i^2 + 3(\sum_1^n c_j d_j)^2 (n+1)^{-1}]^{-\frac{1}{2}} \,.$$

Then the process $(A_n Y_n(t))_{t\in[0,1]}$ converges weakly to a process of the form $(tZ)_{t\in[0,1]}$, where Z is a real random variable with normal distribution $N(0, c^2)$.

THEOREM 2. *Suppose that f satisfies the assumptions of Theorem 1 and*

(i) $\lim_{\Delta\to 0} \Delta^{-2} \int\int_{-\Delta}^{\Delta} [f(x-y) - f(x)]^2 \,dy\,dx = 0$.

Let $B_n = 2(\sum_1^n c_i d_i)\|f\|_2^2$. If in the definition of $Y_n(t)$ the quantity $E[T_n(t)]$ is replaced by $-tB_n$ then the conclusion of Theorem 1 holds.

REMARK. Condition (i) of Theorem 2 is satisfied if, for example,

(1) f is such that $|f(x+t) - f(x)| \leqq |t|^\alpha h(x)$, with $\alpha > \frac{1}{2}$ and $h(x) \in L_2(-\infty, +\infty)$,

or

(2) f is absolutely continuous and $f' \in L_2(-\infty, +\infty)$.

1.4 *Preliminaries.* In order to prove Theorem 1 it is enough to show (see Billingsley (1968), Theorem 15.1) that

(A) the finite dimensional distributions of $A_n Y_n(t)$ converge weakly to the corresponding distributions of tZ,

(B) the sequence $\{A_n Y_n\}$ is tight.

To prove relations (A) and (B) two successive approximations of the process $Y_n(t) = T_n(t) - E[T_n(t)]$ are used. First $T_n(t)$ is written as a sum of two processes $U_n(t)$ and $V_n(t)$ (say), the last one being asymptotically negligible. Then the process $U_n(t) - E[U_n(t)]$ is approximated by its projection (in the sense of Hájek (1968)). At this point the main problem consists in finding a bound for the variance of the difference. Such a bound is given in Proposition 1 of the Appendix.

2. The finite dimensional distributions. Consider the process $T_n(t)$ of Section

1.2 and define the following processes:

$$U_n(t) = (n+1)^{-1} \sum_{i,j\neq i}^n c_i\{I(|X_j - td_j| < X_i - td_i) - I(|X_j| < X_i) + I(|X_j| < -X_i) - I(|X_j - td_j| < -X_i + td_i)\},$$

$$V_n(t) = 2(n+1)^{-1} \sum_i^n c_i[I(X_i > td_i) - I(X_i > 0)],$$

$$X_n(t) = U_n(t) - E[U_n(t)].$$

Then with probability one, $T_n(t) = U_n(t) + V_n(t)$ and

$$Y_n(t) = X_n(t) + V_n(t) - E[V_n(t)].$$

We first show that the process $V_n(t)$ is asymptotically negligible.

LEMMA 2.1. *Under the assumptions of Theorem* 1,

$$E[\sup\{|V_n(t)| : t \in [0,1]\}] \to 0 \qquad \textit{as } n \to \infty.$$

PROOF. Since $|d_i t| \leqq Mn^{-\frac{1}{2}}$, for all i and all $t \in [0, 1]$,

$$|I(X_i > td_i) - I(X_i > 0)| \leqq I(-Mn^{-\frac{1}{2}} \leqq X_i \leqq Mn^{-\frac{1}{2}}).$$

Hence

$$\sup\{|V_n(t)| : t \in [0,1]\} \leqq 2M(n+1)^{-1} \sum_i^n I(-Mn^{-\frac{1}{2}} \leqq X_i \leqq Mn^{-\frac{1}{2}}).$$

Therefore,

$$E[\sup\{|V_n(t)| : t \in [0,1]\}] \leqq 2M \int_{-Mn^{-\frac{1}{2}}}^{Mn^{-\frac{1}{2}}} f(x)\,dx.$$

By assumption, $\int f^2(x)\,dx < \infty$ and the lemma follows.

Denote by $\hat{X}_n(t)$ the projection of $X_n(t)$, and for simplicity of notation let $\eta_{ij} = d_i + d_j$, $\gamma_{ij} = d_i - d_j$. Since $X_n(t)$ is centered at expectation, $\hat{X}_n(t) = \sum_i^n E(X_n(t) \mid X_i)$ and it is a matter of easy computation to show that

$$\hat{X}_n(t) = \hat{U}_n(t) - E[\hat{U}_n(t)],$$

where

$$\hat{U}_n(t) = (n+1)^{-1} \sum_{i,j\neq i}^n c_i[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij}) + F(-X_i) - F(X_i)] + (n+1)^{-1} \sum_{i,j\neq i}^n c_j[F(|X_i - td_i| - td_j) - F(|X_i - td_i| + td_j)].$$

According to Proposition 1 of the Appendix, $E(X_n(t) - \hat{X}_n(t))^2 \leqq C|t|n^{-\frac{1}{2}}$.

