The Annals of Statistics
1973, Vol. 1, No. 6, 1046–1060

CENTRAL LIMIT THEOREM FOR WILCOXON RANK STATISTICS PROCESS

BY JANA JUREČKOVÁ

Charles University, Prague

The rank statistics $S_{\Delta N} = N^{-1} \sum_{i=1}^{N} c_{Ni} R_{Ni}^{\Delta}$, with R_{Ni}^{Δ} being the rank of $X_{Ni} + \Delta d_{Ni}$, $i = 1, 2, \cdots, N$ and $X_{N1}, \cdots, X_{NN}$ being the random sample from the basic distribution with density function f, are considered as a random process with Δ in the role of parameter. Under some assumptions on c_{Ni}'s, d_{Ni}'s and on the underlying distribution, it is proved that the process $\{S_{\Delta N} - S_{0N} - ES_{\Delta N}; 0 \leqq \Delta \leqq 1\}$, being properly standardized, converges weakly to the Gaussian process with covariances proportional to the product of parameter values. Under additional assumptions, Δb_N can be written instead of $ES_{\Delta N}$, where $b_N = \sum_{i=1}^{N} c_{Ni} d_{Ni} \int f^2(x)\, dx$. As an application, this result yields the asymptotic normality of the standardized form of the length of a confidence interval for regression coefficient based on statistic $S_{\Delta N}$.

1. Summary and introduction. Let $(X_1, X_2, \cdots, X_N)$ be an independent random sample from a distribution with finite Fisher's information and let us consider the statistics

$$S_{\Delta N} = N^{-1} \sum_{i=1}^{N} c_i R_{Ni}^{\Delta}$$

where $R_{N1}^{\Delta}, R_{N2}^{\Delta}, \cdots, R_{NN}^{\Delta}$ is the vector of ranks for random variables $X_1 + \Delta d_1$, $X_2 + \Delta d_2, \cdots, X_N + \Delta d_N$; Δ, c_i and d_i, $i = 1, \cdots, N$ are real constants. Then $\{S_{\Delta N}; 0 \leqq \Delta \leqq 1\}$ can be considered as a random process with realizations being right-continuous functions of Δ. The residuals $S_{\Delta N} - S_{0N}$ for more general scores were investigated recently and their uniform asymptotic linearity in Δ was proved (see [7]); this result is of a law of large numbers type. This result was then extended by this and by other authors to the multiparameter case, to the stronger type of convergence and to other types of rank test statistics. The results are applicable in estimation theory, where they represent the main tool for proving the asymptotic normality of the estimators of regression coefficients based on rank tests statistics (see [8] and [9]).

The present paper represents a further step in the investigations of residuals $S_{\Delta N} - S_{0N}$; the main result is that the process $\{S_{\Delta N} - S_{0N}; 0 \leqq \Delta \leqq 1\}$, being properly standardized, converges weakly to Gaussian process with covariances proportional to the product of parameter values. The result can be extended without difficulties for $\Delta \in [-M, M]$. The author succeeded in proving the weak convergence for Wilcoxon rank statistics only, but she believes a similar technique

Received June 1972; revised February 1973.

Key words and phrases. Nonparametrics, Wilcoxon rank test statistic against regression alternatives, asymptotic behavior of rank test statistics as a function of regression parameter, asymptotic distribution of nonparametric estimate of regression coefficient, space $D[0, 1]$ of right-continuous functions.

would yield analogous results for a wider class of score-generating functions, at least of those with bounded second or third derivative; the difficulties to be overcome are rather of a technical type. Similar theorems would hold also for signed-rank statistics.

As an application, a confidence interval for regression parameter based on statistic $S_{\Delta N}$ is considered and the asymptotic normality of the standardized form of its length is deduced.

2. Notation and basic assumptions. We shall consider for any positive integer N:

1° an independent random sample $(X_{N1}, X_{N2}, \cdots, X_{NN})$ from a distribution whose cdf F has finite Fisher's information, i.e.

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

where f is the density of the distribution;

2° a real vector $(c_{N1}, c_{N2}, \cdots, c_{NN})$ of regression constants such that

$$\sum_{i=1}^{N} c_{Ni} = 0\,, \qquad \sum_{i=1}^{N} c_{Ni}^2 = 1\,, \tag{2.1}$$

$$\lim_{N\to\infty} \max_{1\leqq i\leqq N} c_{Ni}^2 = 0\,; \tag{2.2}$$

3° a real vector $(d_{N1}, d_{N2}, \cdots, d_{NN})$ such that

$$\sum_{i=1}^{N} d_{Ni} = 0\,, \qquad \sum_{i=1}^{N} d_{Ni}^2 = 1\,, \tag{2.3}$$

$$\lim_{N\to\infty} \max_{1\leqq i\leqq N} d_{Ni}^2 = 0\,. \tag{2.4}$$

4° We suppose that the vectors $(c_{N1}, \cdots, c_{NN})$ and $(d_{N1}, \cdots, d_{NN})$ are such that

$$\lim_{N\to\infty} \left(\sum_{i=1}^{N} c_{Ni} d_{Ni}\right) = b^2 > 0 \tag{2.5}$$

and that

$$\lim_{N\to\infty} \left[\max_{1\leqq i\leqq N} (c_{Ni}^2 d_{Ni}^2)\left(\sum_{i=1}^{N} c_{Ni}^2 d_{Ni}^2\right)^{-1}\right] = 0\,. \tag{2.6}$$

5° Δ is a real parameter, $0 \leqq \Delta \leqq 1$.

6° Let

$$R_{Ni}^{\Delta} = \sum_{j=1}^{N} u(X_{Ni} - X_{Nj} + \Delta(d_{Ni} - d_{Nj})) \tag{2.7}$$

where

$$\begin{aligned} u(x) &= 1 \quad \text{for} \quad x \geqq 0 \\ &= 0 \quad \text{for} \quad x < 0\,. \end{aligned}$$

REMARK. For simplicity of notation, we shall omit indices N in X_{Ni}, c_{Ni}, d_{Ni} and R_{Ni}^{Δ} in the sequel; we hope that this simplification will not cause confusion.

We shall consider the statistics

$$S_{\Delta N} = N^{-1} \sum_{i=1}^{N} c_i R_i^{\Delta}\,. \tag{2.8}$$

Let us denote

$$A_N^2 = \sum_{i=1}^{N} c_i^2 d_i^2 + 3N^{-1}\left(\sum_{i=1}^{N} c_i d_i\right)^2\,. \tag{2.9}$$

We are prepared to formulate the first result of the paper.

