Volume 9, No. 4 (2000), pp. 399–414 Allerton Press, Inc.

MATHEMATICAL METHODS OF STATISTICS

SOME RESULTS FOR EMPIRICAL PROCESSES OF LOCALLY DEPENDENT ARRAYS

JOHN H. J. EINMAHL[1] AND FRITS H. RUYMGAART[2]

[1]Dept of Math. and Computing Science, Eindhoven Univ. of Technology & EURANDOM
PO Box 513, 5600 MB Eindhoven, The Netherlands
E-mail: j.h.j.einmahl@tue.nl

[2]Dept of Mathematics and Statistics, Texas Tech University
Lubbock, TX 79409, USA
E-mail: ruymg@math.ttu.edu

In this paper we derive some fundamental properties of locally dependent arrays of order $m(n)$, where $m(n)$ is allowed to tend to infinity with the sample size n. More specifically we consider a central limit theorem, an exponential inequality for the local fluctuations of the empirical process, and weak convergence of the empirical process. Locally dependent arrays are of independent interest, but they may also serve as useful approximations to other stochastic processes. Some applications are indicated.

Key words: local dependence of order $m(n)$, central limit theorem, empirical process, local fluctuation inequality, weak convergence.

AMS 1991 Subject Classification: Primary 62G30, secondary 60F05.

1. Introduction and Motivation

On the one hand mixing concepts are elegant and powerful tools for suitable description of the dependence structure of large classes of time series. On the other hand, however, mixing may be hard to intuitively understand and to deal with, as it may not even apply to otherwise well behaved processes like certain linear processes (Andrews [1]). A useful alternative to the analysis of time series via mixing is provided by approximation with a simpler process. The approximating

[1]Research partially supported by NATO Collaborative Research Grant CRG 960097 and by the Netherlands Organization for Scientific Research (NWO).

[2]Research partially supported by European Union HCM grant ERB CHRX-CT 940693.

processes that will be considered here are so-called locally dependent processes or, rather, locally dependent arrays, to be defined in Section 2. Such processes are also of interest in their own right.

Versions of locally dependent processes have been regularly studied in the literature: see, e.g., Hoeffding [11], Billingsley [5], Berk [4], Zetterqvist [20]. The name has been coined by Barbour [2]; see also Reinert [17]. Independently these processes were studied in Chanda and Ruymgaart [8, 9], Nieuwenhuis and Ruymgaart [13], and Nieuwenhuis [12], where they were introduced under less appealing names and where the emphasis was on approximation of linear processes. Using an exponential fluctuation inequality for the approximating process, these authors obtained for instance rates of convergence for density and autocovariance estimators. This fluctuation inequality was not sharp enough to yield tightness of the empirical process of the approximating array. In Section 4 we will sharpen the inequality and employ it in Section 5 to prove weak convergence of the empirical process of a class of locally dependent arrays. Portnoy [16] exploited approximation by means of locally dependent processes studying regression quantiles in non-stationary time series.

In this paper we will focus on locally dependent arrays as a topic of independent interest. Since the definition allows for rather strong local dependence a simple yet representative example of a locally dependent array that will serve as an illustration throughout this paper is given by

$$(1.1) \qquad X_{n,i} = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \varepsilon_{i+k-1}, \qquad i = 1, 2, \ldots, n, \qquad m = m(n),$$

where $\varepsilon_1, \varepsilon_2, \ldots$ is an infinite sequence of i.i.d. variables with mean 0 and variance 1. We will consider the triangular array $X_{n,1}, \ldots, X_{n,n}$, and let $m = m(n)$ depend on n in such a way that

$$(1.2) \qquad m \to \infty \quad \text{and} \quad \frac{m}{n} \to 0 \quad \text{as} \quad n \to \infty.$$

Both these conditions make sense from an approximation perspective. The most important property of the array (1.1) is that its elements are $(m-1)$-dependent. It should be noted, however, that the order of the dependence m grows indefinitely with n.

Although we don't want to dwell on the aspect of approximation, we should mention that linear processes and, more generally, processes with Volterra expansions, provide a natural motivation for locally dependent arrays. For instance, if we have a sample from a linear process

$$(1.3) \qquad X_i = \sum_{k=-\infty}^{\infty} a_k \varepsilon_{i-k}, \qquad i = 1, \ldots, n,$$

where the ε_i are i.i.d. mean 0 and variance 1, say, an approximation is given by the random variables

$$(1.4) \qquad X_{n,i} = \sum_{k=-m}^{m} a_k \varepsilon_{i-k}, \qquad i = 1, \ldots, n, \quad m = m(n).$$

In a large sample setting this approximation will be only useful if $m \to \infty$, as $n \to \infty$. On the other hand, in order to take advantage of the $2m$-dependence of the $X_{n,i}$ we will need $m/n \to 0$ as $n \to \infty$.