Since A_n is bounded, the last inequality implies that $A_n|X_n(t) - \hat{X}_n(t)| \to_P 0$ as $n \to \infty$, for every $t \in [0, 1]$. Thus according to Lemma 2.1 the process $A_n Y_n(t)$ satisfies Relation (A) of Section 1.4 if the process $A_n \hat{X}_n(t)$ does.

That $A_n \hat{X}_n(t)$ satisfies (A) is shown next. For this, let

$$W_n = -2(n+1)^{-1} \sum_{i,j\neq i}^n c_i d_i[f(X_i) - ||f||_2^2] - 2(n+1)^{-1} \sum_{i,j\neq i}^n c_j d_j[f(X_i) - ||f||_2^2],$$

$$W_n(t) = tW_n.$$

The variable $A_n W_n$ converges weakly to a variable Z which is normally distributed $N(0, c^2)$ with $c^2 = 4[\int f^3(x)\,dx - (\int f^2(x)\,dx)^2]$. This follows from Assumptions (a)—(c) of Section 1.2. Consequently the finite dimensional distributions of $A_n W_n(t)$ converge weakly to those of the process tZ. The same is true for the process $A_n \hat{X}_n(t)$ because of the following.

LEMMA 2.2. *Under the assumptions of Theorem* 1,

$$E(W_n(t) - \hat{X}_n(t))^2 \to 0 \qquad as \quad n \to \infty\,, \quad for\ every \quad t \in [0, 1]\,.$$

PROOF. Recalling the definition of η_{ij} and γ_{ij} we have:

$$\begin{aligned} W_n(t) = &-t(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i \gamma_{ij}[f(X_i) - \|f\|_2^2] \\ &- t(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i \eta_{ij}[f(X_i) - \|f\|_2^2] \\ &- 2t(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_j d_j[f(X_i) - \|f\|_2^2]\,. \end{aligned}$$

Since the X_i are independent the following inequality can be easily proved:

$$\begin{aligned} &E(W_n(t) - \hat{X}_n(t))^2 \\ &\quad \leqq 3(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i^2 E[F(X_i - t\gamma_{ij}) - F(X_i) + t\gamma_{ij} f(X_i)]^2 \\ &\qquad + 3(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i^2 E[F(-X_i + t\eta_{ij}) - F(-X_i) - t\eta_{ij} f(X_i)]^2 \\ &\qquad + 3(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_j^2 E[F(|X_i - td_i| - td_j) - F(|X_i - td_i| + td_j) \\ &\qquad + 2td_j f(X_i)]^2\,. \end{aligned}$$

Since $|c_i| \leqq M$, for all i, the last expression is smaller than

$$\begin{aligned} &3M^2(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n \int [\int_0^{t\gamma_{ij}} (f(x-y) - f(x))\, dy]^2 f(x)\, dx \\ &\quad + 3M^2(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n \int [\int_0^{t\eta_{ij}} f(-x+y) - f(-x))\, dy]^2 f(x)\, dx \\ &\quad + 6M^2(n+1)^{-1} \textstyle\sum_{i,j\neq i}^n \{\int [\int_{t\gamma_{ij}}^{t\eta_{ij}} (f(-x+y) - f(-x)\, dy]^2 f(x)\, dx \\ &\quad + \int [\int_{t\gamma_{ij}}^{t\eta_{ij}} (f(x-y) - f(x))\, dy]^2 f(x)\, dx\}\,. \end{aligned}$$

Now let $\varepsilon > 0$ be an arbitrary number. By hypothesis,

$$\sup_{1\leqq i,j\leqq n} |t\eta_{ij}| + \sup_{1\leqq i,j\leqq n} |t\gamma_{ij}| \leqq 4Mn^{-\frac{1}{2}}\,,$$

for all $t \in [0, 1]$. Hence we can apply Lemma 2 of the Appendix and for n large enough: $E(W_n(t) - \hat{X}_n(t))^2 \leqq \varepsilon\, 120M^4$. ε being arbitrary, Lemma 2.2 follows.

3. The tightness of the sequence $\{A_n Y_n\}$. Since $A_n Y_n(0) = 0$, according to Theorem 15.5 of Billingsley, the family $\{A_n Y_n\}$ is tight if

(C) $\lim_{m\to\infty} \limsup_n P\{w(A_n Y_n, m^{-1}) > \varepsilon\} = 0$, for every $\varepsilon > 0$.

Letting X_n, $\hat{X}_n$, V_n as defined before, $A_n Y_n(t)$ can be written as $A_n[X_n(t) - \hat{X}_n(t)] + A_n \hat{X}_n(t) + A_n[V_n(t) - EV_n(t)]$. For simplicity write $Z_n(t)$ for $X_n(t) - \hat{X}_n(t)$.

By Lemma 2.1 and since A_n is bounded, in order to prove relation (C) it is enough to show that

(D) $\lim_{m\to\infty} \limsup_n P\{w(\hat{X}_n, m^{-1}) > \varepsilon\} = 0$, for every $\varepsilon > 0$,

and

(E) $\lim_{m\to\infty} \limsup_n P\{w(Z_n, m^{-1}) > \varepsilon\} = 0$, for every $\varepsilon > 0$.