THEOREM 2.1. *Under assumptions* 1° *through* 6°, *the random process*

$$(2.10)\qquad \begin{aligned}\mathscr{S}_{\Delta N} &= A_N^{-1}[S_{\Delta N} - S_{0N} - ES_{\Delta N}] \\ &= A_N^{-1}[S_{\Delta N} - S_{0N} - N^{-1} \textstyle\sum_{i=1}^N c_i \sum_{j=1}^N \int [F(x + \Delta(d_i - d_j)) \\ &\qquad - F(x)]\, dF(x)]\end{aligned}$$

converges weakly for $\Delta \in [0, 1]$ *to the Gaussian process* $\{G(\Delta) : 0 \leqq \Delta \leqq 1\}$ *with*

$$(2.11)\qquad EG(\Delta) = 0\,, \qquad \operatorname{Cov}(G(\Delta_1), G(\Delta_2)) = C\Delta_1\Delta_2$$

for any $0 \leqq \Delta_1 \leqq \Delta_2 \leqq 1$, *where*

$$(2.18)\qquad C = \int f^3(x)\, dx - (\int f^2(x)\, dx)^2\,.$$

3. Proof of Theorem 2.1. It follows from the proof of Theorem 2.1 of [7] that $S_{\Delta N}$ are for $N = 1, 2, \cdots$ step functions with probability 1, so that their definition may be completed at the points of discontinuity so as to be right-continuous. The realizations of $S_{\Delta N}$ and also of $\mathscr{S}_{\Delta N}$ then will belong to the space $D[0, 1]$ of functions on $[0, 1]$ that are right-continuous and have left-hand limits. By Theorem 15.1 of [3], $\{\mathscr{S}_{\Delta N} : 0 \leqq \Delta \leqq 1\}$ converges in distribution to $\{G(\Delta) : 0 \leqq \Delta \leqq 1\}$ if the sequence of distributions of $\mathscr{S}_{\Delta N}$ is tight and if all finite-dimensional distributions of $\mathscr{S}_{\Delta N}$ are asymptotically normal with corresponding parameters. For this, the most convenient sufficient condition is provided by Theorem 15.4 of [3].

The proof of Theorem 2.1 thus consists of two steps. The proof of the asymptotic normality of finite-dimensional distributions is based on Hájek's projection method (see [4]) and the proof of tightness is based on an extension of Theorem 12.1 of [3].

(i) Proof of asymptotic normality of finite-dimensional distributions of $\mathscr{S}_{\Delta N}$.

Let us denote for any fixed Δ and any fixed N

$$(3.1)\qquad \hat{S}_{\Delta N} = \textstyle\sum_{i=1}^N E(S_{\Delta N} \mid X_i) - (N-1)ES_{\Delta N}\,.$$

For any fixed positive integer K, let $(\lambda_1, \lambda_2, \cdots, \lambda_K)$ be any vector of real numbers and $(\Delta_1, \Delta_2, \cdots, \Delta_K)$ be the vector of parameter values such that $0 \leqq \Delta_1 < \Delta_2 < \cdots < \Delta_K \leqq 1$. Let

$$(3.2)\qquad Z_N = \textstyle\sum_{\nu=1}^K \lambda_\nu (S_{\Delta_\nu N} - S_{0N})$$

and

$$(3.3)\qquad \hat{Z}_N = \textstyle\sum_{\nu=1}^K \lambda_\nu (\hat{S}_{\Delta_\nu N} - \hat{S}_{0N})\,.$$

By Lemma 4.1 of [4], we have

$$(3.4)\qquad E\hat{Z}_N = EZ_N = \textstyle\sum_{\nu=1}^K \lambda_\nu ES_{\Delta_\nu N}$$

and

$$(3.5)\qquad E(Z_N - \hat{Z}_N)^2 = \operatorname{Var} Z_N - \operatorname{Var} \hat{Z}_N\,.$$

If we are able to prove

$$\lim_{N\to\infty}[(\operatorname{Var} Z_N - \operatorname{Var} \hat{Z}_N)(\operatorname{Var} \hat{Z}_N)^{-1}] = 0 \tag{3.6}$$

then the asymptotic normality of $\hat{Z}_N$ with parameters $(E\hat{Z}_N, \operatorname{Var} \hat{Z}_N)$ will imply the asymptotic normality of Z_N with the same parameters.

First of all, let us compute $\operatorname{Var} Z_N$. We have

$$\begin{aligned}\operatorname{Var} Z_N = {} & \sum_{\nu=1}^{K} \lambda_\nu^2 \operatorname{Var}(S_{\Delta_\nu N} - S_{0N}) \\ & + \sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{K} \sum_{\mu=1}^{K} \lambda_\nu \lambda_\mu \operatorname{Cov}(S_{\Delta_\nu N} - S_{0N}, S_{\Delta_\mu N} - S_{0N}) .\end{aligned} \tag{3.7}$$

Let us denote for fixed Δ

$$B_{ij} = u(X_i - X_j + \Delta(d_i - d_j)) - u(X_i - X_j) ; \qquad i, j = 1, 2, \cdots, N \tag{3.8}$$

and

$$H_i = \sum_{j=1}^{N} B_{ij} . \tag{3.9}$$

We may then write

$$\begin{aligned}\operatorname{Var}(S_{\Delta N} - S_{0N}) = {} & N^{-2} \sum_{i=1}^{N} c_i^2 \operatorname{Var} H_i \\ & + N^{-2} \sum_{\substack{j=1 \\ j\neq k}}^{N} \sum_{k=1}^{N} c_j c_k \operatorname{Cov}(H_j, H_k) .\end{aligned} \tag{3.10}$$

It holds that

$$B_{ij} = -B_{ji} \qquad \text{with probability} \quad 1 \tag{3.11}$$

and

$$B_{ii} = 0 , \qquad i, j = 1, 2, \cdots, N .$$

Further,

$$\begin{aligned} B_{ij} &= 1 && \text{if and only if} \quad -\Delta(d_i - d_j) \leqq X_i - X_j < 0 \\ &= -1 && \text{if and only if} \quad 0 \leqq X_i - X_j < -\Delta(d_i - d_j) \\ &= 0 && \text{in other cases.} \end{aligned} \tag{3.12}$$

(3.9), (3.11) and (3.12) imply

$$\begin{aligned} EH_i &= \sum_{j=1}^{N} \int [F(x + \Delta(d_i - d_j)) - F(x)]\, dF(x) \\ EB_{ij}^2 &= \int |F(x + \Delta(d_i - d_j)) - F(x)|\, dF(x) \\ E(B_{ij} B_{ik}) &= \int [F(x + \Delta[d_i - d_j)) - F(x)] \\ & \qquad \times [F(x + \Delta(d_i - d_k)) - F(x)]\, dF(x) \qquad \text{for } j \neq k . \end{aligned} \tag{3.13}$$

It follows from (3.13)

$$\begin{aligned} \operatorname{Var} H_i = {} & \sum_{j=1}^{N} \{\int |F(x + \Delta(d_i - d_j)) - F(x)|\, dF(x) \\ & \times [1 - \int |F(x + \Delta(d_i - d_j)) - F(x)|\, dF(x)]\} \\ & + \sum_{\substack{j=1 \\ j\neq k}}^{N} \sum_{k=1}^{N} \{\int [F(x + \Delta(d_i - d_j)) - F(x)] \\ & \times [F(x + \Delta(d_i - d_k)) - F(x)]\, dF(x) \\ & - \int [F(x + \Delta(d_i - d_j)) - F(x)]\, dF(x) \\ & \times \int [F(x + \Delta(d_i - d_k)) - F(x)]\, dF(x)\} . \end{aligned} \tag{3.14}$$