In Section 6 we will return to this example because it presents an interesting illustration of the theory which holds some surprises. We restrict ourselves to a_k decaying as a power of k. It then turns out that the approximation cannot be good for processes that are long range dependent (Beran [3]). This means, for instance, that we cannot infer asymptotic normality of the sample average of the original process from the sample average of the approximating process. Since it is known from the literature that the former asymptotic normality may very well fail to hold, there is no contradiction with our asymptotic normality result for the latter (Sections 3 and 6).

Our main result is weak convergence of the empirical process of locally dependent arrays (Section 5). For the weak convergence of the finite dimensional distributions (fidi's) of the empirical process asymptotic normality of row averages is of fundamental importance. This problem is considered in Section 3 where a theorem in Berk [4] is generalized. It should be stressed that we deliberately want to avoid any of the usual mixing conditions so that our results are intended to be complementary to those in Bosq [6], Peligrad [14], and Bradley [7] as far as sums of random variables are concerned and to those in Shao and Yu [18] regarding the empirical process. For the tightness we will derive a fluctuation inequality, already mentioned above, in Section 4.

2. Locally Dependent Arrays

A triangular array $\{X_{n,i}, i = 1, \ldots, n, n \in \mathbb{N}\}$ is called *locally dependent of order* m if the variables in the n-th row

$$X_{n,1}, \ldots, X_{n,n} \qquad \text{are } (m-1)\text{-dependent,} \tag{2.1}$$

for some integer $1 \le m \le n$. As part of the definition we will also require that

$$m = m(n) \quad \text{with} \quad \frac{m}{n} \to 0 \quad \text{as} \quad n \to \infty. \tag{2.2}$$

We are especially interested in the case where $m \to \infty$ as $n \to \infty$ (cf. (1.2)), but if the arrays are not used for approximation this condition will not be needed. Local dependence of order $m = 1$ reduces to independence.

Example 2.1. The array in (1.1) is locally dependent of order m. The array is stationary in the strict sense. In order to quantify the degree of dependence let us calculate the correlation function. For convenience we will restrict ourselves to the unrestricted array $i \in \mathbb{N}$. Because $\mathbf{Var}\, X_{n,i} = 1$ the correlation function equals the covariance function

$$\gamma_n(h) = \mathbf{E} X_{n,i} X_{n,i+h}, \qquad h = 0, 1, \ldots \tag{2.3}$$

Clearly $\gamma_n(0) = 1$, and for $1 \le h \le m-1$ we have

$$\gamma_n(h) = \frac{1}{m} \sum_{k=1}^{m} \sum_{\ell=1}^{m} \delta_{k,\ell+h} = 1 - \frac{h}{m}. \tag{2.4}$$

Summarizing, we have

$$\gamma_n(h) = \begin{cases} 1 - h/m, & h = 0, \dots, m-1, \\ 0, & h = m, m+1, \dots \end{cases} \tag{2.5}$$

3. A Central Limit Theorem

In this section asymptotic normality of row sums will be proved. At this level of generality even in the central limit theorem for independent summands the existence of the second moment is not sufficient. Here we opt for the Lyapunov condition because it is much more convenient to work with in practice than other conditions. This entails in particular that the row variables are supposed to have a moment of order slightly larger than 2. It will be briefly pointed out that in the case of strictly stationary rows, existence of a second moment suffices, provided that a rather awkward Lindeberg-type condition is also fulfilled. Before listing sets of sufficient conditions, let us first consider the special array (1.1) to get an insight into a possible order of magnitude of row sums.

Example 3.1. For array (1.1) we have $(1 \le i \le i + k \le n)$

$$\mathbf{Var}\Big(\sum_{j=1}^{k} X_{n,i+j}\Big) = \mathbf{Var}\Big(\sum_{j=1}^{k} X_{n,j}\Big) = k + 2\sum_{i=1}^{k\wedge m}(k-i)\Big(1 - \frac{i}{m}\Big). \tag{3.1}$$

This means that

$$\mathbf{Var}\Big(\sum_{j=1}^{k} X_{n,i+j}\Big) \le C(k \wedge m)k, \tag{3.2}$$

and

$$\mathbf{Var}\Big(\sum_{j=1}^{k} X_{n,i+j}\Big) \ge C(k \wedge m)k, \tag{3.3}$$

for some generic $C \in (0, \infty)$.