Consider first relation (D). According to Proposition 2 of the Appendix,

$$E(\hat{X}_n(t) - \hat{X}_n(s))^2 \leqq D(t-s)^2\,, \qquad \text{for } t, s \in [0, 1]\,.$$

As it is proved in Billingsley (Theorem 12.3), this implies relation (D).

PROOF OF RELATION (E). Let e_n be the smallest integer greater or equal to $n^{\frac{1}{4}}$. For $j = 0, 1, \cdots, m-1$, $i = 0, 1, \cdots, e_n$, define

$$t_{ij} = m^{-1}(j + e_n^{-1}i), \qquad I_j = [t_{0j}, t_{0(j+1)}], \qquad J_{ij} = [t_{ij}, t_{(i+1)j}].$$

By the triangle inequality, relation (E) holds if

(F) $\max_{0 \leq j \leq m-1} \sup \{|Z_n(t) - Z_n(t_{0j})| : t \in I_j\} \to_P 0$

as $n \to \infty$, and then $m \to \infty$.

Assumption (d) of Section 1.2 is now used to prove (F). Van Eeden has shown that under this assumption, the process $T_n(t)$ is with probability one a monotone step function. Then writing down Z_n in terms of the processes T_n, V_n, U_n and $\hat{X}_n$ the following inequality can be easily proved:

$$\begin{aligned} \sup \{|Z_n(t) - Z_n(t_{0j})| : t \in I_j\} & \\ \leqq 2 \max \{|Z_n(t_{ij}) - Z_n(t_{0j})| : 1 \leqq i \leqq e_n\} & \\ + 2 \sup \{|V_n(t)| : t \in [0, 1]\} + w(\hat{X}_n, m^{-1}) & \\ + \max_{0 \leq i \leq e_n - 1} \sup \{|E[U_n(t) - U_n(s)]| : t, s \in J_{ij}\}. & \end{aligned}$$

We already know that the second and third terms converge to zero in probability as $n \to \infty$ and then $m \to \infty$. So in order to prove (F) it is enough to show that, as $n \to \infty$ and then $m \to \infty$,

(G) $\max_{0 \leq j \leq m-1} \max \{|Z_n(t_{ij}) - Z_n(t_{0j})| : 1 \leqq i \leqq e_n\} \to_P 0,$

and

(K) $\max_{0 \leq i \leq e_n - 1; 0 \leq j \leq m-1} \sup \{|E[U_n(t) - U_n(s)]| : t, s \in J_{ij}\} \to 0.$

PROOF OF RELATION (G). Let C be the constant appearing in Proposition 1 of the Appendix. Apply Lemma II.7 in Antille (1972). For this, replace there R_i by $Z_n(t_{ij}) - Z_n(t_{(i-1)j})$ and u_i by $Cn^{-\frac{1}{2}}m^{-1}e_n^{-1}$, $i = 1, 2, \cdots, e_n$, to get:

$$E[\max \{|S_i| : i = 1, 2, \cdots, e_n\}]^2 \leqq (\log_2 4e_n)^2 Cn^{-\frac{1}{2}}m^{-1}.$$

The conclusion is justified by Proposition 1 of the Appendix. But $S_i = \sum_{k=1}^{i} R_k = Z_n(t_{ij}) - Z_n(t_{0j})$. Therefore using Markov inequality,

$$P\{\max_{0 \leq j \leq m-1} \max \{|Z_n(t_{ij}) - Z_n(t_{0j})| : 1 \leqq i \leqq e_n\} > \varepsilon\} \leqq \varepsilon^{-2}Cn^{-\frac{1}{2}}(\log_2 4e_n)^2.$$

This proves relation (G).

PROOF OF RELATION (K). Let $t, s \in [t_{ij}, t_{(i+1)j}]$ and let U_n^1, U_n^2, P_{ij} be as defined in the proof of Proposition 1 of the Appendix. Then $U_n(t) - U_n(s) = U_n^1 + U_n^2$. In terms of P_{ij}, U_n^1 can be written as $(n+1)^{-1} \sum_{i,j \neq i}^{n} c_i P_{ij}$. For all i, c_i is bounded by M. Hence $|E(U_n^1)| \leqq (n+1)^{-1} M \sum_{i,j \neq i}^{n} E(|P_{ij}|)$. According to Lemma 4.1 (Appendix), $E(P_{ij})^2 \leqq 4M\|f\|_2^2 |t - s| n^{-\frac{1}{2}}$, for all i, j. Since P_{ij} is a difference of two indicator functions, the same bound applies to $E(|P_{ij}|)$ and therefore $|E(U_n^1)| \leqq 4M^2\|f\|_2^2 n^{-\frac{1}{2}}|t - s|$. But $|t - s| \leqq m^{-1}e_n^{-1}$ for $t, s \in [t_{ij}, t_{(i+1)j}]$. This implies $|E(U_n^1)| \leqq 4M^2\|f\|_2^2 m^{-1}$. The same is true for U_n^2 and relation (K) is proved.