(3.13) together with the independence of B_{ij}, B_{kl} for i, j, k, l being different numbers further implies

$$
\begin{aligned}
\operatorname{Cov}(H_j, H_k) = \textstyle\sum_{i=1}^N \{ & \textstyle\int [F(x+\Delta(d_j-d_i)) - F(x)] \\
& \times [F(x+\Delta(d_k-d_i)) - F(x)]\, dF(x) \\
& - \textstyle\int [F(x+\Delta(d_j-d_i)) - F(x)]\, dF(x) \\
& \times \textstyle\int [F(x+\Delta(d_k-d_i)) - F(x)]\, dF(x)\} \\
& - \textstyle\int |F(x+\Delta(d_j-d_k)) - F(x)|\, dF(x) \\
& \times [1 - \textstyle\int |F(x+\Delta(d_j-d_k)) - F(x)|\, dF(x)] \\
& - \textstyle\sum_{i=1, i\neq k}^N \{\int [F(x+\Delta(d_j-d_i)) - F(x)] \\
& \times [F(x+\Delta(d_j-d_k)) - F(x)]\, dF(x) \\
& - \textstyle\int [F(x+\Delta(d_j-d_i)) - F(x)]\, dF(x) \\
& \times \textstyle\int [F(x+\Delta(d_j-d_k)) - F(x)]\, dF(x)\} \\
& - \textstyle\sum_{i=1, i\neq j}^N \{\int [F(x+\Delta(d_k-d_i)) - F(x)] \\
& \times [F(x+\Delta(d_k-d_j)) - F(x)]\, dF(x) \\
& - \textstyle\int [F(x+\Delta(d_k-d_i)) - F(x)]\, dF(x) \\
& \times \textstyle\int [F(x+\Delta(d_k-d_j)) - F(x)]\, dF(x)\}\,.
\end{aligned} \tag{3.15}
$$

From (3.10), (3.14) and (3.15), we get the variance

$$
\begin{aligned}
& \operatorname{Var}(S_{\Delta N} - S_{0N}) \\
& \quad = N^{-2} \textstyle\sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j)\{\int |F(x+\Delta(d_i-d_j)) - F(x)|\, dF(x) \\
& \qquad \times [1 - \textstyle\int |F(x+\Delta(d_i-d_j)) - F(x)|\, dF(x)] \\
& \qquad + N^{-2} \textstyle\sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N (c_i - c_j)(c_i - c_k) \\
& \qquad \times \{\textstyle\int [F(x+\Delta(d_i-d_j)) - F(x)] \\
& \qquad \times [F(x+\Delta(d_i-d_k)) - F(x)]\, dF(x) \\
& \qquad - \textstyle\int [F(x+\Delta(d_i-d_j)) - F(x)]\, dF(x) \\
& \qquad \times \textstyle\int [F(x+\Delta(d_i-d_k)) - F(x)]\, dF(x)\}\,.
\end{aligned} \tag{3.16}
$$

Now we shall deal with the covariance term in (3.7). Analogously, we define

$$H_i^{(\nu)} = \textstyle\sum_{j=1}^N B_{ij}^{(\nu)} = \sum_{j=1}^N [u(X_i - X_j + \Delta_\nu(d_i - d_j)) - u(X_i - X_j)]\,; \qquad \nu = 1, \cdots, K\,.$$

It holds that for $\nu \neq \mu$

$$
\begin{aligned}
& E(B_{ij}^{(\nu)} \cdot B_{ij}^{(\mu)}) = \textstyle\int |F(x + \min(\Delta_\nu, \Delta_\mu)(d_i - d_j)) - F(x)|\, dF(x) \qquad \text{and} \\
& E(B_{ij}^{(\nu)} \cdot B_{ik}^{(\mu)})_{j\neq k} = \textstyle\int [F(x+\Delta_\nu(d_i-d_j)) - F(x)] \\
& \qquad\qquad \times [F(x+\Delta_\mu(d_i-d_k)) - F(x)]\, dF(x)\,.
\end{aligned} \tag{3.17}
$$

Regarding the equality

$$
\begin{aligned}
& E[(S_{\Delta_\nu N} - S_{0N})(S_{\Delta_\mu N} - S_{0N})] \\
& \quad = N^{-2} \textstyle\sum_{i=1}^N c_i^2 E(H_i^{(\nu)} H_i^{(\mu)}) + N^{-2} \sum_{j=1, j\neq k}^N \sum_{k=1}^N c_j c_k E(H_j^{(\nu)} H_k^{(\mu)})
\end{aligned} \tag{3.18}
$$

we get that the covariances are

$$
\begin{aligned}
\operatorname{Cov}&(S_{\Delta_\nu N} - S_{0N}, S_{\Delta_\mu N} - S_{0N}) \\
&= N^{-2} \textstyle\sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j) \\
&\quad \times \{\int |F(x + \min(\Delta_\nu, \Delta_\mu)(d_i - d_j)) - F(x)|\, dF(x) \\
&\quad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
&\quad \times \int [F(x + \Delta_\mu(d_i - d_j)) - F(x)]\, dF(x)\} \\
&\quad + N^{-2} \textstyle\sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N (c_i - c_j)(c_i - c_k) \\
&\quad \times \{\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)] \\
&\quad \times [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x) \\
&\quad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
&\quad \times \int [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x)\}\,.
\end{aligned} \tag{3.19}
$$

The finite result follows from (3.7), (3.16) and (3.19):

$$
\begin{aligned}
\operatorname{Var} Z_N &= N^{-2} \textstyle\sum_{\nu=1}^K \lambda_\nu^2 \sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j) \\
&\quad \times \{\int |F(x + \Delta_\nu(d_i - d_j)) - F(x)|\, dF(x) \\
&\quad \times [1 - \int |F(x + \Delta_\nu(d_i - d_j)) - F(x)|\, dF(x)]\} \\
&\quad + N^{-2} \textstyle\sum_{\nu=1}^K \sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N \lambda_\nu^2 (c_i - c_j)(c_i - c_k) \\
&\quad \times \{\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)] \\
&\quad \times [F(x + \Delta_\nu(d_i - d_k)) - F(x)]\, dF(x) \\
&\quad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
&\quad \times \int [F(x + \Delta_\nu(d_i - d_k)) - F(x)]\, dF(x)\} \\
&\quad + N^{-2} \textstyle\sum_{\nu=1}^K \sum_{\mu=1, \nu\neq\mu}^K \lambda_\nu \lambda_\mu \sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j) \\
&\quad \times \{\int |F(x + \min \Delta_\nu, \Delta_\mu)(d_i - d_j)) - F(x)|\, dF(x) \\
&\quad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
&\quad \times \int [F(x + \Delta_\mu(d_i - d_j)) - F(x)]\, dF(x)\} \\
&\quad + N^{-2} \textstyle\sum_{\nu=1}^K \sum_{\mu=1, \nu\neq\mu}^K \lambda_\nu \lambda_\mu \sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N (c_i - c_j)(c_i - c_k) \\
&\quad \times \{\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)] \\
&\quad \times [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x) \\
&\quad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
&\quad \times \int [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x)\}\,.
\end{aligned} \tag{3.20}
$$

Now we shall find $\operatorname{Var} \hat{Z}_N$. We may write

$$\hat{Z}_N = \textstyle\sum_{i=1}^N Y_{iN} = \sum_{i=1}^N \sum_{\nu=1}^K \lambda_\nu Y_{iN}^{(\nu)} \tag{3.21}$$

where Y_{iN}, $i = 1, 2, \cdots, N$ are independent random variables and

$$
\begin{aligned}
Y_{iN}^{(\nu)} = N^{-1} \textstyle\sum_{j=1}^N \{&(c_i - c_j)[F(X_i + \Delta_\nu(d_i - d_j)) - F(X_i)] \\
&- c_i \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x)\}\,.
\end{aligned} \tag{3.22}
$$