Assumption 3.1. The locally dependent array defined in (2.1) and (2.2) satisfies

$$\mathbf{E}|X_{n,i}|^{2+\delta} \le C, \qquad \mathbf{E}X_{n,i} = 0, \quad i = 1, \dots, n, \quad n \in \mathbb{N}, \tag{3.4}$$

for some C, $\delta \in (0, \infty)$. Moreover, we have $s_n^2 = \mathbf{Var}\left(\sum_{i=1}^{n} X_{n,i}\right) > 0$, and for some $K \in (0, \infty)$

$$\mathbf{Var}\Big(\sum_{j=1}^{k} X_{n,i+j}\Big) \le K\frac{k}{n}s_n^2. \tag{3.5}$$

Furthermore we have

$$nm^{1+\delta}/s_n^{2+\delta} \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.6}$$

Theorem 3.1. *Let Assumption* 3.1 *be satisfied. We then have*

$$\frac{1}{s_n}\sum_{i=1}^{n} X_{n,i} \to_d N(0,1) \qquad \text{as} \quad n \to \infty. \tag{3.7}$$

Proof. Let us choose integers $q = q(n)$ such that $q > 2m$, $q/m \to \infty$, and $q/n \to 0$ as $n \to \infty$. It is convenient to introduce blocks and gaps defined by

$$B_1 = \sum_{i=1}^{q} X_{n,i}, \quad G_1 = \sum_{i=q+1}^{q+m-1} X_{n,i}, \quad B_2 = \sum_{i=q+m}^{2q+m-1}, \quad \text{etc.} \tag{3.8}$$

This means that we can write

$$\sum_{i=1}^{n} X_{n,i} = \sum_{j=1}^{\nu} B_j + \sum_{j=1}^{\nu} G_j + R, \tag{3.9}$$

where $\nu = \lfloor n/(q+m) \rfloor$ and R contains the last $n - \nu(q+m)$ of the $X_{n,i}$. It is obvious that the B_i are mutually independent, and that also the G_i are mutually independent.

We see at once from (3.5) that

$$\mathbf{Var}\left(\frac{R}{s_n}\right) \le K \cdot \frac{q+m}{n} \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.10}$$

Similarly it follows that

$$\mathbf{Var}\left(\frac{\sum_{j=1}^{\nu} G_j}{s_n}\right) = \sum_{j=1}^{\nu} \mathbf{Var}\left(\frac{G_j}{s_n}\right) \le \left\lfloor \frac{n}{q+m} \right\rfloor \cdot K \cdot \frac{m}{n} \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.11}$$

It follows from (3.10) and (3.11) that

$$\frac{1}{s_n}\Big(\sum_{j=1}^{\nu} G_j + R\Big) \to_p 0 \qquad \text{as} \quad n \to \infty. \tag{3.12}$$

Apparently, to prove (3.7) it suffices to prove that

$$\frac{1}{s_n}\sum_{j=1}^{\nu} B_j \to_d N(0,1) \qquad \text{as} \quad n \to \infty, \tag{3.13}$$

in view of (3.12). We will use C as a generic constant.

To prove this the Lyapunov condition will be employed. Let us first note that obviously $\nu(q+m)/n \to 1$, so that $\nu q/n \to 1$, because $m/q \to 0$. Since $\left\{\mathbf{Var}\left(\sum_{j=1}^{\nu} B_j + \sum_{j=1}^{\nu} G_j + R\right)\right\}/s_n^2 = 1$, it follows from (3.5), (3.10), and (3.11) that

$$\frac{\sum_{j=1}^{\nu} \mathbf{Var}\, B_j}{s_n^2} \to 1 \qquad \text{as} \quad n \to \infty. \tag{3.14}$$

Minkowski's inequality and (3.4) imply

$$\sum_{j=1}^{\nu} \mathbf{E}|B_j|^{2+\delta} \le C\nu q^{2+\delta}. \tag{3.15}$$

It follows from (3.14) and (3.15) that

$$\frac{\sum_{j=1}^{\nu} \mathbf{E}|B_j|^{2+\delta}}{(\sum_{j=1}^{\nu} \operatorname{Var} B_j)^{1+\delta/2}} \le C\frac{\nu q^{2+\delta}}{s_n^{2+\delta}} \le C\frac{\nu q \cdot q^{1+\delta}}{s_n^{2+\delta}} \le C\frac{nq^{1+\delta}}{s_n^{2+\delta}}. \tag{3.16}$$

Since (3.10)–(3.12) hold true for each choice of q as long as $q/m \to \infty$ and $q/n \to 0$, and because by (3.6) it is possible to choose q's satisfying these requirements such that also $nq^{1+\delta}/s_n^{2+\delta} \to 0$, the desired result follows. □

Remark 3.1. Let us first briefly comment on the moment condition. Even for triangular arrays of independent variables the existence of a second moment is not sufficient. Hence it is not surprising that when dealing with a fairly arbitrary array of dependent random variables an extra condition is needed. Aiming at Lyapunov's version of the central limit theorem leads to (3.4) and (3.6). Clearly some sharpness is lost in (3.15) by applying the Minkowski inequality. To get an insight into the effect of this inequality let us apply it to the second moment rather than the $(2+\delta)$-moment. This would yield an upper bound of order $\nu q^2 \sim nq$ rather than s_n^2 as in (3.14). Observe that s_n^2 will be of smaller order than nq. On the other hand, in the model of Example 3.1 we have $s_n^2 \sim mn$. Although this is of smaller order than nq because $q/m \to \infty$, this difference does not seem to be too significant because q may be chosen close to m.

