Z. Wahrscheinlichkeitstheorie verw. Gebiete
70, 395–410 (1985)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Limit Theorems for Generalized Sequential Rank Statistics

F. Lombard[1]* and D.M. Mason[2]**

[1] University of South Africa, Department of Statistics, P.O. Box 392, Pretoria 001, Republic of South Africa
[2] University of Delaware, Department of Mathematical Sciences, 501 Ewing Hall, Newark, DE 19716, USA

Summary. A general class of statistics based on sequential ranks is introduced. Under suitable regularity conditions, an almost sure representation and invariance principle are established for this class. In particular, it is shown that these statistics can obey invariance principles that are radically different from those obeyed by the usual full rank statistics.

1. Introduction

Let (X_i, Y_i), $i \geqq 1$, denote a sequence of independent bivariate random vectors, where each X_i is assumed to have a continuous distribution function. Let R_i denote the rank of X_i among $X_1, \ldots, X_i$, i.e.

$$R_i = \sum_{j=1}^{i} I(X_j \leqq X_i) \quad \text{for } i \geqq 1, \tag{1}$$

where $I(\cdot)$ is the indicator function. A general linear sequential rank statistic is defined for $n \geqq 2$ by

$$M_n = \sum_{i=2}^{n} c_i Y_i J(R_i/(i+1)) \tag{2}$$

where $\{c_i\}$ is a sequence of constants and J is a function defined on the interval $(0, 1)$.

The asymptotic distribution of various special cases of (2) has been the subject of investigation by a number of authors. Parent (1963) and Reynolds (1975) investigated a signed sequential rank statistic, which is obtained from (2) by setting $c_i = 1$, $J(u) = u$, $Y_i = \operatorname{sign}(Z_i)$ and $X_i = |Z_i|$, $Z_1, Z_2, \ldots$, being a sequence of i.i.d. random variables. Their work is concerned with sequential tests, based

* Research supported by the CSIR and the University of South Africa
** Research supported by the University of Delaware Research Foundation, Grant #8325530015

on this version of M_n, of the hypothesis that the common distribution of the Z_i's is symmetric about 0. In particular, Reynolds obtains an invariance principle for the case when the Z_i's are assumed to be i.i.d. with a distribution not necessarily symmetric about 0. Sen (1978) derives an invariance principle for a class of statistics, which he calls *rank discounted partial sums*. These statistics are obtained from (2) by setting $c_i \equiv 1$ and $Y_i = g(X_i)$, where g is a continuous function satisfying a boundedness condition. He also assumes that the X_i's are i.i.d.

Now let $R_{n,i}$ denote the rank of X_i among $X_1, \ldots, X_n$, i.e.

$$R_{n,i} = \sum_{j=1}^{n} I(X_j \leqq X_i) \qquad \text{for } 1 \leqq i \leqq n,\ n \geqq 1.$$

Note that $R_i = R_{ii}$ for $i \geqq 1$. A general linear full rank statistic is defined by

$$T_n = \sum_{i=1}^{n} e_i Y_i J(R_{n,i}/(n+1)) \tag{3}$$

where $\{e_i\}$ is a sequence of constants. Mason (1981) considers the version of M_n obtained by setting $Y_i \equiv 1$ and assuming that $X_1, X_2, \ldots,$ are i.i.d. It is known that in this situation M_n consists of independent summands. Mason exploits a close connection between M_n and T_n to obtain, among other results, an invariance principle and a law of the iterated logarithm for T_n from the corresponding (easily established) results for M_n. Lombard (1981, 1983) uses the Hájek projection technique to establish another close relationship between M_n and T_n which implies, in particular, that hypothesis tests based on these statistics will have the same distribution, hence Pitman efficiency, under sequences of local alternatives. On the other hand, in Mason (1984) it is shown that hypothesis tests based on T_n and M_n will, in general, have differing Bahadur efficiencies under fixed alternatives. His conclusion is that tests based on T_n are more efficient in this sense than those based on M_n.

Theorems 1 and 2 in Sect. 2 below imply that the asymptotic distributions of M_n and T_n, though the same under sequences of local alternatives, may be quite different under fixed alternatives. This is in line with the Bahadur efficiency study of Mason (1984). Theorem 1 below gives an almost sure representation of M_n as a sum of independent *non-identically* distributed random variables plus a remainder term which converges almost surely to zero at a certain rate. Sen (1981, p. 135) has given an almost sure representation for the signed-rank version of T_n (obtained by setting $e_i \equiv 1$, $Y_i = \operatorname{sign}(Z_i)$, $X_i = |Z_i|$ in (3), whith $Z_1, Z_2, \ldots$ i.i.d.) as a sum of i.i.d. random variables plus a remainder term. Utilizing this representation, he has shown that an appropriately constructed continuous time stochastic process based on $\{T_n\}$ converges weakly to a Brownian motion. It is a consequence of Theorem 2 below that the corresponding process based on the signed rank version of M_n converges weakly to a nonstationary mean zero continuous Gaussian process $(Z(t), 0 \leqq t \leqq 1)$ with covariance function given for $0 \leqq s,\ t \leqq 1$ by

$$\rho(s,t) = (s \wedge t) \left\{ a + b \operatorname{sign}(s-t) \log\left(\frac{s}{t}\right) \right\}, \tag{4}$$

where a and b are positive constants. Thus, here the limiting process is not a Brownian motion, though it has the same distribution as a Brownian motion for each fixed t. Our Theorem 2 also corrects Theorem 3.1 of Reynolds (1975) as well as Theorem 1 of Sen (1978).

In the two-sample case, where X_i either has c.d.f H_1 or c.d.f. H_2, the appropriate version of T_n is obtained by substituting $Y_i \equiv 1$ and

$$e_i = \begin{cases} 1 & \text{if } X_i \text{ has c.d.f. } H_1 \\ 0 & \text{if } X_i \text{ has c.d.f. } H_2 \end{cases} \tag{5}$$

in (3). Lai (1975) has given an almost sure representation of T_n in terms of two sums of i.i.d. random variables, and has shown that a suitably constructed continuous time version of $\{T_n\}$ converges weakly to a Brownian motion.