4. Proof of Theorem 2. Let A_n, B_n be as defined in Theorems 1 and 2. In view of Lemma 2.1 it is sufficient to show that

$$\sup\{|E[U_n(t)] + tB_n| : t \in [0, 1]\} \to 0 \qquad \text{as } n \to \infty .$$

By easy calculation we get:

$$\begin{aligned} E[U_n(t)] &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^{n} c_i E[F(X_i - t\gamma_{ij}) - F(X_i)] \\ &\quad + (n+1)^{-1} \textstyle\sum_{i,j\neq i}^{n} c_i E[F(-X_i) - F(-X_i + t\eta_{ij})] \\ &= E_1(t) + E_2(t) \quad \text{(say)}. \end{aligned}$$

Then tB_n can be written as

$$t(n+1)^{-1} \textstyle\sum_{i,j\neq i}^{n} c_i\gamma_{ij}\|f\|_2^2 + t(n+1)^{-1} \textstyle\sum_{i,j\neq i}^{n} c_i\eta_{ij}\|f\|_2^2 = tB_{n1} + tB_{n2} \quad \text{(say)}.$$

Now we only prove that

$$\sup\{|E_1(t) + tB_{n1}| : t \in [0, 1]\} \to 0 \qquad \text{as } n \to \infty ,$$

the proof for $E_2(t) + tB_{n2}$ being similar.

Let $\Gamma_{ij}(t) = E[F(X_i - t\gamma_{ij}) - F(X_i) + t\gamma_{ij} f(X_i)]$ and assume (without loss of generality) that $\gamma_{ij} > 0$. Then

$$\begin{aligned} |\Gamma_{ij}(t)| &= \left|\int \left[\int_0^{t\gamma_{ij}} (f(x) - f(x-y))\, dy\right] f(x)\, dx\right| \\ &= 2^{-1} \int\int_0^{t\gamma_{ij}} [f(x-y) - f(x)]^2\, dy\, dx . \end{aligned}$$

By assumption, $t\gamma_{ij} \leqq 2Mn^{-\frac{1}{2}}$ and therefore

$$|\Gamma_{ij}(t)| \leqq 2^{-1} \int\int_{-\Delta_n}^{\Delta_n} [f(x-y) - f(x)]^2\, dy\, dx ,$$

where $\Delta_n = 2Mn^{-\frac{1}{2}}$. The last inequality is valid for any $t \in [0, 1]$ and all i, j. Hence

$$\begin{aligned} |E_1(t) + tB_{n1}| &\leqq (n+1)^{-1} \textstyle\sum_{i,j\neq i}^{n} |c_i||\Gamma_{ij}(t)| \\ &\leqq Mn2^{-1} \int\int_{-\Delta_n}^{\Delta_n} [f(x-y) - f(x)]^2\, dy\, dx . \end{aligned}$$

for all $t \in [0, 1]$.

By hypothesis (i) of Theorem 2 the last expression tends to zero and the proof is complete.

APPENDIX

As we showed before the process $Y_n(t)$ can be written as $Z_n(t) + \hat{X}_n(t) + [V_n(t) - EV_n(t)]$, with $Z_n(t) = X_n(t) - \hat{X}_n(t)$ and $[V_n(t) - E(V_n(t))]$ asymptotically negligible. The basic result in this section is Proposition 1, where a bound for the variance of $Z_n(t) - Z_n(s)$ is given. A similar result for the process $\hat{X}_n$ is contained in Proposition 2.

PROPOSITION 1. *Let* $L = \sup\{|f(x)| : x \in [-1, +1]\}$. *Then, under the assumptions of Theorem* 1, *for any* $t, s \in [0, 1]$ *and* n *large enough,*

$$E(Z_n(t) - Z_n(s))^2 \leqq C|t - s|n^{-\frac{1}{2}} ,$$

where $C = 400M^3\|f\|_2^2 + 800L^2M^4 + 3200M^4 \int f^3(x)\, dx$.

PROPOSITION 2. *Let L be as in Proposition* 1. *Then under the assumptions of Theorem* 1,

$$E[\hat{X}_n(t) - \hat{X}_n(s)]^2 \leqq D(t-s)^2 , \qquad \textit{for any} \quad t, s \in [0, 1] ,$$

where $D = 80M^2[8L^2M^2 + 32M^2 \int f^3(x)\, dx]$.