Remark 3.2. It should be noted that we have not thus far assumed any stationarity. This assumption, rather common in time series, is satisfied rowwise for the arrays in (1.1) and (1.4). As far as a useful reduction of Assumption 3.1 is concerned, we do not gain a lot by assuming that the array is rowwise strictly stationary and that the second rather than $(2+\delta)$-moments exist; condition (3.5) will still be needed.

In fact under these assumptions we can proceed as in the proof of Theorem 3.1 and see that it suffices to show (3.13), and that (3.14) still holds true. Due to the stationarity for each n the $B_1, \ldots, B_\nu$ are now i.i.d. with common c.d.f. G_n, say. It follows from (3.5) that the common variance of the B_j satisfies

$$\sigma_n^2 = \int_{-\infty}^{\infty} x^2 dG_n(x) \le K\frac{q}{n}s_n^2. \tag{3.17}$$

According to the Lindeberg–Feller central limit theorem it suffices to show that, for each $\varepsilon > 0$,

$$\begin{aligned} \frac{1}{s_n^2}\sum_{j=1}^{\nu} \mathbf{E}B_j^2 1_{\{|B_j| \ge \varepsilon s_n\}} &= \frac{\nu}{s_n^2}\int_{\{|x|\ge \varepsilon s_n\}} x^2\, dG_n(x) \\ &\le K\frac{\nu}{s_n^2}\frac{q}{n}s_n^2\frac{1}{\sigma_n^2}\int_{\{|x|\ge \varepsilon s_n\}} x^2\, dG_n(x) \\ &\le K\frac{\int_{\{|x|\ge\varepsilon s_n\}} x^2\, dG_n(x)}{\int_{-\infty}^{\infty} x^2\, dG_n(x)} \to 0 \quad \text{as} \quad n \to \infty. \end{aligned} \tag{3.18}$$

Although s_n^2 is of larger order than σ_n^2 as can be seen from (3.17), the last expression on the right in (3.18) will not in general tend to zero. Hence (3.18) must be added as a condition which is rather awkward. To avoid a condition of this type we will almost inevitably have to return to condition (3.4). It should be emphasized that for our main application, Theorem 5.1 below on the weak convergence of empirical processes, any kind of moment condition is automatically fulfilled, since the 'building blocks' of empirical processes are bounded, anyway.

Remark 3.3. By comparing the variance s_n^2 with the variance in the i.i.d. case one might call s_n^2/n the excess of the variance. The excess may be smaller than 1 in cases of negative dependence but is of order m, i.e., of the same order as the dependence for array (1.1); see also (3.2) and (3.3). In cases where the order of the excess is as big as the order of the dependence we have

$$(3.19) \qquad \frac{nm^{1+\delta}}{s_n^{2+\delta}} = O\left(\frac{nm^{1+\delta}}{(nm)^{1+\delta/2}}\right) = O\left(\left(\frac{m}{n}\right)^{\delta/2}\right) \to 0 \qquad \text{as} \quad n \to \infty,$$

by (2.2), so that then (3.6) is automatically fulfilled.

Example 3.2. Let us return to model (1.1) and recall that (3.2) entails that condition (3.5) is fulfilled in this case. Since the order of the excess equals that of the dependence we have seen in Remark 3.3 that (3.6) is also fulfilled. If the ε_i are i.i.d. standard normal it follows that each $X_{n,i}$ has a standard normal distribution, so that (3.4) will be satisfied as well. If, more generally, we only assume that $\mathbf{E}|\varepsilon_i|^{2+\delta} < \infty$, we can show by a non-uniform Berry–Esseen inequality (see, e.g., Shorack and Wellner [19], p. 849, Theorem 3) and the fact that $\mathbf{E}|Y|^\alpha = \int_0^\infty \mathbf{P}(|Y| > y^{1/\alpha})\, dy$, that $\mathbf{E}|X_{n,i}|^{2+\delta'} < \infty$, for some $0 < \delta' < \delta$. Hence (3.4) is satisfied.

In the spirit of this example and Remark 3.3, the following result is a convenient, useful specialization of Theorem 3.1.