The analogue of T_n in the sequential rank case is obtained from (2) by setting $Y_i \equiv 1$ and

$$c_i = e_i - (i-1)^{-1} \sum_{j=1}^{i-1} e_j \quad \text{for } i \geqq 2 \tag{6}$$

with the e_i's defined as in (5) above. (See Mason (1981, Sect. 2) and Lombard (1981, Sect. 4) for the motivation of the choice of the c_i's given in (6).) Although it is possible to obtain an invariance principle for the two-sample version of M_n, based on the almost sure representation given in Theorem 1 below, the technical details are quite lengthy and the resulting covariance function of the limiting process depends on functions of s and t which are not expressible in closed form. This problem will be considered elsewhere.

In Sect. 2, we state and discuss the theorems. The proofs are detailed in Sect. 3.

2. Statements of Main Results

First, we must introduce some necessary notion and assumptions.

Let (X_i, Y_i) with bivariate distribution function G_i, $i \geqq 1$, be a sequence of independent bivariate random vectors and let F_i denote the marginal distribution function of X_i. The empirical distribution function of $X_1, \ldots, X_i$ is

$$\tilde{F}_i(x) = i^{-1} \sum_{j=1}^{i} I(X_j \leqq x), \qquad -\infty < x < \infty$$

and this has expectation

$$\bar{F}_i(x) = i^{-1} \sum_{j=1}^{i} F_j(x), \qquad -\infty < x < \infty.$$

Since it does not cause any additional complications in the proof of Theorem 1 below, we will consider the following general $k(\geqq 1)$-sample situation:

(A) There is a finite set $\{V_1, \ldots, V_k\}$ of c.d.f.'s, a set of numbers $\rho_1, \ldots, \rho_k \in (0, 1)$ which sum to 1, and a function j: $\{1, 2, \ldots\} \to \{1, 2, \ldots, k\}$ such that $G_i = V_{j(i)}$ for all $i \geqq 1$ and such that as $m \to \infty$,

$$\rho_{r,m} = \frac{1}{m} \sum_{i=1}^{m} I(j(i) = r) \to \rho_r \qquad \text{for } r = 1, 2, \ldots, k.$$

Furthermore, we assume that

(B) For every $i \geqq 1$, F_i is continuous and Y_i has a finite absolute moment of order $1/\alpha$ for some $0 < \alpha < 1/2$, i.e.

$$\max_{1 \leqq l \leqq k} \iint |y|^{1/\alpha} dV_l(x, y) < \infty.$$

We are interested in linear sequential rank statistics of the form given in (2) above, where $c_2, c_3, \ldots,$ is a sequence of constants satisfying the condition

$$\text{(C)} \qquad \max_{2 \leqq i \leqq n} c_i^2 / C_n^2 = O\left(\frac{1}{n}\right)$$

with

$$C_n^2 = \sum_{i=2}^{n} c_i^2 \qquad \text{for } n \geqq 2 \tag{7}$$

and $J(u)$, $u \in (0, 1)$ is a score function satisfying the Chernoff-Savage (1958)-type condition.

$$\text{(D)} \qquad |J^{(i)}(u)| \equiv \left| \frac{d^i}{du^i} J(u) \right| \leqq M(u(1-u))^{-1/2-i+\delta}; \qquad i = 0, 1, 2$$

with $0 < \alpha < \delta < \frac{1}{2}$ for the α in (B) above, and M a finite positive constant.

Let

$$\mu_n = \sum_{i=2}^{n} c_i \iint y J(\bar{F}_i(x)) dG_i(x, y),$$

$$B_{1,n} = \sum_{i=2}^{n} c_i Y_i J(\bar{F}_i(X_i)) - \mu_n, \tag{8}$$

$$B_{2,n} = \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \frac{c_k}{k+1} \iint y \{I(X_i \leqq x) - F_i(x)\} J^{(1)}(\bar{F}_k(x)) dG_k(x, y)$$

and write

$$M_n - \mu_n = B_{1,n} + B_{2,n} + \mathscr{R}_n \equiv B_n + \mathscr{R}_n. \tag{9}$$

Notice that B_n is a sum of independent random variables. The following theorem gives the aforementioned almost sure representation.

Theorem 1. *If conditions* (A) *through* (D) *above hold,*

$$C_n^{-1} \mathscr{R}_n \to 0 \quad \text{a.s.}$$

The proof is postponed until Sect. 3.

Suppose it can be shown that the sequence of processes

$$(C_n^{-1} B_{[nt]}, 0 \leqq t \leqq 1) \tag{10}$$

($B_0 = B_1 \equiv 0$ and $[x]$ = integer part of x) converges weakly in the topology induced by the metric d_0 defined on $D[0,1]$ (see Chapt. 2 of Billingsley (1968) for details) to a process

$$\mathscr{Z} \equiv (Z(t), 0 \leqq t \leqq 1).$$

Theorem 1 then implies that the sequence of processes

$$\mathscr{Z}_n \equiv (C_n^{-1}(M_{[nt]} - \mu_{[nt]}), 0 \leqq t \leqq 1) \tag{11}$$

($M_0 = M_1 = \mu_0 = \mu_1 \equiv 0$) will also converge weakly to the process $\mathscr{Z}$.

Even though B_n is a sum of independent random variables, weak convergence of $C_n^{-1} B_{[nt]}$ is in general extremely difficult to establish without further assumptions regarding the distributions G_i or the form of the constants c_i. This becomes fairly obvious upon close inspection of (8).