Regarding (3.21) and (3.22), we get by direct computation

$$\begin{aligned}
\operatorname{Var} \hat{Z}_N = {} & N^{-2} \textstyle\sum_{\nu=1}^{K} \lambda_\nu^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} (c_i - c_j)(c_i - c_k) \\
& \times \{\textstyle\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)] \\
& \times [F(x + \Delta_\nu(d_i - d_k)) - F(x)]\, dF(x) \\
& - \textstyle\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
& \times \textstyle\int [F(x + \Delta_\nu(d_i - d_k)) - F(x)]\, dF(x)\} \\
& + N^{-2} \textstyle\sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{K} \sum_{\mu=1}^{K} \lambda_\nu \lambda_\mu \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} (c_i - c_j)(c_i - c_k) \\
& \times \{\textstyle\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)] \\
& \times [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x) \\
& - \textstyle\int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x) \\
& \times \textstyle\int [F(x + \Delta_\mu(d_i - d_k)) - F(x)]\, dF(x)\} .
\end{aligned} \tag{3.23}$$

We are now able to prove (3.6). By (2.4), it holds

$$\Delta^{-1}(d_i - d_j)^{-1}[F(x + \Delta(d_i - d_j)) - F(x)] \to f(x) \tag{3.24}$$

for N tending to infinity for any $\Delta \neq 0$ uniformly in $i, j = 1, 2, \cdots, N$ for which $d_i - d_j \neq 0$. On the other hand, assumption 1° implies that $f(x)$ is bounded for $x \in (-\infty, \infty)$, so that, in view of the mean value theorem, there is a positive constant C^* such that

$$|\Delta^{-1}(d_i - d_j)^{-1}[F(x + \Delta(d_i - d_j)) - F(x)]| \leqq C^* \tag{3.25}$$

for $N = 1, 2, \cdots, x \in (-\infty, \infty)$ and uniformly for all $i, j = 1, \cdots, N$ for which $\Delta(d_i - d_j) \neq 0$. (3.23), (3.24) and (3.25) together with Lebesgue's theorem and with the boundedness of

$$A_N^{-2}N^{-2} \textstyle\sum_{i=1}^{N} (\sum_{j=1}^{N} |(c_i - c_j)(d_i - d_j)|)^2$$

imply

$$\begin{aligned}
& \lim_{N\to\infty} [A_N^{-2} \operatorname{Var} \hat{Z}_N] \\
& = C(\textstyle\sum_{\nu=1}^{K} \lambda_\nu \Delta_\nu)^2 \cdot \lim_{N\to\infty} \{A_N^{-2}N^{-2} \sum_{i=1}^{N} (\sum_{j=1}^{N} (c_i - c_j)(d_i - d_j))^2\} \\
& = C(\textstyle\sum_{\nu=1}^{K} \lambda_\nu \Delta_\nu)^2 .
\end{aligned} \tag{3.26}$$

By (3.5) and (3.26),

$$\liminf [\operatorname{Var} Z_N \cdot A_N^{-2}C^{-1}(\textstyle\sum_{\nu=1}^{K} \lambda_\nu \Delta_\nu)^{-2}] \geqq 1 . \tag{3.27}$$

If we prove

$$\limsup [\operatorname{Var} Z_N \cdot A_N^{-2}C^{-1}(\textstyle\sum_{\nu=1}^{K} \lambda_\nu \Delta_\nu)^{-2}] \leqq 1 \tag{3.28}$$

then (3.6) will follow.

For the first member of the right-hand side of (3.20), we have in view of (3.24) and (3.25)

$$\begin{aligned}
0 & \leqq \limsup [A_N^{-2}N^{-2} \textstyle\sum_{\nu=1}^{K} \lambda_\nu^2 \sum_{i=1}^{N} \sum_{j=1}^{N} |c_i^2 - c_i c_j| \\
& \qquad \times \{\textstyle\int |F(x + \Delta_\nu(d_i - d_j)) - F(x)|\, dF(x) \\
& \qquad \times (1 - \textstyle\int |F(x + \Delta_\nu(d_i - d_j)) - F(x)|\, dF(x)\}] \\
& \leqq \textstyle\sum_{\nu=1}^{K} \lambda_\nu^2 \Delta_\nu \int f^2(x)\, dx \\
& \qquad \times \lim_{N\to\infty} [N^{-2}A_N^{-2} \textstyle\sum_{i=1}^{N} \sum_{j=1}^{N} |(c_i^2 - c_i c_j)(d_i - d_j)|] = 0
\end{aligned} \tag{3.29}$$

(see assumptions 2°, 3° and 4°).

The upper limit of the product of A_N^{-2} and of the second member of the righ-hand side of (3.20) is

$$(C \textstyle\sum_{\nu=1}^K \lambda_\nu^2 \Delta_\nu^2) \limsup [N^{-2} A_N^{-2} \sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N (c_i - c_j) \times (c_i - c_k)(d_i - d_j)(d_i - d_k)] = C \sum_{\nu=1}^K \lambda_\nu^2 \Delta_\nu^2 . \tag{3.30}$$

The upper limit of the A_N^{-2}-product of the third member of the right-hand side of (3.20) is

$$\begin{aligned} &\textstyle\sum_{\nu=1, \nu\neq\mu}^K \sum_{\mu=1}^K \lambda_\nu \lambda_\mu \{\min(\Delta_\nu, \Delta_\mu) \int f^2(x)\, dx \\ &\quad \times \lim_{N\to\infty} [N^{-2} A_N^{-2} \sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j)|(d_i - d_j)|] \\ &\quad - \Delta_\nu \Delta_\mu (\int f^2(x)\, dx)^2 \\ &\qquad \times \lim_{N\to\infty} [N^{-2} A_N^{-2} \sum_{i=1}^N \sum_{j=1}^N (c_i^2 - c_i c_j)(d_i - d_j)^2]\} = 0 . \end{aligned} \tag{3.31}$$

Finally, the limit of the A_N^{-2}-product of the last member of the right-hand side of (3.20) is

$$C \textstyle\sum_{\nu=1, \nu\neq\mu}^K \sum_{\mu=1}^K \lambda_\nu \lambda_\mu \Delta_\nu \Delta_\mu \lim_{N\to\infty} \{A_N^{-2} N^{-2} \sum_{i=1}^N \sum_{j=1, j\neq k}^N \sum_{k=1}^N (c_i - c_j) \times (c_i - c_k)(d_i - d_j)(d_i - d_k)\} = C \sum_{\nu=1, \nu\neq\mu}^K \sum_{\mu=1}^K \lambda_\nu \lambda_\mu \Delta_\nu \Delta_\mu . \tag{3.32}$$

(3.29)—(3.32) imply

$$\limsup [A_N^{-2} \cdot \operatorname{Var} Z_N] \leqq C(\textstyle\sum_{\nu=1}^K \lambda_\nu \Delta_\nu)^2 , \tag{3.33}$$

and this together with (3.27) implies (3.6).