Corollary 3.1. *Let* (3.4) *be satisfied. Furthermore assume*

$$(3.20) \qquad \frac{s_n^2}{n(2m-1)} \to \sigma^2 \in [0, \infty),$$

then

$$(3.21) \qquad \frac{1}{\sqrt{n(2m-1)}} \sum_{i=1}^{n} X_{n,i} \to_d N(0, \sigma^2) \qquad \textit{as} \quad n \to \infty.$$

Proof. In case $\sigma^2 > 0$, it suffices to show that (3.5) and (3.6) are satisfied. Condition (3.5), for n large enough, follows from (3.20) and the mere fact that the $X_{n,i}$ are $(m-1)$-dependent. Condition (3.6) is trivially fulfilled, see (3.19). In case $\sigma^2 = 0$ the proof follows along the same lines as that of Theorem 3.1, but instead of showing that the Lyapunov condition holds, a direct application of the Chebyshev inequality suffices. □

4. A Fluctuation Inequality for the Empirical Process

Fluctuation inequalities for empirical processes are important tools in the analysis of nonparametric density, regression, and autocovariance functions. They are also very useful for proving tightness of the process. The results in this section are valid for arbitrary but fixed sample size. The only assumption that will be needed is that the variables in each row have the same distribution.

Assumption 4.1. For each $n \in \mathbb{N}$ the random variables $X_{n,1}, \ldots, X_{n,n}$ are identically distributed, i.e.,

$$\mathbf{P}\{X_{n,i} \leq x\} = F_n(x), \qquad x \in \mathbb{R}, \qquad i = 1, \ldots, n. \tag{4.1}$$

For the empirical c.d.f. we will employ the usual notation

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty,x]}(X_{n,i}), \qquad x \in \mathbb{R}, \tag{4.2}$$

and the discrepancy between empirical and actual c.d.f. will be denoted

$$\Delta_n = \widehat{F}_n - F_n. \tag{4.3}$$

Theorem 4.1. *Let $a_0 < b_0$ with $0 < F_n(b_0) - F_n(a_0) \leq \frac{1}{2}$. Then we have, for any $\varepsilon \in (0,1)$,*

$$\begin{aligned} &\mathbf{P}\{\sup_{a_0 \leq a < b \leq b_0} |\Delta_n(b) - \Delta_n(a)| \geq \lambda\} \\ &\leq C(\varepsilon) \exp\left(\frac{-(1-\varepsilon)n\lambda^2}{2m\{F_n(b_0) - F_n(a_0)\}} \psi\left(\frac{\sqrt{n}\lambda}{\sqrt{m \lfloor \frac{n}{m} \rfloor} \{F_n(b_0) - F_n(a_0)\}}\right)\right), \quad \lambda \geq 0. \end{aligned} \tag{4.4}$$

where $0 < C(\varepsilon) < \infty$ is a constant depending on ε only, and

$$\psi(x) = 2x^{-2}\{(1+x)\log(1+x) - x\}, \quad x > 0, \quad \textit{and} \quad \psi(0) = 1. \tag{4.5}$$

Remark 4.1. The function ψ has the following properties: ψ is decreasing and continuous; $\psi(x) \sim (2 \log x)/x$ as $x \to \infty$; $\psi(x) \geq 1/(1 + x/3)$. See pp. 440–441 in Shorack and Wellner [19]. Note also that the condition $F_n(b_0) - F_n(a_0) \leq \frac{1}{2}$ can be weakened to $F_n(b_0) - F_n(a_0) \leq c_0 < 1$; see Einmahl [10], p. 10.

Proof. Let us write, for brevity, $I_0 := (a_0, b_0]$, $I := (a, b]$, $F_n\{I_0\} := F_n(b_0) - F_n(a_0)$, $\widehat{F}_n\{I_0\} := \widehat{F}_n(b_0) - \widehat{F}_n(a_0)$, $\Delta_n\{I_0\} := \Delta_n(b_0) - \Delta_n(a_0)$, and similarly for I. First note that

$$\mathbf{P}\{\sup_{I \subset I_0} |\Delta_n\{I\}| \geq \lambda\} \leq \mathbf{P}\{\sup_{I \subset I_0} \Delta_n\{I\} \geq \lambda\} + \mathbf{P}\{\sup_{I \subset I_0} (-\Delta_n\{I\}) \geq \lambda\}. \tag{4.6}$$