With the following simplifications, we are able to give a weak convergence theorem for a version of the sequence of processes given in (11):

Suppose that $c_i \equiv 1$ and that (X_i, Y_i) for $i \geqq 1$ are *identically* distributed. Put $F_i \equiv F$ and $G_i \equiv G$. We then find that

$$\begin{aligned}
&\mu_n = (n-1) \iint y J(F(x))\, dG(x,y) \equiv (n-1)\mu, \\
&B_{1,n} = \sum_{i=2}^{n} \{Y_i J(F(X_i)) - \mu\} \equiv \sum_{i=2}^{n} \eta(Y_i, X_i) \quad \text{and} \\
&B_{2,n} = \sum_{i=1}^{n-1} (d_n - d_i) \iint y \{I(X_i \leqq x) - F(x)\} J^{(1)}(F(x))\, dG(x,y) \\
&\quad \equiv \sum_{i=1}^{n-1} (d_n - d_i)\, \xi(X_i),
\end{aligned} \tag{12}$$

with

$$d_j = \sum_{k=0}^{j} \frac{1}{k+1}; \quad j = 0, 1, \ldots . \tag{13}$$

Let

$$\sigma_\xi^2 = E\xi^2(X_i), \quad \sigma_\eta^2 = E\eta^2(Y_i, X_i), \quad \text{and} \quad \sigma_{\eta\xi} = E\{\eta(Y_i, X_i)\,\xi(X_i)\}. \tag{14}$$

Theorem 2. *Suppose that (X_i, Y_i) for $i \geqq 1$ are i.i.d. and that assumptions* (B) *and* (D) *hold. Let $c_i \equiv 1$ in* (2). *Then the processes $(Z_n(t), 0 \leqq t \leqq 1)$ defined as in* (11) *above converge weakly in the Skorokhod $D[0,1]$-topology to a path-continuous nonstationary mean zero Gaussian process $(Z(t), 0 \leqq t \leqq 1)$ with covariance function given in* (4) *above, with*

$$a = \sigma_\eta^2 + 2\sigma_\xi^2 + 2\sigma_{\eta\xi}$$

and

$$b = \sigma_\xi^2 + \sigma_{\eta\xi}. \tag{15}$$

The proof is postponed until Sect. 3.

The limiting process Z in Theorem 2 may be represented as

$$Z(t) = \int_0^t \left(\sigma_\eta - \sigma_\eta^{-1} \sigma_{\eta\xi} \log\left(\frac{u}{t}\right)\right) dW_1(u) + \sigma_\eta^{-1} \sqrt{\sigma_\eta^2 \sigma_\xi^2 - \sigma_{\eta\xi}^2} \int_0^t \log\left(\frac{u}{t}\right) dW_2(u) \tag{16}$$

where W_1 and W_2 are independent standard Brownian motions and the integrals are interpreted in Ito's sense. It is easy to verify, using for example Theorem (5.1.1) of Arnold (1973), that the process defined by (16) is path-continuous with covariance function given by (4) and (15).

The results of Reynolds (1975) and Sen (1978) imply, for their versions of the statistic given in (2), that the sequence of processes $\mathscr{X}_n$ converges weakly to a Brownian motion. Our Theorem 2 is in obvious conflict with their results. A close inspection of their proofs reveals, however, that they incorrectly evaluated the limiting finite-dimensional distributions. (Note that in Sen (1978), the smoothness assumptions that we given on J in (D) are transferred to the distribution of Y_i, but this does not affect the validity of our discussion.) Using a different method, Müller-Funk (1983) has studied the special case $Y_i = \text{sign}(X_i)$ under conditions comparable to ours. Our Theorem 2, specialized to this case, agrees with his result.

One implication of Theorem 2 from a statistical viewpoint is that analytical approximations to the asymptotic power, under fixed alternatives, of sequential tests based on these statistics would not be readily available. Explicit expressions for boundary crossing probabilities of Gaussian processes with covariance function given by (4) are, to the best of our knowledge, not readily available.

3. Proofs of the Theorems

Proof of Theorem 1. Since for every $i \geqq 1$,

$$R_i = i\hat{F}_i(X_i)$$

we obtain by Taylor expansion that

$$\begin{aligned} M_n - \mu_n = {} & \sum_{i=2}^{n} c_i Y_i J(\bar{F}_i(X_i)) - \mu_n \\ & + \sum_{i=2}^{n} c_i Y_i \left\{ \frac{i}{i+1} \hat{F}_i(X_i) - \bar{F}_i(X_i) \right\} J^{(1)}(\bar{F}_i(X_i)) \\ & + \frac{1}{2} \sum_{i=2}^{n} c_i Y_i \left\{ \frac{i}{i+1} \hat{F}_i(X_i) - \bar{F}_i(X_i) \right\}^2 J^{(2)} \left(\theta_i \frac{i}{i+1} \hat{F}_i(X_i) + (1-\theta_i) \bar{F}_i(X_i) \right) \\ \equiv {} & B_{1,n} + A_n + \sum_{i=2}^{n} c_i \mathscr{R}_{i,1} \end{aligned}$$

with $0 < \theta_i < 1$ for $i = 1, \ldots, n$. Next

$$\begin{aligned} A_n = {} & \sum_{i=2}^{n} \left(\frac{i-1}{i+1} \right) c_i Y_i \{ \hat{F}_{i-1}(X_i) - \bar{F}_{i-1}(X_i) \} J^{(1)}(\bar{F}_i(X_i)) \\ & + \sum_{i=2}^{n} c_i Y_i \left\{ \frac{i}{i+1} \hat{F}_i(X_i) - \frac{i-1}{i+1} \hat{F}_{i-1}(X_i) - \bar{F}_i(X_i) \right. \\ & \left. + \frac{i-1}{i+1} \bar{F}_{i-1}(X_i) \right\} J^{(1)}(\bar{F}_i(X_i)), \end{aligned}$$

which, after a little manipulation, is seen to equal

$$\sum_{i=2}^{n}\left(\frac{i-1}{i+1}\right)c_i\iint y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))d\delta_{(X_i,Y_i)}(x,y)$$
$$+\sum_{i=2}^{n}c_iY_i\frac{1}{i+1}\{1-F_i(X_i)-\bar{F}_i(X_i)\}J^{(1)}(\bar{F}_i(X_i))$$
$$\equiv D_n+\sum_{i=2}^{n}c_i\mathscr{R}_{i,2},$$

where, for every pair of real vectors (a_1, a_2) and (b_1, b_2) we define

$$\delta_{(a_1,a_2)}(b_1,b_2)=1 \quad \text{if } b_1\geqq a_1\, b_2\geqq a_2;\ =0 \quad \text{otherwise.}$$