It remains to prove that $\hat{Z}_N$ has asymptotically normal distribution ($E\hat{Z}_N$, $\operatorname{Var} \hat{Z}_N$). If $\sum_{\nu=1}^K \lambda_\nu \Delta_\nu = 0$, then $A_N^{-2} \operatorname{Var} \hat{Z}_N \to 0$ and the criterion of degenerate convergence is satisfied for the sequence $\hat{Z}_N - E\hat{Z}_N$ (see [11]). Let $\sum_{\nu=1}^K \lambda_\nu \Delta_\nu \neq 0$ and denote

$$\sigma_N^2 = \operatorname{Var} \hat{Z}_N ; \tag{3.34}$$

then the Lindeberg condition (see [11]) for $\hat{Z}_N = \sum_{i=1}^N Y_{iN}$ takes on the form

$$\sigma_N^{-2} \textstyle\sum_{i=1}^N \int_{|x|\geqq \varepsilon\sigma_N} x^2\, dP(Y_{iN} - EY_{iN} \leqq x) \to 0 \tag{3.35}$$

where Y_{iN} are given by (3.21) and (3.22). By (3.26), there exists N_0 such that for all $N \geqq N_0$

$$\tfrac{1}{2} A_N^2 C (\textstyle\sum_{\nu=1}^K \lambda_\nu \Delta_\nu)^2 \leqq \sigma_N^2 \leqq \tfrac{3}{2} A_N^2 C (\sum_{\nu=1}^K \lambda_\nu \Delta_\nu)^2 , \tag{3.36}$$

so that (3.35) will hold if

$$2A_N^{-2} C^{-1} (\textstyle\sum_{\nu=1}^K \lambda_\nu \Delta_\nu)^{-2} \sum_{i=1}^N \int_{|x|\geqq \varepsilon B_N} x^2\, dP(Y_{iN} - EY_{iN} \leqq x) \to 0 \tag{3.37}$$

where $B_N = A_N (C/2)^{\frac{1}{2}} |\sum_{\nu=1}^K \lambda_\nu \Delta_\nu|$.

Let us investigate integration domains in (3.37). By (3.21), (3.22) and (3.25), there exists a constant $C_0 > 0$ such that

$$\begin{aligned} (3.38)\quad & A_N^{-1}|Y_{iN} - EY_{iN}| \\ & = N^{-1}A_N^{-1}|\sum_{\nu=1}^K \lambda_\nu \sum_{j=1}^N (c_i - c_j)\{[F(X_i + \Delta_\nu(d_i - d_j)) - F(X_i)] \\ & \qquad - \int [F(x + \Delta_\nu(d_i - d_j)) - F(x)]\, dF(x)\}| \\ & \leqq N^{-1}A_N^{-1}C_0 \sum_{\nu=1}^K |\lambda_\nu \Delta_\nu| \sum_{j=1}^N |(c_i - c_j)(d_i - d_j)| \\ & \leqq 2C_0 A_N^{-1} \sum_{\nu=1}^K |\lambda_\nu \Delta_\nu| N^{-\frac{1}{2}}[Nc_i^2 d_i^2 + c_i^2 \sum_{j=1}^N d_j^2 \\ & \qquad + d_i^2 \sum_{j=1}^N c_j^2 + \sum_{j=1}^N c_j^2 d_j^2]^{\frac{1}{2}} . \end{aligned}$$

The right-hand side of (3.38) tends to zero for N tending to infinity uniformly for $i = 1, 2, \cdots, N$ in view of (2.2), (2.4), (2.5) and (2.6), so that there exists N_1 such that for all $N > N_1$

$$(3.39)\qquad \max_{1 \leqq i \leqq N} |Y_{iN} - EY_{iN}| A_N^{-1} < (\varepsilon/2)(C/2)^{\frac{1}{2}} |\sum_{\nu=1}^K \lambda_\nu \Delta_\nu| ,$$

and thus

$$(3.40)\qquad P\{|Y_{iN} - EY_{iN}| \geqq \varepsilon B_N\} = 0$$

for $N > N_1$ and $i = 1, 2, \cdots, N$. It means that Lindeberg's condition (3.35) is trivially satisfied, and this together with (3.6) and (3.26) implies that

$$(3.41)\qquad \mathscr{L}\{A_N^{-1}(Z_N - EZ_N)\} \to N(0, C(\sum_{\nu=1}^K \lambda_\nu \Delta_\nu)^2) ,$$

and thus the asymptotic distribution of $(\mathscr{S}_{\Delta_1 N}, \mathscr{S}_{\Delta_2 N}, \cdots, \mathscr{S}_{\Delta_K N})$ is K-dimensional normal with null means and covariance matrix $(\sigma_{\nu\mu})_{\nu,\mu=1}^K$ with $\sigma_{\nu\mu} = C\Delta_\nu \Delta_\nu$; $\nu, \mu = 1, 2, \cdots, K$.

(ii) Proof of tightness of the sequence of distributions of $\mathscr{S}_{\Delta N}$.

By Theorem 15.4 of [3], this sequence is tight if for each positive ε and $\eta > 0$ there exist a δ, $0 < \delta < 1$ and an integer N_0 such that

$$(3.42)\qquad P\{w''(\mathscr{S}_N, \delta) \geqq \varepsilon\} \leqq \eta , \qquad N \geqq N_0, \quad \text{where}$$

$$(3.43)\qquad w''(\mathscr{S}_N, \delta) = \sup \min \{|\mathscr{S}_{\Delta N} - \mathscr{S}_{\Delta_1 N}|, |\mathscr{S}_{\Delta_2 N} - \mathscr{S}_{\Delta N}|\}$$

where the supremum extends over Δ_1, Δ and Δ_2 satisfying

$$(3.44)\qquad \Delta_1 \leqq \Delta \leqq \Delta_2 , \qquad \Delta_2 - \Delta_1 \leqq \delta .$$

For proving (3.42), we shall need the following slight generalization of Theorem 12.1 of [3].

LEMMA 3.1. *For any pair of positive integers N, m, let the sequence of* rv's $\xi_{N1}^{(m)}, \xi_{N2}^{(m)}, \cdots, \xi_{Nm}^{(m)}$ *be given. Let* $S_{Nk}^{(m)} = \xi_{N1}^{(m)} + \cdots + \xi_{Nk}^{(m)}$ $(S_{N0}^{(m)} = 0)$ *for* $k = 1, 2, \cdots, m$, *and put*

$$(3.45)\qquad M'_{Mm} = \max_{0 \leqq k \leqq m} \min \{|S_{Nk}^{(m)}|, |S_{Nm}^{(m)} - S_{Nk}^{(m)}|\} .$$

For each m, let there exist nonnegative numbers $u_1^{(m)}, \cdots, u_m^{(m)}$ such that

$$(3.46)\qquad \limsup P\{|S_{Nj}^{(m)} - S_{Ni}^{(m)}| \geqq \lambda, |S_{Nk}^{(m)} - S_{Nj}^{(m)}| \geqq \lambda\} \leqq \lambda^{-2}(\sum_{i \leqq l \leqq k} u_l^{(m)})^2 , \quad 0 \leqq i \leqq j \leqq k \leqq m$$

holds for all $\lambda > 0$ *uniformly for* $m = 1, 2, \cdots$. *Then, for all positive* λ *and uniformly in* m, *it holds that*

$$\limsup P\{M'_{Nm} \geqq \lambda\} \leqq K\lambda^{-2}(u_1^{(m)} + \cdots + u_m^{(m)})^2 \tag{3.47}$$

where K *is a constant independent on* m.