PROOF OF PROPOSITION 1. Let t, s be fixed. Then $U_n(t) - U_n(s)$ can be written as $U_n^1 + U_n^2$, where

$$\begin{aligned} U_n^1 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i[I(|X_j - td_j| < X_i - td_i) \\ &\qquad - I(|X_j - sd_j| < X_i - sd_i)] , \\ U_n^2 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i[I(|X_j - sd_j| < -X_i + sd_i) \\ &\qquad - I(|X_j - td_j| < -X_i + td_i)] \end{aligned}$$

and $\hat{U}_n(t) - \hat{U}_n(s)$ as $\hat{U}_n^1 + \hat{U}_n^2$, where

$$\begin{aligned} \hat{U}_n^1 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i\{I(X_i > td_i)[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij})] \\ &\qquad - I(X_i > sd_i)[F(X_i - s\gamma_{ij}) - F(-X_i + s\eta_{ij})]\} \\ &\qquad + (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_j[F(|X_i - sd_i| + sd_j) - F(|X_i - td_i| + td_j)] , \\ \hat{U}_n^2 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i\{I(X_i < td_i)[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij})] \\ &\qquad - I(X_i < sd_i)[F(X_i - s\gamma_{ij}) - F(-X_i + s\eta_{ij})]\} \\ &\qquad + (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_j[F(|X_i - td_i| - td_j) - F(|X_i - sd_i| - sd_j)] . \end{aligned}$$

Now by definition:

$$\begin{aligned} &Z_n(t) - Z_n(s) \\ &\quad = [U_n^1 - \hat{U}_n^1 - E(U_n^1 - \hat{U}_n^1)] + [U_n^2 - \hat{U}_n^2 - E(U_n^2 - \hat{U}_n^2)] \\ &\quad = Z_{n1} + Z_{n2} \quad \text{(say)}. \end{aligned}$$

Clearly, Proposition 1 is proved if we show that

$$E(Z_{ni})^2 \leqq C/4|t-s|n^{-\frac{1}{2}} , \quad i = 1, 2 .$$

The proofs for $i = 1, 2$, being similar consider the case $i = 1$ (say).

The process $\hat{U}_n^1 - E\hat{U}_n^1(\hat{U}_n^2 - E\hat{U}_n^2)$ is the projection of $U_n^1 - EU_n^1(U_n^2 - EU_n^2)$. Therefore, $E(Z_{n1})^2 = \operatorname{Var}(U_n^1) - \operatorname{Var}(\hat{U}_n^1)$. Next we give explicit formulas for the variances. To simplify the notation define the following new quantities. Let

$$\begin{aligned} P_{ij} &= I(|X_j - td_j| < X_i - td_i) - I(|X_j - sd_j| < X_i - sd_i) , \\ G_{ij} &= I(X_i > td_i)[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij})] \\ &\qquad - I(X_i > sd_i)[F(X_i - s\gamma_{ij}) - F(-X_i + s\eta_{ij})] , \\ H_{ij} &= F(|X_i - sd_i| + sd_j) - F(|X_i - td_i| + td_j) . \end{aligned}$$

In terms of these new variables,

$$\begin{aligned} U_n^1 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i P_{ij} , \\ \hat{U}_n^1 &= (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_i G_{ij} + (n+1)^{-1} \textstyle\sum_{i,j\neq i}^n c_j H_{ij} . \end{aligned}$$

A direct computation shows that:

$$\begin{aligned}
\operatorname{Var}(U_n^1) &= 2(n+1)^{-2} \textstyle\sum_{i,j\neq i,1\neq i}^{n} c_i^2 \operatorname{Cov}(P_{ij}, P_{i1}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{i,j\neq i,k\neq i}^{n} c_i c_k \operatorname{Cov}(P_{ij}, P_{ki}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{k,i\neq k,1\neq k}^{n} c_i c_k \operatorname{Cov}(P_{ik}, P_{k1}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{j,i\neq j,k\neq j}^{n} c_i c_k \operatorname{Cov}(P_{ij}, P_{kj}) \\
\operatorname{Var}(\hat{U}_n^1) &= 2(n+1)^{-2} \textstyle\sum_{i,j\neq i,1\neq i}^{n} c_i^2 \operatorname{Cov}(G_{ij}, G_{i1}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{i,j\neq i,k\neq i}^{n} c_i c_k \operatorname{Cov}(G_{ij}, H_{ik}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{i,j\neq i,1\neq i}^{n} c_i c_j \operatorname{Cov}(H_{ij}, G_{i1}) \\
&\quad + 2(n+1)^{-2} \textstyle\sum_{j,i\neq j,k\neq i}^{n} c_i c_k \operatorname{Cov}(H_{ji}, H_{jk}) .
\end{aligned}$$

The following lemma gives bounds for the covariances appearing in the above sum.

LEMMA 1. *Let P_{ij}, G_{ij}, H_{ij} be as defined above. Let* $L = \sup\{|f(x)| : x \in [-1, 1]\}$. *Then under the assumptions of Theorem* 1:

(1) $E(P_{ij}^2) \leqq 4M\|f\|_2^2|t-s|n^{-\frac{1}{2}}$
(2) $E(G_{ij}^2) \leqq (8L^2M^2 + 32M^2 \int f^3(x)\,dx)(t-s)^2n^{-1}$,
(3) $E(H_{ij}^2) \leqq 32M^2(\int f^3(x)\,dx)(t-s)^2n^{-1}$,

for all $1 \leqq i, j \leqq n$ *and n large enough.*

PROOF. (1) Let i, j be fixed and define:

$$\alpha = |X_j - td_j| + td_i\,, \qquad \beta = |X_j - sd_j| + sd_i\,, \qquad \gamma = |t-s|n^{-\frac{1}{2}}M\,.$$