Below, we shall also write

$$\delta_{a_1}(b_1)=\delta_{a_1,a_2}(b_1,\infty).$$

Finally

$$D_n=\sum_{i=2}^{n}\left(\frac{i-1}{i+1}\right)c_i\iint y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))dG_i(x,y)$$
$$+\sum_{i=2}^{n}\left(\frac{i-1}{i+1}\right)c_i\iint y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))d(\delta_{(X_i,Y_i)}(x,y)-G_i(x,y))$$
$$=\sum_{i=2}^{n}c_i\iint y\frac{1}{i+1}\left\{\sum_{k=1}^{i-1}(\delta_{X_k}(x)-F_k(x))\right\}J^{(1)}(\bar{F}_i(x))dG_i(x,y)+\sum_{i=2}^{n}c_i\mathscr{R}_{i,3}$$
$$=\sum_{k=1}^{n-1}\sum_{i=k+1}^{n}\frac{c_i}{i+1}\iint y\{\delta_{X_k}(x)-F_k(x)\}J^{(1)}(\bar{F}_i(x))dG_i(x,y)+\sum_{i=2}^{n}c_i\mathscr{R}_{i,3}$$
$$=\sum_{i=1}^{n-1}\sum_{k=i+1}^{n}\frac{c_k}{k+1}\iint y\{\delta_{X_i}(x)-F_i(x)\}J^{(1)}(\bar{F}_k(x))dG_k(x,y)+\sum_{i=2}^{n}c_i\mathscr{R}_{i,3}$$
$$=B_{2,n}+\sum_{i=2}^{n}c_i\mathscr{R}_{i,3}.$$

Thus, we obtain the representation given in (9), viz.,

$$M_n-\mu_n=B_{1,n}+B_{2,n}+\mathscr{R}_n$$

with

$$\mathscr{R}_n=\sum_{i=2}^{n}c_i(\mathscr{R}_{i,1}+\mathscr{R}_{i,2}+\mathscr{R}_{i,3})\equiv\Sigma_{1,n}+\Sigma_{2,n}+\Sigma_{3,n}.$$

The proof consists of showing that

$$C_n^{-1}\Sigma_{i,n}\to 0 \text{ a.s.} \qquad \text{for each } 1\leqq i\leqq 3. \tag{17}$$

In order to establish (17), we require the following auxilliary lemmas:

Lemma 1. *Let* $\left\{S_n=\sum_{i=1}^{n} X_i, \mathscr{F}_n\colon n\geqq 1\right\}$ *be an* L_p *martingale for some* $p\in[1,2]$. *If for some sequence of positive constants* $b_n\uparrow\infty$

$$\sum_{n\geqq 1} b_n^{-p} E|X_n|^p<\infty,$$

then

$$b_n^{-1} S_n\to 0 \quad \text{a.s.}$$

(See problem 1, p. 244, of Chow and Teicher (1978).)

Lemma 2. *(Marcinkiewicz-Zygmund law of large numbers.) Let* $X_1, X_2, \ldots,$ *be an* L_p *sequence of* i.i.d.r.v. *for some* $p\in(0,1)$. *Then*

$$n^{-1/p}\sum_{i=1}^{n} X_i\to 0 \quad \text{a.s.}$$

(*See p.* 122 *of Chow and Teicher* (1978).)

For every $\gamma>0$, let

$$\phi_\gamma(u)=(u(1-u))^{-\gamma}, \qquad u\in(0,1).$$

Lemma 3. *If* $U_1, U_2, \ldots,$ *is an* i.i.d. *sequence of Uniform* $[0,1]$ r.v., *then for every* $\alpha>\beta>1$

$$n^{-\alpha}\sum_{i=1}^{n}\phi_\beta(U_i)\to 0 \quad \text{a.s.}$$

Proof. The result follows from Lemma 2 upon noting that

$$E(\phi_\beta(U_1))^{1/\alpha}=E(\phi_{\beta/\alpha}(U_1))<\infty. \quad \square$$

Lemma 4. *Let* $U_i=F_i(X_i)$ *for all* $i\geqq 1$. *For every* $\gamma>0$ *there exists a constant* $K(\gamma)$ *such that*

$$\phi_\gamma(\bar{F}_i(X_i))\leqq K(\gamma)\phi_\gamma(U_i).$$

Proof. For $r=1,\ldots,k$, let

$$n_r=\min\{i\colon G_i=V_r\} \quad \text{and} \quad K=\min_{1\leqq r\leqq k}\inf_{m\geqq n_r}\rho_{r,m}.$$

Clearly, K is positive and finite. Choose any $i\geqq 1$ and assume, without loss of generality, that $G_i=V_1$. Denote the first marginal of V_i by W_i for $i=1,\ldots,k$. Then

$$\bar{F}_i(X_i)=\rho_{1,i}W_1(X_i)+\sum_{r=2}^{k}\rho_{r,i}W_r(X_i)\geqq\rho_{1,i}W_1(X_i)=\rho_{1,i}U_i$$

and

$$\begin{aligned}1-\bar{F}_i(X_i)&=1-\rho_{1,i}W_1(X_i)-\sum_{r=2}^{k}\rho_{r,i}W_r(X_i)\geqq 1-\rho_{1,i}W_1(X_i)-\sum_{r=2}^{k}\rho_{r,i}\\&=\rho_{1,i}-\rho_{1,i}U_i=\rho_{1,i}(1-U_i).\end{aligned}$$