Let us apply this lemma to the system of rv's

$$\xi_{Ni}^{(m)} = \mathscr{S}_N(\Delta + (i/m)\delta) - \mathscr{S}_N(\Delta + ((i-1)/m)\delta)\,, \qquad i = 1, \cdots, m \tag{3.48}$$

where $\mathscr{S}_N(\Delta)$ is written instead of $\mathscr{S}_{\Delta N}$; Δ and δ are fixed numbers from $[0, 1]$. Let

$$M''_{Nm} = \max\min\{|\mathscr{S}_N(\Delta + (j/m)\delta) - \mathscr{S}_N(\Delta + (i/m)\delta)|, \\ |\mathscr{S}_N(\Delta + (k/m)\delta) - \mathscr{S}_N(\Delta + (j/m)\delta)|\} \tag{3.49}$$

where the maximum extends over $0 \leqq i \leqq j \leqq k \leqq m$. In view of (3.28) and (3.33)

$$\lim_{N\to\infty} \operatorname{Var}(\mathscr{S}_{\Delta_2 N} - \mathscr{S}_{\Delta_1 N}) = C(\Delta_2 - \Delta_1)^2 \tag{3.50}$$

and it follows from the proof of (3.33) that the convergence (3.50) is uniform for $0 \leqq \Delta_1, \Delta_2 \leqq 1$. Regarding Schwarz's inequality, we see from (3.50) that

$$\begin{aligned} \limsup P\{&|\mathscr{S}_N(\Delta_0 + (j/m)\delta) - \mathscr{S}_N(\Delta_0 + (i/m)\delta)| \geqq \lambda, \\ &|\mathscr{S}_N(\Delta_0 + (k/m)\delta) - \mathscr{S}_N(\Delta_0 + (j/m)\delta)| \geqq \lambda\} \\ &\leqq C[(k-i)/m]^2\lambda^{-2}\delta^2 \end{aligned} \tag{3.51}$$

uniformly for $\Delta_0 \in [0, 1]$, $0 \leqq i \leqq j \leqq k \leqq m$ and $m = 1, 2, \cdots$. By Lemma 3.1,

$$\limsup P(M''_{Nm} \geqq \varepsilon) \leqq CK\varepsilon^{-2}\delta^2 \tag{3.52}$$

which implies (3.42). Theorem 2.1 is proved.

4. Modification of means. Under additional assumptions, we can give a simple form to the asymptotic center of the process $S_{\Delta N} - S_{0N}$. Namely, we suppose

1°° $(X_{N1}, X_{N2}, \cdots, X_{NN})$ is an independent random sample from a distribution with cdf F and density f such that the integrals

$$\int ([f'(x-t)]^2/f(x))\,dx \tag{4.1}$$

are bounded for $|t| \leqq \delta$ for some $\delta > 0$.

2°° $(c_{N1}, c_{N2}, \cdots, c_{NN})$ is a vector of regression constants such that

$$\textstyle\sum_{i=1}^N c_{Ni} = 0\,, \qquad N = 2, 3, \cdots \tag{4.2}$$

$$\lim_{N\to\infty}[\max_{1\leqq i\leqq N} c_{Ni}^2(\textstyle\sum_{i=1}^N c_{Ni}^2)^{-1}] = 0\,. \tag{4.3}$$

3°° $(d_{N1}, d_{N2}, \cdots, d_{NN})$, $N = 1, 2, \cdots$ are vectors such that

$$\textstyle\sum_{i=1}^N d_{Ni} = 0\,, \tag{4.4}$$

$$N\max_{1\leqq i\leqq N} d_{Ni}^2(\textstyle\sum_{i=1}^N d_{Ni}^2)^{-1} = O(1)\,, \tag{4.5}$$

$$\textstyle\sum_{i=1}^N d_{Ni}^2 = O(N^{-1+\eta}) \qquad \text{where} \quad 0 < \eta < 1\,. \tag{4.6}$$

4°° The vectors $(c_{N1}, c_{N2}, \cdots, c_{NN})$ and $(d_{N1}, d_{N2}, \cdots, d_{NN})$ satisfy the conditions

$$\lim_{N\to\infty} \{(\sum_{i=1}^N c_{Ni} d_{Ni})[(\sum_{i=1}^N c_{Ni}^2)(\sum_{i=1}^N d_{Ni}^2)]^{-\frac{1}{2}}\} = b^2 > 0 \tag{4.7}$$

and

$$\lim_{N\to\infty} [\max_{1\leq i\leq N} (c_{Ni}^2 d_{Ni}^2)(\sum_{i=1}^N c_{Ni}^2 d_{Ni}^2)^{-1}] = 0 . \tag{4.8}$$

We shall prove the theorem.

THEOREM 4.1. *Under assumptions* 1°°—4°°, 5° *and* 6°, *the random process*

$$\mathscr{S}_{\Delta N}^* = A_N^{-1}(S_{\Delta N} - S_{0N} - \Delta b_N) , \qquad 0 \leqq \Delta \leqq 1 \tag{4.9}$$

with A_N *given by* (2.9) *and*

$$b_N = (\sum_{i=1}^N c_{Ni} d_{Ni}) \int f^2(x) \, dx \tag{4.10}$$

converges weakly to the Gaussian process $\{G(\Delta): 0 \leqq \Delta \leqq 1\}$ *satisfying* (2.11) *and* (2.12).

PROOF. Let us denote

$$c_{Ni}^* = c_{Ni}(\sum_{j=1}^N c_{Nj}^2)^{-\frac{1}{2}} , \qquad d_{Ni}^* = d_{Ni}(\sum_{j=1}^N d_{Nj}^2)^{-\frac{1}{2}} . \tag{4.11}$$

Then c_{Ni}'s and d_{Ni}'s satisfy assumptions 1°—4°, so that all steps of the proof of the Theorem 2.1 are valid for them. On the other hand, all convergences contained in the proof of Theorem 2.1 except (3.29) and (3.31) are also valid for c_{Ni}'s and d_{Ni}'s satisfying 1°°—4°°, as may be shown by reducing the corresponding fractions by $(\sum_{i=1}^N c_{Ni}^2) \cdot (\sum_{i=1}^N d_{Ni}^2)$. As for (3.29) and (3.31), the convergence

$$\lim_{N\to\infty} [N^{-2} A_N^{-2} \sum_{i=1}^N \sum_{j=1}^N |c_i^2 - c_i c_j||d_i - d_j|] = 0 \tag{4.12}$$

has to be proved for c_{Ni}'s and d_{Ni}'s satisfying 1°°—4°°. But assumptions 1°°—4°° together with (4.11) imply

$$N^{-2} A_N^{*-2} \sum_{i=1}^N \sum_{j=1}^N |c_i^{*2} - c_i^* c_j^*)(d_i^* - d_j^*)| = O(N^{-\frac{1}{2}}) \tag{4.13}$$

where $A_N^{*2} = \sum_{i=1}^N (c_i^* d_i^*)^2 + 3N^{-1}(\sum_{i=1}^N c_i^* d_i^*)^2$, so that

$$\begin{aligned} &N^{-2} A_N^{-2} \sum_{i=1}^N \sum_{j=1}^N |(c_i^2 - c_i c_j)(d_i - d_j)| \\ &\quad = N^{-2} A_N^{*-2} \sum_{i=1}^N \sum_{j=1}^N |(c_i^{*2} - c_i^* c_j^*)(d_i^* - d_j^*)|(\sum_{i=1}^N d_i^2)^{-\frac{1}{2}} = O(N^{-\eta/2}) \end{aligned}$$

by (4.6), so that (4.12) is right. Theorem 2.1 thus holds also for c_{Ni}'s and d_{Ni}'s satisfying 1°°—4°°, and all that it remains to prove is

$$\begin{aligned} \lim_{N\to\infty} A_N^{-2}[N^{-1} \sum_{i=1}^N c_i \sum_{j=1}^N \{\int [F(x + \Delta(d_i - d_j)) - F(x)] \, dF(x) \\ - \Delta(d_i - d_j) \int f^2(x) \, dx\}]^2 = 0 . \end{aligned} \tag{4.14}$$