We will only consider the first term on the right in (4.6), the second one can be treated similarly. In fact it is easy to see that it suffices to prove the inequality with '$\geq \lambda$' replaced by '$> \lambda$'. Set $\nu = [n/m]$ and $\delta = n/m - \nu \in [0,1)$. Write $p = \frac{1}{m}\sqrt{\nu/(\nu+\delta)}$ and $\tilde{p} = \frac{1}{m}\sqrt{(\nu+1)/(\nu+\delta)}$. Note that it is elementary to show that

$$m(1-\delta)p + m\delta\tilde{p} \leq 1.$$

Because of the $(m-1)$-dependence of the $X_{n,i}$, it is not hard to see that

$$\Delta_n = p\sum_{j=1}^{m(1-\delta)} \sqrt{\frac{m\nu}{n}}\Delta_{j,\nu} + \tilde{p}\sum_{j=1}^{m\delta} \sqrt{\frac{m(\nu+1)}{n}}\Delta_{j,\nu+1},$$

where the $\Delta_{j,\nu}$ and $\Delta_{j,\nu+1}$ are centered empirical distribution functions based on i.i.d. samples from F_n, with sample sizes ν and $\nu+1$ respectively (cf., e.g., (2.8) in Chanda and Ruymgaart [8]). Set

$$T_j = \sqrt{\frac{m\nu}{n}} \sup_{I \subset I_0} \Delta_{j,\nu}\{I\} \quad \text{and} \quad \tilde{T}_j = \sqrt{\frac{m(\nu+1)}{n}} \sup_{I \subset I_0} \Delta_{j,\nu+1}\{I\}.$$

Now we have by the Markov inequality and the Jensen inequality, for $t > 0$,

$$\begin{aligned}
(4.7)\quad \mathbf{P}\Big\{ \sup_{I \subset I_0} \Delta_n\{I\} > \lambda \Big\} &\leq \mathbf{P}\Big\{ p\sum_{j=1}^{m(1-\delta)} T_j + \tilde{p}\sum_{j=1}^{m\delta} \tilde{T}_j > \lambda \Big\} \\
&\leq e^{-t\lambda}\mathbf{E}\exp\Big(t\Big(p\sum_{j=1}^{m(1-\delta)} T_j + \tilde{p}\sum_{j=1}^{m\delta}\tilde{T}_j \Big)\Big) \\
&\leq e^{-t\lambda}\Big(p\sum_{j=1}^{m(1-\delta)} \mathbf{E}\exp(tT_j) + \tilde{p}\sum_{j=1}^{m\delta} \mathbf{E}\exp(t\tilde{T}_j) \Big) \\
&\leq e^{-t\lambda}\big(m(1-\delta)p\mathbf{E}\exp(tT_1) + m\delta\tilde{p}\mathbf{E}\exp(t\tilde{T}_1) \big).
\end{aligned}$$

We have used here the inequality

$$\begin{aligned}
\mathbf{E}\exp(tT_1) \geq \exp(t\mathbf{E}T_1) &= \exp\Big(t\sqrt{\frac{m\nu}{n}}\mathbf{E}\sup_{I \subset I_0} \Delta_{1,\nu}\{I\} \Big) \\
&\geq \exp\Big(t\sqrt{\frac{m\nu}{n}}\mathbf{E}\Delta_{1,\nu}\{I_0\} \Big) = 1.
\end{aligned}$$

Hence the last expression in (4.7) is bounded from above by

$$\max\Big(\mathbf{E}\exp\big(t(T_1 - \lambda)\big),\ \mathbf{E}\exp\big(t(\tilde{T}_1 - \lambda)\big) \Big).$$

So it suffices to bound the two terms in the maximum by the right side of the inequality. We will confine ourselves to the first term; the second one can be treated in the same way. Writing $s = t\sqrt{m/n}$ and $Y = T_1\sqrt{n/m}$ we have

$$(4.8)\qquad \mathbf{E}\exp\big(t(T_1 - \lambda)\big) = \mathbf{E}\exp\Big(s\sqrt{\frac{n}{m}}\Big(Y\sqrt{\frac{m}{n}} - \lambda \Big)\Big) = e^{-s\lambda\sqrt{n/m}}\mathbf{E}e^{sY}.$$

Note that $Y = \sup_{I \subset I_0} \nu \Delta_{\nu,1}\{I\}$. It is shown in Einmahl [10], Chapter 2, that

$$\mathbf{P}\{Y > \lambda\} \leq 8\mathbf{P}\Big\{Z + \sqrt{8F_n\{I_0\}} > \lambda\Big\}, \qquad \lambda \in \mathbb{R},$$

where $Z = \big(\sqrt{\tfrac{1}{\nu}}(V - \nu F_n\{I_0\})\big)$, with V a Poisson $(\nu F_n\{I_0\})$ random variable. So $\mathbf{P}\{e^{sY} > e^{s\lambda}\} \leq 8\mathbf{P}\{e^{s(Z+\sqrt{8F_n\{I_0\}})} > e^{s\lambda}\}$, and hence

$$\text{(4.9)} \qquad \begin{aligned} \mathbf{E}e^{sY} &= \int_0^\infty \mathbf{P}\{e^{sY} > x\}\,dx \leq 8\int_0^\infty \mathbf{P}\{e^{s(Z+\sqrt{8F_n\{I_0\}})} > x\}\,dx \\ &= 8\mathbf{E}\exp\{s(Z + \sqrt{8F_n\{I_0\}})\}. \end{aligned}$$