Thus $\bar{F}_i(X_i)(1-\bar{F}_i(X_i)) \geqq K^2 U_i(1-U_i)$, and since the function $z \to z^{-\gamma}$ is nonincreasing, it follows that

$$\phi_\gamma(\bar{F}_i(X_i)) \leqq K^{-2\gamma} \phi_\gamma(U_i). \quad \square$$

Next, let

$$\Delta_n^{(1)} = \sup_{n^{-1} \leqq \bar{F}_n(s) \leqq 1-n^{-1}} \{n^{1/2} |\hat{F}_n(s) - \bar{F}_n(s)| / \{\bar{F}_n(s)(1-\bar{F}_n(s))\}^{1/2}\},$$

$$\Delta_n^{(2)} = \sup_{\bar{F}_n(s),\, 1-\bar{F}_n(s) < n^{-1}} \{n |\hat{F}_n(s) - \bar{F}_n(s)|\},$$

$$r_n^{(1)} = \sup_{s \geqq X_{1,n}} \{\bar{F}_n(s)/\hat{F}_n(s)\}, \quad \text{and}$$

$$r_n^{(2)} = \sup_{s < X_{n,n}} \{(1-\bar{F}_n(s))/(1-\hat{F}_n(s))\},$$

where $X_{1,n} < \ldots < X_{n,n}$ are the order statistics of $X_1, \ldots, X_n$.

Lemma 5. (a) For every $\gamma \geqq 0$ *there exists a* $0 < K < \infty$ *and* $n_0 \geqq 1$ *such that for every* $n \geqq n_0$

$$P(\Delta_n^{(i)} > K \log n) < n^{-(1+\gamma)}; \qquad i = 1, 2.$$

(b) *For every* $\lambda > 1$,

$$P(r_n^{(i)} > \lambda) \leqq 15 \lambda^2 \exp(-\lambda); \qquad i = 1, 2.$$

(See Theorem 3.3 of Ruymgaart and van Zuijlen (1978) and Theorem 2.7 of van Zuijlen (1980).)
Write

$$r_n = \sup_{X_{1n} \leqq s < X_{nn}} \{\bar{F}_n(s)(1-\bar{F}_n(s)) / \{\hat{F}_n(s)(1-\hat{F}_n(s))\}\}.$$

Lemma 6. *There exists a constant* $0 < K < \infty$ *such that, with probability one,*

$$\max\{\Delta_n^{(1)}, \Delta_n^{(2)}, r_n^{1/2}\} \leqq K \log n$$

for all n sufficiently large.

Proof. The proof follows easily from the inequalities given in Lemma 5 together with the Borel-Cantelli lemma. $\square$

Lemma 7.

$$r_n^* \equiv \sup_{X_{1n} \leqq s \leqq X_{nn}} \left\{ \frac{\bar{F}_n(s)(1-\bar{F}_n(s))}{\frac{n}{n+1} \hat{F}_n(s) \left(1 - \frac{n}{n+1} \hat{F}_n(s)\right)} \right\} \leqq 4 r_n \qquad \text{for all } n \geqq 2.$$

Proof. Elementary.

Lemma 8. *Let* $\{b_\nu\}$ *be a sequence of constants and* $\{a_{\nu,n}\colon \nu = 1, \ldots, n,\ n \geqq 1\}$ *a triangular array of constants such that for some finite constant* b,

$$b_\nu \to b,$$

and for each fixed $n_0 \geqq 1$

$$\max_{1 \leqq \nu \leqq n_0} |a_{\nu,n}|/n \to 0. \tag{18}$$

If for some finite constant a,

$$n^{-1} \sum_{\nu=1}^{n} a_{\nu,n} \to a, \tag{19}$$

then

$$n^{-1} \sum_{\nu=1}^{n} a_{\nu,n} b_\nu \to ab.$$

Proof. Elementary.

Lemma 9. *Let* $X_1, X_2, \ldots,$ *be a sequence of random variables such that for some* $r>0$

$$n^{-r} S_n \equiv n^{-r} \sum_{\nu=1}^{n} X_\nu \to 0 \quad \text{a.s.} \tag{20}$$

Then for every $\alpha>0$ *and* $\beta \geqq 0$ *such that* $\alpha+\beta=r$

$$n^{-\alpha} \sum_{\nu=1}^{n} \nu^{-\beta} X_\nu \to 0 \quad \text{a.s.} \tag{21}$$

Proof. Set $b_\nu = \nu^{-r} S_\nu$ for $\nu = 1, 2, \ldots,$ and for each $n \geqq 2$, let

$$a_{\nu,n} = n^{1-\alpha}(\nu^{-\beta} - (\nu+1)^{-\beta})\nu^r$$

for $\nu = 1, \ldots, n-1$ and $a_{n,n} = n$.

Observe that

$$n^{-1} \sum_{\nu=1}^{n} a_{\nu,n} b_\nu = n^{-\alpha} \sum_{\nu=1}^{n} \nu^{-\beta} X_\nu.$$

By (20), $b_\nu \to 0$ a.s. Notice that for $1 \leqq \nu \leqq n-1$,

$$n^{-1} a_{\nu,n} \leqq \beta n^{-\alpha} \nu^{\alpha-1}$$

so that we see that condition (18) of Lemma 8 is satisfied. Finally, a standard integral approximation shows that

$$n^{-1} \sum_{\nu=1}^{n} a_{\nu,n} \to r\alpha^{-1}.$$

Hence condition (19) of Lemma 8 also holds. Thus (21) follows from Lemma 8. □

Now we are ready to complete the proof of Theorem 1. Throughout the proof below, the symbol K is a generic finite positive constant.