Then

$$\begin{aligned}
E(P_{ij}^2) &= E[I(\alpha < X_i) - I(\beta < X_i)]^2 \leqq E[I(\min\{\alpha,\beta\} \leqq X_i \leqq \max\{\alpha,\beta\})] \\
&\leqq E[I(\alpha < X_i < \alpha+\gamma)] + E[I(\beta < X_i < \beta+\gamma)]\,.
\end{aligned}$$

Therefore,

$$E(P_{ij}^2) \leqq E[F(\alpha+\gamma) - F(\alpha)] + E[F(\beta+\gamma) - F(\beta)]\,.$$

Now

$$\begin{aligned}
E[F(\alpha+\gamma) - F(\alpha)] &= E[I(X_j < td_j)(F(\alpha+\gamma) - F(\alpha))] \\
&\quad + E[I(X_j \geqq td_j)(F(\alpha+\gamma) - F(\alpha))] \\
&= \int_{-\infty}^{td_j} [F(-x + t\eta_{ij} + \gamma) - F(-x + t\eta_{ij})] f(x)\,dx \\
&\quad + \int_{td_j}^{\infty} [F(x + t\gamma_{ij} + \gamma) - F(x + t\gamma_{ij})] f(x)\,dx \\
&\leqq \int \Big[\int_0^{\gamma} f(-x + t\eta_{ij} + y)\,dy\Big] f(x)\,dx \\
&\quad + \int \Big[\int_0^{\gamma} f(x + t\gamma_{ij} + y)\,dy\Big] f(x)\,dx\,.
\end{aligned}$$

Apply the Fubini theorem and the Schwarz inequality to show that the last expression is smaller than $2\gamma\|f\|_2^2$. The same result holds for $E[F(\beta+\gamma) - F(\beta)]$ and inequality (1) follows.

(2) Let i, j be fixed and suppose without loss of generality that $td_i < sd_i$.

Then by definition,

$$\begin{aligned} G_{ij} &= I(td_i < X_i < sd_i)[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij})] \\ &\quad + I(X_i > sd_i)[F(X_i - t\gamma_{ij}) - F(-X_i + t\eta_{ij}) \\ &\quad - F(X_i - s\gamma_{ij}) + F(-X_i + s\eta_{ij})] \\ &= G_{ij}^1 + G_{ij}^2 \quad \text{(say)}. \end{aligned}$$

By the Schwarz inequality, $E(G_{ij}^2) \leqq 2E(G_{ij}^1)^2 + 2E(G_{ij}^2)^2$. So, it is sufficient to give bounds for $E(G_{ij}^k)^2$, $k = 1, 2$. We begin with G_{ij}^1.

By assumption, the density is symmetric around the origin. Since there is only a finite number of jump discontinuities, f must be continuous on some symmetric interval around zero. Then on this interval, the distribution function F is derivable with derivate f. Use this to obtain:

$$E(G_{ij}^1)^2 \leqq E[I(td_i < X_i < sd_i)4(X_i - td_i)^2 L^2]\,, \qquad \text{for} \quad n \quad \text{large enough.}$$

The right-hand side of the inequality is smaller than $4L^2(s - t)^2 d_i^{\,2}$. Hence $E(G_{ij}^1)^2 \leqq 4L^2M^2(s - t)^2 n^{-1}$, since $|d_i| \leqq Mn^{-\frac{1}{2}}$ for all i. Consider now the second term G_{ij}^2. By the Schwarz inequality,

$$\begin{aligned} E(G_{ij}^2)^2 &\leqq 2E[F(X_i - t\gamma_{ij}) - F(X_i - s\gamma_{ij})]^2 \\ &\quad + 2E[F(-X_i + t\Delta_{ij}) - F(-X_i + s\Delta_{ij})]^2\,. \end{aligned}$$

Suppose without loss of generality that $\gamma_{ij} > 0$ and $t > s$. Then using the Schwarz and Hölder inequalities,

$$\begin{aligned} &E[F(X_i - t\gamma_{ij}) - F(X_i - s\gamma_{ij})]^2 \\ &\qquad = \int [\textstyle\int_{s\gamma_{ij}}^{t\gamma_{ij}} f(x - y)\, dy]^2 f(x)\, dx \\ &\qquad = \int [\textstyle\int_{s\gamma_{ij}}^{t\gamma_{ij}} f(x - y)\, dy][\int_{s\gamma_{ij}}^{t\gamma_{ij}} f(x - z)\, dz] f(x)\, dx \\ &\qquad \leqq \textstyle\int_{s\gamma_{ij}}^{t\gamma_{ij}} \int_{s\gamma_{ij}}^{t\gamma_{ij}} [\int f(x - y) f(x - z) f(x)\, dx]\, dy\, dz \\ &\qquad \leqq \textstyle\int_{s\gamma_{ij}}^{t\gamma_{ij}} \int_{s\gamma_{ij}}^{t\gamma_{ij}} [\int f^2(x - y) f(x)\, dx]^{\frac{1}{2}} [\int f^2(x - z) f(x)\, dx]^{\frac{1}{2}}\, dy\, dz \\ &\qquad = (\textstyle\int_{s\gamma_{ij}}^{t\gamma_{ij}} [\int f^2(x - y) f(x)\, dx]^{\frac{1}{2}}\, dy)^2 \\ &\qquad \leqq (t - s)^2 \gamma_{ij}^2 \textstyle\int f^3(x)\, dx \\ &\qquad \leqq 4M^2(\textstyle\int f^3(x)\, dx)(t - s)^2 n^{-1}\,, \end{aligned}$$

since by definition $\gamma_{ij} = d_i - d_j$.