For $\Delta(d_i - d_j) \geqq 0$, we have

$$\begin{aligned} &|\int [F(x + \Delta(d_i - d_j)) - F(x) - \Delta(d_i - d_j) f(x)] \, dF(x)| \\ &\quad = |\int_0^{\Delta(d_i - d_j)} (\int [f(x + u) - f(x)] f(x) \, dx) \, du| \\ &\quad \leqq \int_0^{\Delta(d_i - d_j)} \{\int_0^u (\int |f'(x + z) f(x)| \, dx) \, dz\} \, du \\ &\quad \leqq \int_0^{\Delta(d_i - d_j)} \{\int_0^u [\int ([f'(x + z)]^2 / f(x)) \, dx \cdot \int f^3(y) \, dy]^{\frac{1}{2}} \, dz\} \, du . \end{aligned} \tag{4.15}$$

In view of assumptions $1^{\circ\circ}$ and $3^{\circ\circ}$ it follows that there exists an integer N_0 and a constant $C^* > 0$ such that for $N \geqq N_0$

$$|\int [F(x + \Delta(d_i - d_j)) - F(x) - \Delta(d_i - d_j)f(x)]\, dF(x)| \leqq C^*\Delta^2(d_i - d_j)^2 . \tag{4.16}$$

For $\Delta(d_i - d_j) < 0$, we get an analogous conclusion.

(4.16) implies that for $N \geqq N_0$

$$\begin{aligned} A_N^{-2}(ES_{\Delta N} - \Delta b_N)^2 & \\ \leqq C^{*2}\Delta^4 A_N^{-2}N^{-2}[\textstyle\sum_{i=1}^N \sum_{j=1}^N |c_i|(d_i - d_j)^2]^2 & \\ = C^{*2}\Delta^4 A_N^{*-2}N^{-2}[\textstyle\sum_{i=1}^N \sum_{j=1}^N |c_i^*|(d_i^* - d_j^*)^2]^2 \cdot (\sum_{i=1}^N d_i^2) & \\ = O(1) \cdot O(N^{-1+\eta}) . & \qquad \square \end{aligned} \tag{4.17}$$

5. Application. In this section, we shall use the following notation: for any positive integer N, let $X_{N1}, X_{N2}, \cdots, X_{NN}$ be independent random variables, X_{Ni} having cdf

$$F(x - \Delta^0 d_{Ni}) , \qquad i = 1, 2, \cdots, N \tag{5.1}$$

such that the integrals

$$\int \{[f'(x - t)]^2/f(x)\}\, dx \tag{5.2}$$

are bounded for $|t| \leqq \delta$ for some $\delta > 0$.

Suppose that the constants d_{Ni}, $i = 1, 2, \cdots, N$ are such that

$$\textstyle\sum_{i=1}^N d_{Ni} = 0 , \qquad \sum_{i=1}^N d_{Ni}^2 = 1 , \qquad N = 2, 3, \cdots , \tag{5.3}$$

$$\max_{1\leqq i\leqq N} |d_{Ni}| = O(N^{-\frac{1}{2}}) \tag{5.4}$$

and

$$\lim_{N\to\infty} [\max_{1\leqq i\leqq N} d_{Ni}^4(\textstyle\sum_{i=1}^N d_{Ni}^4)^{-1}] = 0 . \tag{5.5}$$

Let us consider the statistics

$$S_N(X_1, \cdots, X_N) = S_N(X_i) = N^{-1} \textstyle\sum_{i=1}^N d_{Ni} R_{Ni} \tag{5.6}$$

where R_{Ni} is the rank of X_{Ni}, $i = 1, 2, \cdots, N$, i.e.

$$R_{Ni} = \textstyle\sum_{j=1}^N u(X_i - X_j) . \tag{5.7}$$

For any positive integer N, consider the confidence set

$$D_N = \{\Delta : |S_N(X_i - \Delta d_i)| \leqq \varphi_\alpha\} \qquad \text{where} \tag{5.8}$$

$$\varphi_\alpha = 12^{-\frac{1}{2}}\Phi^{-1}((\alpha + 1)/2) , \qquad 0 < \alpha < 1 . \tag{5.9}$$

Φ^{-1} is the inverse cdf of the standard normal distribution. We have the following

LEMMA 5.1. *Under* (5.1)—(5.9),

$$\lim_{N\to\infty} P_{\Delta^0}(D_N \ni \Delta^0) = \alpha \tag{5.10}$$

holds for any α, $0 < \alpha < 1$.

PROOF.

$$P_{\Delta^0}(D_N \ni \Delta^0) = P_{\Delta^0}(|S_N(X_i - \Delta^0 d_i)| \leqq \varphi_\alpha)$$
$$= P_0(|S_N(X_i)| \leqq \varphi_\alpha) \to 2\Phi(12^{\frac{1}{2}}\varphi_\alpha) - 1 = \alpha$$

where the convergence follows from Theorem V.1.6.a and Lemma V.1.6.a of [6].

Noting that the statistic $S_N(X_i - \Delta d_i)$ is a non-increasing function of Δ for fixed $X_1, \cdots, X_N$ with probability 1 (see Theorem 2.1 of [7]), we can write

$$D_N = (T_N^-, T_N^+)\,, \tag{5.11}$$

where

$$T_N^- = \sup\{\Delta : S_N(X_i - \Delta d_i) > \varphi_\alpha\} \tag{5.12}$$

and

$$T_N^+ = \inf\{\Delta : S_N(X_i - \Delta d_i) < -\varphi_\alpha\}\,. \tag{5.13}$$

The first result concerning the asymptotic behavior of the length $T_N^+ - T_N^-$ of the confidence interval is the following

LEMMA 5.2. *Under* (5.1)—(5.9),

$$\{(T_N^+ - T_N^-)[2\Phi^{-1}(\tfrac{1}{2}(\alpha + 1))\sigma_\Delta]^{-1} - 1\} \to_{P_{\Delta^0}} 0 \tag{5.14}$$

for $N \to \infty$, where σ_Δ^2 is the asymptotic variance of the estimate of Δ^0 of Hodges–Lehmann type based on $S_N(X_i)$, i.e.

$$\sigma_\Delta^2 = [12(\textstyle\int f^2(x)\,dx)^2]^{-1}\,. \tag{5.15}$$

PROOF. (For σ_Δ^2, see e.g. [1] or [8].) We have for any real t

$$\begin{aligned} \lim_{N\to\infty} P_{\Delta^0}(T_N^- > t) &= \lim_{N\to\infty} P_{\Delta^0}\{(S_N(X_i - td_i) > \varphi_\alpha\} \\ &= \lim_{N\to\infty} P_0\{S_N(X_i + (\Delta^0 - t)d_i) > \varphi_\alpha\} \\ &= 1 - \Phi(12^{\frac{1}{2}}[\varphi_\alpha - (\Delta^0 - t)\textstyle\int f^2(x)\,dx]) \end{aligned} \tag{5.16}$$