Claim 1.

$$C_n^{-1} \Sigma_{3,n} \to 0 \quad \text{a.s.}$$

Proof. For every $v>0$, it follows from the c_r-inequality, Hölder's inequality and the independence, for all $i\geqq 2$, of $\hat{F}_{i-1}(x)$ and X_i that

$$\begin{aligned}
&E|\mathscr{R}_{i,3}|^{1+v}\\
&\leqq \left(\frac{i-1}{i+1}\right)2^v E|\iint y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))d\delta_{(X_i,Y_i)}(x,y)|^{1+v}\\
&\quad+\left(\frac{i-1}{i-1}\right)2^v E|\iint y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))dG_i(x,y)|^{1+v}\\
&\leqq 2^v E\iint|y\{\hat{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))|^{1+v}\delta_{(X_i,Y_i)}(x,y)\\
&\quad+2^v E\iint|y\{\tilde{F}_{i-1}(x)-\bar{F}_{i-1}(x)\}J^{(1)}(\bar{F}_i(x))|^{1+v}dG_i(x,y)\\
&=2^{v+1}\iint|y|^{1+v}|J^{(1)}(\bar{F}_i(x))|^{1+v}E|\tilde{F}_{i-1}(x)-\bar{F}_{i-1}(x)|^{1+v}dG_i(x,y).
\end{aligned}$$

Since

$$E(\tilde{F}_{i-1}(x)-\bar{F}_{i-1}(x))^2\leqq(i-1)^{-2}\sum_{j=1}^{i}F_j(x)(1-F_j(x)),$$

which, by concavity of the function $u\to u(1-u)$ for $u\in(0,1)$, is

$$\leqq i(i-1)^{-2}\bar{F}_i(x)(1-\bar{F}_i(x))\leqq 4i^{-1}\bar{F}_i(x)(1-\bar{F}_i(x)),$$

we obtain from Hölder's inequality for any $0<v\leqq 1$, that

$$E|\tilde{F}_{i-1}(x)-\bar{F}_{i-1}(x)|^{1+v}\leqq\{E|\tilde{F}_{i-1}(x)-\bar{F}_{i-1}(x)|^2\}^{(1+v)/2}\leqq 4i^{-(1+v)/2}\phi_\tau(\bar{F}_i(x))$$

where $\tau=-(1+v)/2$. Hence, by (A), (D) and Lemma 4,

$$E|\mathscr{R}_{i,3}|^{1+v}\leqq Ki^{-(1+v)/2}E\{|Y_i|^{1+v}\phi_\gamma(U_i)\}$$

with

$$\gamma=(1+v)(\tfrac{3}{2}-\delta)-\frac{(1+v)}{2}=(1+v)(1-\delta).$$

Using Hölder's inequality again, this is

$$\leqq Ki^{-(1+v)/2}\{E|Y_i|^{1/\alpha}\}^{\alpha(1+v)}\{E\phi_{\gamma_*}(U_i)\}^{1-\alpha(1+v)},$$

where $\gamma_*=\gamma/(1-\alpha(1+v))$. Note that for $0<v<1$, both $\alpha(1+v)<1$ and $\gamma_*=\gamma/(1-\alpha(1+v))<1$ if and only if $(1+v)(1+\alpha-\delta)<1$, which will be the case for all sufficiently small $v>0$. Thus, for $v>0$ small enough, using (B) we have

$$E|\mathscr{R}_{i,3}|^{1+v}\leqq Ki^{-(1+v)/2}. \tag{22}$$

For $n\geqq 2$, let $\mathscr{F}_n$ denote the σ-field generated by $X_1,\ldots,X_n$. The independence, for all $i\geqq 2$, of $\tilde{F}_{i-1}(x)$ and X_i implies that $E(\mathscr{R}_{i,3}|\mathscr{F}_{i-1})=0$ for all $i\geqq 2$, so that

$$\left\{\left(\sum_{i=2}^{n}c_i\mathscr{R}_{i,3},\mathscr{F}_n\right):n\geqq 2\right\}$$

is a martingale. From (22), we see that it is, in fact, an $L_{1+\nu}$-martingale for some $\nu\in(0,1)$. Since by assumption (C) and (22),

$$\sum_{i\geqq 2} C_i^{-(1+\nu)} E|c_i \mathscr{R}_{i,3}|^{1+\nu} \leqq K \sum_{i\geqq 2} i^{-1-\nu} < \infty,$$

an application of Lemma 1 completes the proof. □

Claim 2.

$$C_n^{-1}\Sigma_{2,n}\to 0 \quad \text{a.s.}$$

Proof. Since $\gamma\equiv\frac{3}{2}-\delta<\frac{3}{2}-\alpha$, it follows from Lemma 3 that

$$n^{-3/2+\alpha}\sum_{i=2}^{n}\phi_\gamma(U_i)\to 0 \quad \text{a.s.}$$

Furthermore, assumption (B) implies that $|Y_i|\leqq i^\alpha$ for all sufficiently large i, a.s. Hence, by Lemma 4 and Lemma 9

$$C_n^{-1}\left|\sum_{i=2}^{n} c_i\mathscr{R}_{i,2}\right|\leqq Kn^{-1/2}\sum_{i=2}^{n} i^{-1+\alpha}\phi_\gamma(U_i)\to 0 \quad \text{a.s.} \quad \square$$

Claim 3.

$$C_n^{-1}\Sigma_{1,n}\to 0 \quad \text{a.s.}$$

Proof. By assumption (D), the convexity of ϕ_γ for all $\gamma>0$, the fact that $0<\theta_i<1$ and from Lemma 6 and Lemma 7 we have, almost surely, that

$$\begin{aligned}
&\left|J^{(2)}\left(\theta_i\frac{i}{i+1}\tilde{F}_i(X_i)+(1-\theta_i)\bar{F}_i(X_i)\right)\right|\\
&\leqq M\phi_{5/2-\delta}\left(\theta_i\frac{i}{i+1}\tilde{F}_i(X_i)+(1-\theta_i)\bar{F}_i(X_i)\right)\\
&\leqq M\phi_{5/2-\delta}\left(\frac{i}{i+1}\tilde{F}_i(X_i)\right)+\phi_{5/2-\delta}(\bar{F}_i(X_i))\\
&\leqq M\{4^{5/2-\delta}r_i^{5/2-\delta}\phi_{5/2-\delta}(\bar{F}_i(X_i))+\phi_{5/2-\delta}(\bar{F}_i(X_i))\}\\
&\leqq K(\log i)^{5-2\delta}\phi_{5/2-\delta}(\bar{F}_i(X_i)),
\end{aligned}$$

for all sufficiently large i. (Recall the definition of r_i given just prior to Lemma 6.)