The same result holds for $E[F(-X_i + t\eta_{ij}) - F(-X_i + s\eta_{ij})]^2$. Hence $E(G_{ij}^2)^2 \leqq 16M^2(\int f^3(x)\, dx)(t - s)^2 n^{-1}$ and inequality (2) follows.

(3) Let i, j fixed and define:

$$\alpha = |t - s|\,|d_i| + sd_j\,, \qquad \beta = |t - s|\,|d_i| + td_j\,.$$

Then

$$\begin{aligned} E(H_{ij})^2 &= E[F(|X_i - sd_i| + sd_j) - F(|X_i - td_i| + td_j)]^2 \\ &\leqq 2E[F(|X_i - td_i| + \alpha) - F(|X_i - td_i| + td_j)]^2 \\ &\quad + 2E[F(|X_i - sd_i| + \beta) - F(|X_i - sd_i| + sd_j)]^2\,. \end{aligned}$$

The first term is smaller than

$$2 \int [\int_{td_j}^{\alpha} f(-x + td_i + y)\, dy]^2 f(x)\, dx + 2 \int [\int_{td_j}^{\alpha} f(x - td_i + y)\, dy]^2 f(x)\, dx \leqq 16M^2(\int f^3(x)\, dx)(t - s)^2 n^{-1} .$$

(See the proof of inequality (2).)

The second term is bounded by the same quantity. Hence $E(H_{ij})^2 \leqq 32M^2(\int f^3(x)\, dx)(t - s)^2 n^{-1}$ and the proof of Lemma 1 is complete.

We now return to the variance $U_n^1 - \hat{U}_n^1$. This variance is equal to the difference of the variances. Use the explicit formulas we gave before to show that:

$$\begin{aligned}\operatorname{Var}(U_n^1 - \hat{U}_n^1) \leqq\ & 4(n + 1)^{-2} \textstyle\sum_{i,j\neq i}^n c_i^2 \operatorname{Var}(P_{ij}) \\ & + 4(n + 1)^{-2} \textstyle\sum_{i,j\neq i}^n |c_i|\,|c_j|\,|\operatorname{Cov}(P_{ij}, P_{ji})| \\ & + 4(n + 1)^{-2} \textstyle\sum_{k,i\neq k}^n |c_i|\,|c_k|\,|\operatorname{Cov}(P_{ik}, P_{ki})| \\ & + 4(n + 1)^{-2} \textstyle\sum_{i,i\neq j}^n c_i^2 \operatorname{Var}(P_{ij}) \\ & + 4(n + 1)^{-2} \textstyle\sum_{i,j\neq i}^n c_i^2 \operatorname{Var}(G_{ij}) \\ & + 4(n + 1)^{-2} \textstyle\sum_{i,j\neq i}^n |c_i|\,|c_j|\,|\operatorname{Cov}(G_{ij}, H_{ij})| \\ & + 4(n + 1)^{-2} \textstyle\sum_{i,k\neq i}^n |c_i|\,|c_k|\,|\operatorname{Cov}(H_{ki}, G_{ik})| \\ & + 4(n + 1)^{-2} \textstyle\sum_{j,i\neq j}^n c_i^2 \operatorname{Var}(H_{ji}) .\end{aligned}$$

By Schwarz's inequality and Lemma 4.1, the last sum is smaller than

$$80M^3\|f\|_2^2|t - s|n^{-\frac{1}{2}} + 160M^4L^2(t - s)^2 n^{-1} + 800M^4(\int f^3(x)\, dx)(t - s)^2 n^{-1} ,$$

and Proposition 1 follows.

PROOF OF PROPOSITION 2. Let t, s be fixed and $\hat{U}_n^1$, $\hat{U}_n^2$ as defined in the proof of Proposition 1. By the Schwarz inequality,

$$\operatorname{Var}(\hat{U}_n(t) - \hat{U}_n(s)) \leqq 2 \operatorname{Var}(\hat{U}_n^1) + 2 \operatorname{Var}(\hat{U}_n^2) .$$