So from (4.8) and (4.9) we see that we have to bound

$$\text{(4.10)} \qquad 8\exp\Big\{-s\Big(\lambda\sqrt{\frac{n}{m}} - \sqrt{8F_n\{I_0\}}\Big)\Big\}\mathbf{E}e^{sZ}.$$

Since this holds true for every $s > 0$, the best result for (4.7) is obtained by minimizing (4.10) over s. Exploiting a well-known result for the moment generating function for Poisson random variables we see that minimization yields

$$\begin{aligned} &8\exp\Big(-\frac{(\lambda\sqrt{\frac{n}{m}} - \sqrt{8F_n\{I_0\}})^2}{2F_n\{I_0\}}\psi\Big(\frac{\lambda\sqrt{\frac{n}{m}} - \sqrt{8F_n\{I_0\}}}{\sqrt{\nu}F_n\{I_0\}}\Big)\Big) \\ &\qquad \leq 8\exp\Big(-\frac{(\lambda\sqrt{\frac{n}{m}} - \sqrt{8F_n\{I_0\}})^2}{2F_n\{I_0\}}\psi\Big(\frac{\lambda\sqrt{n}}{\sqrt{m\nu}F_n\{I_0\}}\Big)\Big), \end{aligned}$$

since $\psi \downarrow$. For $\lambda \geq 2\sqrt{8\frac{m}{n}F_n\{I_0\}}/\varepsilon$, this expression is bounded by

$$\text{(4.11)} \qquad 8\exp\Big(-\frac{(1-\varepsilon)n\lambda^2}{2mF_n\{I_0\}}\psi\Big(\frac{\lambda\sqrt{n}}{\sqrt{m\lfloor n/m\rfloor}F_n\{I_0\}}\Big)\Big).$$

Now consider $0 \leq \lambda < 2\sqrt{8\frac{m}{n}F_n\{I_0\}}/\varepsilon$. Then there exists $C(\varepsilon) \in (0,\infty)$ such that

$$\text{(4.12)} \qquad \mathbf{P}\{\sup_{I \subset I_0} \Delta_n\{I\} > \lambda\} \leq 1 \leq C(\varepsilon)\exp\Big(\frac{-n\lambda^2}{2mF_n\{I_0\}}\Big),$$

and (4.4) follows. □

Example 4.1. We observed already that (4.1) is immediate for model (1.1). It should be noted that in this case for fixed $a_0 < b_0$ inequality (4.4) holds true for sufficiently large n with $F_n(b_0) - F_n(a_0)$ replaced with $\Phi(b_0) - \Phi(a_0)$ and $(1-\varepsilon)$ with $(1-\varepsilon)^2$, due to the central limit theorem (Φ is the standard normal c.d.f.).

5. Weak Convergence of the Empirical Process

For technical reasons that will become clear in the proof, we will restrict ourselves to the situation of Corollary 3.1, in particular (3.20), in order to prove weak convergence in a space of corlol functions, cf. Philipp and Stout [15]. More precisely we need the following

Assumption 5.1. For convenience we assume that all the $X_{n,i}$ take values in $[0,1]$. There exists a c.d.f. F such that in addition to (4.1) we have

$$\sup_{0\le t\le 1} |F_n(t) - F(t)| \to 0 \qquad \text{as} \quad n \to \infty. \tag{5.1}$$

Furthermore there exists a function $H\colon [0,1]\times[0,1] \to [0,1]$ such that

$$\frac{1}{n(2m-1)} \sum \sum_{|i-j|<m} \mathbf{P}\{X_{n,i} \le s, X_{n,j} \le t\} \to H(s,t) \qquad \text{as} \quad n \to \infty \tag{5.2}$$

for $(s,t) \in [0,1]\times[0,1]$ and $m = m(n)$ satisfying (2.2).

Theorem 5.1. *There exists a centered Gaussian process $\mathcal{G}$ with covariance function*

$$\Gamma(s,t) := \mathbf{E}\mathcal{G}(s)\mathcal{G}(t) = H(s,t) - F(s)F(t), \qquad (s,t) \in [0,1]\times[0,1], \tag{5.3}$$

such that

$$\sqrt{\frac{n}{2m-1}}\,\Delta_n \to_d \mathcal{G} \qquad \textit{as} \quad n \to \infty. \tag{5.4}$$

The convergence is in the space $D([0,1])$ endowed with the Skorokhod $\mathcal{J}_1$-topology. If F is continuous, $\mathcal{G}$ has continuous sample paths with probability one.

Proof. To establish (5.4) it suffices to prove suitable weak convergence of the finite-dimensional distributions (fidi's) and tightness (Billingsley [5], Theorem 15.1).