(where we have used Theorem and Lemma V.1.6.a of [6]); thus T_N^- is asymptotically normal

$$(\Delta^0 - \Phi^{-1}(\tfrac{1}{2}(\alpha + 1)) \cdot 12^{-\frac{1}{2}}(\textstyle\int f^2(x)\,dx)^{-1}, (12[\textstyle\int f^2(x)\,dx]^2)^{-1})\,. \tag{5.17}$$

Analogously, we can prove that T_N^+ is asymptotically normal

$$(\Delta^0 + \Phi^{-1}(\tfrac{1}{2}(\alpha + 1)) \cdot 12^{-\frac{1}{2}}(\textstyle\int f^2(x)\,dx)^{-1}, (12[\textstyle\int f^2(x)\,dx]^2)^{-1})\,. \tag{5.18}$$

T_N^- and T_N^+ are thus bounded in probability, and by Theorem 3.1 of [7] and Theorem 2.1 of [5], it holds for any $\varepsilon > 0$ that

$$\begin{aligned} \lim_{N\to\infty} P_{\Delta^0}\{|S_N(X_i - T_N^- d_i) &- S_N(X_i - \Delta^0 d_i) \\ &+ (T_N^- - \Delta^0)\textstyle\int f^2(x)\,dx| \geqq \varepsilon\} = 0 \end{aligned} \tag{5.19}$$

and

$$\begin{aligned} \lim_{N\to\infty} P_{\Delta^0}\{|S_N(X_i - T_N^+ d_i) &- S_N(X_i - \Delta^0 d_i) \\ &+ (T_N^+ - \Delta^0)\textstyle\int f^2(x)\,dx| \geqq \varepsilon\} = 0\,. \end{aligned} \tag{5.20}$$

(5.12), (5.13), (5.17) and (5.18) then imply

(5.21) $$(T_N^+ - T_N^-) \int f^2(x)\,dx - 2.12^{-\frac{1}{2}}\Phi^{-1}(\tfrac{1}{2}(\alpha+1)) \to_{P_{\Delta^0}} 0 .$$ □

The main result concerning the asymptotic behavior of $T_N^+ - T_N^-$ is the following

THEOREM 5.1. *Under* (5.1)—(5.9), *the asymptotic distribution of the sequence*

(5.22) $$A_N^{-1}[(T_N^+ - T_N^-)(2\Phi^{-1}(\tfrac{1}{2}(\alpha+1))\sigma_\Delta)^{-1} - 1]$$

is normal with the parameters

(5.23) $$(0, [\int f^3(x)\,dx - (\int f^2(x)\,dx)^2]\cdot(\int f^2(x)\,dx)^{-2})$$

where $A_N^2 = \sum_{i=1}^N d_i^4 + 3N^{-1}$.

PROOF. By Theorem 4.1 of this paper, it holds that for any η, $0 < \eta < \frac{1}{2}$, and for any Δ_1, Δ_2, $\Delta_2 > 0$

(5.24) $$\lim_{N\to\infty} P_{\Delta^0}\{N^{\frac{1}{2}-\eta}A_N^{-1}C^{-\frac{1}{2}}[S_N(X_i - (\Delta_1 - \Delta_2)N^{-\frac{1}{2}+\eta}d_i) - S_N(X_i - \Delta_1 N^{-\frac{1}{2}+\eta}d_i) - \Delta_2 N^{-\frac{1}{2}+\eta}\int f^2(x)\,dx] \leqq y\} = \Phi(y/\Delta_2)$$

where

(5.25) $$C = \int f^3(x)\,dx - (\int f^2(x)\,dx)^2 .$$

Now, using a technique similar to that of Chapter III of [2], we get from (5.24) and Lemma 5.2

(5.26) $$\lim_{N\to\infty} P_{\Delta^0}\{N^{\frac{1}{2}-\eta}A_N^{-1}C^{-\frac{1}{2}}[S_N(X_i - T_N^- N^{-\frac{1}{2}+\eta}d_i) - S_N(X_i - T_N^+ N^{-\frac{1}{2}+\eta}d_i) - N^{-\frac{1}{2}+\eta}(T_N^+ - T_N^-)\int f^2(x)\,dx] \leqq y\} = \Phi(y/B)$$

where $B = 2\Phi^{-1}(\frac{1}{2}(\alpha+1))\sigma_\Delta$.

Theorem 3.1 of [7] implies that

(5.27) $$\lim_{N\to\infty} P_{\Delta^0}\{\max_{|\Delta-\Delta^0|\leqq C^*} |S_N(X_i - \Delta N^{-\frac{1}{2}+\eta}d_i) - S_N(X_i - \Delta^0 N^{-\frac{1}{2}+\eta}d_i) + (\Delta - \Delta^0)N^{-\frac{1}{2}+\eta}\int f^2(x)\,dx| \geqq \varepsilon\} = 0$$

and

(5.28) $$\lim_{N\to\infty} P_{\Delta^0}\{\max_{|\Delta-\Delta^0|\leqq C^*} N^{-\frac{1}{2}+\eta}|S_N(X_i - \Delta d_i) - S_N(X_i - \Delta^0 d_i) + (\Delta - \Delta^0)\int f^2(x)\,dx| \geqq \varepsilon\} = 0$$

hold for any $\varepsilon > 0$, $C^* > 0$. (5.27) and (5.28) together with the boundedness of T_N^- and T_N^+ in probability imply that

(5.29) $$\lim_{N\to\infty} P_{\Delta^0}\{|[S_N(X_i - T_N^- N^{-\frac{1}{2}+\eta}d_i) - S_N(X_i - T_N^+ N^{-\frac{1}{2}+\eta}d_i)] - N^{-\frac{1}{2}+\eta}[S_N(X_i - T_N^- d_i) - S_N(X_i - T_N^+ d_i)]| \geqq \varepsilon\} = 0$$

holds for any $\varepsilon > 0$, and thus in view of (5.12) and (5.13) we get

(5.30) $$N^{\frac{1}{2}-\eta}[S_N(X_i - T_N^- N^{-\frac{1}{2}+\eta}d_i) - S_N(X_i - T_N^+ N^{-\frac{1}{2}+\eta}d_i)] \to_{P_{\Delta^0}} 2\varphi_\alpha .$$

The final result then follows from (5.26) and (5.30):

$$(5.31)\qquad \lim_{N\to\infty} P_{\Delta^0}\{A_N^{-1}C^{-\frac{1}{2}}[(T_N^+ - T_N^-)\textstyle\int f^2(x)\,dx - 2\varphi_\alpha] \leqq y\} = \Phi(y/B)$$

and the theorem is proved.

REFERENCES

[1] ADICHIE, J. N. (1967). Estimates of regression coefficients based on rank tests. *Ann. Math. Statist.* **38** 894–904.

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

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

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

[5] HÁJEK, J. (1970). Miscellaneous problems of rank test theory. *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.). Cambridge Univ. Press.

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

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

[8] JUREČKOVÁ, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Statist.* **42** 1328–1338.

[9] KOUL, K. L. (1971). Asymptotic behavior of a class of confidence regions based on ranks in regression. *Ann. Math. Statist.* **42** 466–476.

[10] LEHMANN, E. L. (1963). Nonparametric confidence intervals for a shift parameter. *Ann. Math. Statist.* **34** 1507–1512.

[11] LOÈVE, M. (1962). *Probability Theory.* Van Nostrand, New York.

DEPARTMENT OF STATISTICS
CHARLES UNIVERSITY
SOKOLOVSKA 83, PRAGUE 8
CZECHOSLOVAKIA