Use the formula we have for $\operatorname{Var}(\hat{U}_n^1)$, to show that,

$$\begin{aligned}\operatorname{Var}(\hat{U}_n^1) \leqq\ & 2(n + 1)^{-2} \textstyle\sum_{i,j\neq i,1\neq i}^n c_i^2[E(G_{ij})^2 E(G_{i1})^2]^{\frac{1}{2}} \\ & + 2(n + 1)^{-2} \textstyle\sum_{i,j\neq i,k\neq i}^n |c_i|\,|c_k|[E(G_{ij})^2 E(H_{ki})^2]^{\frac{1}{2}} \\ & + 2(n + 1)^{-2} \textstyle\sum_{i,j\neq i,1\neq i}^n |c_i|\,|c_j|[E(H_{ij})^2 E(G_{i1})^2]^{\frac{1}{2}} \\ & + 2(n + 1)^{-2} \textstyle\sum_{i,i\neq j,k\neq j}^n |c_i|\,|c_k|[E(H_{ij})^2 E(H_{jk})^2]^{\frac{1}{2}} .\end{aligned}$$

By Lemma 4.1 the last sum is not greater than

$$16M^2[8L^2M^2 + 32M^2 \int f^3(x)\, dx](t - s)^2 .$$

The same result holds for $\operatorname{Var}(\hat{U}_n^2)$ and Proposition 2 follows.

We now close the Appendix by proving a result used in the proof of Lemma 2.2.

LEMMA 2. *Let $\varepsilon > 0$ be an arbitrary number. Then under the assumptions of Theorem* 1, *there exists an η such that*:

$$\int [\int_{\Delta_1}^{\Delta_2} (f(x + y) - f(x))\, dy]^2 f(x)\, dx \leqq \varepsilon(\Delta_1^2 + \Delta_2^2) ,$$

if $|\Delta_i| \leqq \eta$ for $i = 1, 2$.

PROOF. Consider the case where f has only one jump discontinuity at the point x_0 (say). If f has more but a finite number of such discontinuities, a similar proof can be carried out. Suppose further (without loss of generality) that $\Delta_2 > \Delta_1$.

Use the Fubini theorem and the Schwarz inequality to obtain:

$$\text{(a)} \qquad \int \left[\int_{\Delta_1}^{\Delta_2} f(x+y) - f(x))\, dy\right]^2 f(x)\, dx \\ \leqq (\Delta_2 - \Delta_1) \int_{\Delta_1}^{\Delta_2} \left[\int (f(x+y) - f(x))^2 f(x)\, dx\right] dy\,.$$

Since $\int f^3(x)\, dx < \infty$ there exists an $N (N > |x_0|)$ such that

$$\text{(b)} \qquad \int_N^\infty f^3(x)\, dx < \varepsilon\,.$$

Given this N we can find an $\eta > 0$ ($\eta < N$) and a $\delta > 0$ such that

$$\text{(c)} \qquad \int_{x_0-\delta}^{x_0+\delta} (f(x+y) - f(x))^2 f(x)\, dx \leqq \varepsilon \quad \text{for} \quad |y| \leqq \eta\,, \qquad \text{and}$$

$$\text{(d)} \quad \sup_{|y| \leqq \eta} |f(x+y) - f(x)|^2 \leqq \varepsilon \quad \text{for} \quad x \in [-2N, x_0 - \delta] \cup [x_0 + \delta, 2N]\,.$$

Now write $\int [f(x+y) - f(x)]^2 f(x)\, dx$ as a sum of integrals over the intervals $(-\infty, -2N)$, $(-2N, x_0 - \delta)$, $(x_0 - \delta, x_0 + \delta)$, $(x_0 + \delta, 2N)$, $(2N, \infty)$. Let $|y| \leqq \eta$. Use the Hölder inequality (with $p = 3$, $q = \frac{3}{2}$) to show that the sum of the first and fifth integrals is dominated by 8ε. By (c) the third integral is smaller than ε and by (d) the sum of the second and fourth integrals is not greater than 2ε. It follows that for $|\Delta_i| \leqq \eta$, $i = 1, 2$, the left-hand side of inequality (a) is smaller than $(\Delta_2 - \Delta_1)^2\, 11\varepsilon$ and Lemma 2 is proved.

REFERENCES

ANTILLE, A. (1972). Linéarité asymptotique d'une statistique de rang. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** 309–324.

ANTILLE, A. (1974). A linearized version of the Hodges-Lehmann estimator. *Ann. Statist.* **2** 1308–1313.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

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

JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

JUREČKOVÁ, J. (1973). Central limit theorem for Wilcoxon rank statistics process. *Ann. Statist.* **1** 1046–1060.

KRAFT, C. H. and VAN EEDEN, C. (1972). Linearized rank estimates and signed-rank estimates for the general linear hypothesis. *Ann. Math. Statist.* **43** 42–57.

KOUL, H. L. (1969). Asymptotic behavior of Wilcoxon type confidence regions in multiple linear regression. *Ann. Math. Statist.* **40** 1950–1979.

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

VAN EEDEN, C. (1972). An analogue, for signed-rank statistics of Jurečková's asymptotic linearity theorem for rank statistics. *Ann. Math. Statist.* **43** 791–802.

JOHANN WOLFGANG GOETHE UNIVERSITÄT
FACHBEREICH MATHEMATIK
ROBERT MAYER STRASSE 6–10
6 FRANKFURT AM MAIN
WEST GERMANY