Next, by Lemma 6, we have, almost surely,

$$\begin{aligned}
&\left(\frac{i}{i+1}\tilde{F}_i(X_i)-\bar{F}_i(X_i)\right)^2\\
&\leqq 2\left(\frac{i}{i+1}\tilde{F}_i(X_i)-\frac{i}{i+1}\bar{F}_i(X_i)\right)^2+2\left(\frac{i}{i+1}\bar{F}_i(X_i)-\bar{F}_i(X_i)\right)^2\\
&\leqq 2(\tilde{F}_i(X_i)-\bar{F}_i(X_i))^2+2i^{-2}
\end{aligned}$$

$$
\begin{aligned}
&=2(\tilde{F}_i(X_i)-\bar{F}_i(X_i))^2 I(i^{-1}\leqq \bar{F}_i(X_i)\leqq 1-i^{-1})\\
&\quad +2(\tilde{F}(X_i)-\bar{F}_i(X_i))^2 I(\bar{F}_i(X_i)<i^{-1} \quad \text{or} \quad \bar{F}_i(X_i)>1-i^{-1})+2i^{-2}\\
&\leqq 2i^{-1}\bar{F}_i(X_i)(1-\bar{F}_i(X_i))(\Delta_i^{(1)})^2+2i^{-2}(\Delta_i^{(2)})^2+2i^{-2}\\
&\leqq Ki^{-1}(\log i)^2\bar{F}(X_i)(1-\bar{F}(X_i))+Ki^{-2}(\log i)^2
\end{aligned}
$$

for all sufficiently large i.

Putting these together and using Lemma 4, we have, almost surely,

$$
\begin{aligned}
|\mathscr{R}_{i,1}|&\leqq Ki^{\alpha}(\log i)^{5-2\delta}\phi_{5/2-\delta}(\bar{F}_i(X_i))\{i^{-1}(\log i)^2\bar{F}_i(X_i)(1-\bar{F}_i(X_i))+i^{-2}(\log i)^2\}\\
&\leqq Ki^{-1+\alpha}(\log i)^{7-2\delta}\phi_{3/2-\delta}(U_i)+Ki^{-2+\alpha}(\log i)^{7-2\delta}\phi_{5/2-\delta}(U_i)\\
&\leqq K^{i-1+\alpha+r}\phi_{\beta-1}(U_i)+Ki^{-2+\alpha+r}\phi_{\beta}(U_i)
\end{aligned}
$$

for all sufficiently large i, where $\beta=5/2-\delta$ and r is chosen so that $0<r<\delta-\alpha$. For such a choice of r we have $3/2-r-\alpha>\beta-1$ and $5/2-r-\alpha>\beta$. Hence, by Lemma 3, both

$$n^{-3/2+r+\alpha}\sum_{i=2}^{n}\phi_{\beta-1}(U_i)\to 0 \quad \text{a.s.,}$$

and

$$n^{-5/2+r+\alpha}\sum_{i=2}^{n}\phi_{\beta}(U_i)\to 0 \quad \text{a.s.}$$

Since, by using assumption (C), we have

$$
\begin{aligned}
&C_n^{-1}\left|\sum_{i=2}^{n}c_i\mathscr{R}_{i,1}\right|\\
&\quad\leqq Kn^{-1/2}\sum_{i=2}^{n}i^{-1+r+\alpha}\phi_{\beta-1}(U_i)+Kn^{-1/2}\sum_{i=2}^{n}i^{-2+r+\alpha}\phi_{\beta}(U_i) \quad \text{a.s.}
\end{aligned}
$$

for all sufficiently large i, an application of Lemma 8 completes the proof. □

Thus, (17) follows and Theorem 1 is proven. □

Proof of Theorem 2. First, it is easy to show that the array

$$b_{in}=\eta(Y_{i+1},X_{i+1})+(d_n-d_i)\xi(X_i), 1\leqq i\leqq n-1; \quad n\geqq 2$$

satisfies Lindeberg's condition. (Recall the definitions of η, ξ and the d_i's given in (12) and (13).) Next, let $B_0=B_1=0$ and for $n\geqq 2$ set

$$B_n=\sum_{i=1}^{n-1}b_{i,n}=B_{1,n}+B_{2,n}.$$

(Recall the definition of $B_{1,n}$ and $B_{2,n}$ given in Sect. 2.) We find for all $0<s<t\leqq 1$ and sufficiently large n,

$$
\begin{aligned}
n^{-1}E(B_{[ns]}B_{[nt]})&=\sigma_\eta^2 n^{-1}[ns]+\sigma_{\eta\xi}n^{-1}\sum_{j=1}^{[ns]}(d_{[ns]}-d_j)\\
&\quad+\sigma_{\eta\xi}n^{-1}\sum_{j=1}^{[ns]}(d_{[nt]}-d_j)+\sigma_\xi^2 n^{-1}\sum_{j=1}^{[ns]}(d_{[ns]}-d_j)(d_{[nt]}-d_j),
\end{aligned}
$$

which by an easy limiting argument converges to

$$\sigma_\eta^2 s - \sigma_{\eta\xi}\int_0^s \log\left(\frac{u}{s}\right)du - \sigma_{\eta\xi}\int_0^s \log\left(\frac{u}{t}\right)du + \sigma_\xi^2\int_0^s \log\left(\frac{u}{s}\right)\log\left(\frac{u}{t}\right)du.$$

After some straightforward manipulations, this last expression can be shown to equal

$$s\left\{a+b\,\mathrm{sign}(s-t)\log\left(\frac{s}{t}\right)\right\},$$

where a and b are defined as in the statement of Theorem 2.

Thus, by applying the Cramér-Wold device, the finite dimensional distributions of the process

$$(n^{-1/2}B_{[nt]}, 0\leqq t\leqq 1) \tag{23}$$

converge to those of a Gaussian process with covariance function given by (4) and (15). It remains to prove tightness. Towards this, let

$$l_n(t)=d_{[nt]}-d_n, \qquad 0\leqq t\leqq 1$$

and write

$$B_{[nt]}=B_{1,n}^*(t)+B_{2,n}^*(t) \tag{24}$$

where

$$B_{1,n}^*(t)=\sum_{i=1}^{[nt]-1} b_{in} \qquad \text{for } [nt]\geqq 2;\ =0 \text{ otherwise}$$

and

$$B_{2,n}^*(t)=l_n(t)\sum_{i=1}^{[nt]-1}\xi(X_i) \qquad \text{for } [nt]\geqq 2;\ =0 \text{ otherwise.}$$

The validity of Lindeberg's condition in conjunction with Theorem 3 of Prohorov (1956) shows that the sequence of processes

$$(n^{-1/2}B_{1,n}^*(t)\colon 0\leqq t\leqq 1)$$

is tight. In view of (24), in order to complete the proof, it suffices to show that the sequence of processes

$$(n^{-1/2}B_{2,n}^*(t), 0\leqq t\leqq 1)$$

is tight.

Let $(W(t), 0\leqq t\leqq 1)$ be a standard Brownian motion and let $0<\varepsilon<1$. By Donsker's theorem and the uniform convergence of $l_n(t)$ to $\log t$ on $[\varepsilon, 1]$ we have that the sequence of processes

$$(n^{-1/2}B_{2,n}^*(t), \varepsilon\leqq t\leqq 1)$$

converges weakly to the process

$$(\sigma_\xi W(t)\log t, \varepsilon\leqq t\leqq 1).$$

Next, by the Hàjek-Rényi inequality, for every $\delta>0$ and all n sufficiently large

$$P\{\sup_{0\leqq t\leqq \varepsilon}|n^{-1/2}B^*_{2,n}(t)|>\delta\}\leqq\delta^{-2}\sigma_\xi^2 n^{-1}\sum_{k=1}^{[n\varepsilon]-1}(d_k-d_n)^2,$$

which converges to

$$\delta^{-2}\sigma_\xi^2\int_0^\varepsilon(\log u)^2\,du; \tag{25}$$

and by applying the Birnbaum-Marshall (1961) inequality

$$P\{\sup_{0\leqq t\leqq\varepsilon}|\sigma_\xi W(t)\log t|>\delta\}\leqq\delta^{-2}\sigma_\xi^2\int_0^\varepsilon(\log u)^2\,du.$$

Since expression (25) converges to zero as $\varepsilon\downarrow 0$, we have by applying Theorem 4.2 of Billingsley (1968) in the obvious manner, that the sequence of processes

$$(n^{-1/2}B^*_{2,n}(t), 0\leqq t\leqq 1)$$

converges weakly to the process

$$(W(t)\log t, 0\leqq t\leqq 1).$$

Tightness of the sequence of processes given in (23) is now immediate. □

References

Arnold, L.: Stochastic Differential Equations. New York: J. Wiley 1973

Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968

Birnbaum, Z.W., Marshall, A.W.: Some multivariate Čebišev inequalities with extensions to continuous parameter processes. Ann. Math. Stat. **32**, 687-703 (1961)

Chernoff, H., Savage, I.R.: Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Stat. **29**, 972-994 (1958)

Chow, Y.S., Teicher, H.: Probability Theory. New York: Springer 1978

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

Lombard, F.: An invariance principle for sequential nonparametric test statistics under contiguous alternatives. South African Stat. J. **15**, 129-152 (1981)

Lombard, F.: Asymptotic distributions of rank statistics in the change-point problem. South African Stat. J. **17**, 83-105 (1983)

Mason, D.M.: On the use of a statistic based on sequential ranks to prove limit theorems for simple linear rank statistics. Ann. Stat. **9**, 424-436 (1981)

Mason, D.M.: A Bahadur efficiency comparison between one and two sample rank statistics and their sequential rank statistical analogues. J. Multivariate Anal. **14**, 181-200 (1984)

Müller-Funk, U.: Sequential signed rank statistics. Commun. Stat. Sequential Analysis **2**(2), 123-148 (1983)

Parent, E.A., Jr.: Sequential ranking procedures. Technical report No. **80**, Department of Statistics, Stanford University. (1965)

Prohorov, Yu.V.: Convergence of random processes and limit theorems in probability theory. Theor. Probab. Appl. **1**, 157-214 (1956)

Reynolds, M.R., Jr.: A sequential signed-rank test for symmetry. Ann. Stat. **3**, 382-400 (1975)

Ruymgaart, F.H., Zuijlen, M.C.A. van: On convergence of the remainder term in linear combinations of functions of order statistics in the non-i.i.d. case. Sankhyā, Series A, **40**, 369–387 (1978)

Sen, P.K.: Invariance principles for rank discounted partial sums. Z. Wahrscheinlichkeitstheor. Verw. Geb. **42**, 341-352 (1978)

Sen, P.K.: Sequential Nonparametrics. New York: Wiley 1981

Zuijlen, M.C.A. van: Properties of the empirical distribution function for independent non-identically distributed random vectors. Report **8004**, Math. Inst. Katholieke Universiteit, Nijmegen 1980

Received March 30, 1985