Let us start with the fidi's and choose $0 \le t_1 < \cdots < t_k \le 1$. We need to prove that

$$\sqrt{\frac{n}{2m-1}}(\Delta_n(t_1),\ldots,\Delta_n(t_k)) \to_d (\mathcal{G}(t_1),\ldots,\mathcal{G}(t_k)) \qquad \text{as} \quad n \to \infty. \tag{5.5}$$

According to the Cramér–Wold device it suffices the prove that

$$\sqrt{\frac{n}{2m-1}} \sum_{v=1}^{k} \lambda_v \Delta_n(t_v) \to_d \sum_{v=1}^{k} \lambda_v \mathcal{G}(t_v). \tag{5.6}$$

For this purpose we will apply Corollary 3.1 where for the $X_{n,i}$ we now take

$$\widetilde{X}_{n,i} := \sum_{v=1}^{k} \lambda_v \xi_{n,i}(t_v), \qquad \xi_{n,i}(t_v) := \mathbf{1}_{(-\infty,t_v]}(X_{n,i}) - F_n(t_v). \tag{5.7}$$

Since the $\widetilde{X}_{n,i}$ are bounded and centered, condition (3.4) is automatically fulfilled. To verify condition (3.20), note that

$$(5.8)\quad \frac{1}{n(2m-1)}\mathbf{Var}\,(\widetilde{X}_{n,1}+\cdots+\widetilde{X}_{n,n}) = \frac{1}{n(2m-1)}\mathbf{E}\Big\{\sum_{i=1}^{n}\sum_{v=1}^{k}\lambda_v\xi_{n,i}(t_v)\Big\}^2$$

$$= \frac{1}{n(2m-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\mathbf{E}\xi_{n,i}(t_v)\xi_{n,j}(t_w)$$

$$= \frac{1}{n(2m-1)}\sum\sum_{|i-j|<m}\sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\mathbf{E}\xi_{n,i}(t_v)\xi_{n,j}(t_w)$$

$$\to \sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\{\mathbf{H}(t_v,t_w)-F(t_v)F(t_w)\}\qquad \text{as}\quad n\to\infty.$$

This limit is obviously nonnegative and condition (3.20) follows. So Corollary 3.1 yields the asymptotic normality with limiting variance equal to the quantity at the end of (5.8). This also settles (5.5).

For the tightness in $D([0,1])$ we invoke Billingsley [5], Theorem 15.2. For convenience let us write

$$(5.9)\qquad \mathcal{G}_n(t) := \sqrt{\frac{n}{2m-1}}\Delta_n(t),\qquad t\in[0,1].$$

The first condition that we need to verify is that for each $\eta > 0$ there exists a number $a > 0$ such that

$$(5.10)\qquad \mathbf{P}\big\{\sup_t |\mathcal{G}_n(t)| > a\big\} \le \eta \qquad \text{for all}\quad n\in\mathbb{N}.$$

This follows easily from Theorem 4.1.

Secondly we need to prove that for each $\varepsilon > 0$ and $\eta > 0$ there exist a $\delta \in (0,1)$ and an $n_0 \in \mathbb{N}$ such that

$$(5.11)\qquad \mathbf{P}\big\{w'(\mathcal{G}_n;\delta)\ge\varepsilon\big\} \le \eta \qquad \text{for all}\quad n\ge n_0,$$

where

$$(5.12)\qquad w'(\mathcal{G}_n;\delta) := \inf_{\substack{\text{all finite sets}\\ 0=t_0<t_1<\cdots<t_k=1\\ \text{with } dt_j-t_{j-1}>\delta}}\ \max_{1\le j\le k} w_j(\mathcal{G}_n)$$

and

$$(5.13)\qquad w_j(\mathcal{G}_n) = \sup\big\{|\mathcal{G}_n(s)-\mathcal{G}_n(t)| : s,t\in[t_{j-1},t_j)\big\}.$$

Now choose the t_j in such a way that

$$(5.14)\qquad F(t_j-)-F(t_{j-1}) \le 1/k.$$

Application of Theorem 4.1 yields for a sufficiently large k that

$$\begin{aligned}(5.15)\qquad \mathbf{P}\{\max_j w_j(\mathcal{G}_n)\geq\varepsilon\} &\leq \sum_{j=1}^{k}\mathbf{P}\{w_j(\mathcal{G}_n)\geq\varepsilon\}\\ &\leq C\Big(\frac{1}{2}\Big)k\exp\bigg(\frac{-\varepsilon^2(2m-1)}{2\cdot 2m\cdot(2/k)}\psi\Big(\frac{\varepsilon\sqrt{2m-1}}{\sqrt{m\lfloor n/m\rfloor}(2/k)}\Big)\bigg)\leq\eta,\end{aligned}$$

for all n sufficiently large. For (5.15) we use (5.1) and (5.14) which ensure that $F_n(t_j-)-F_n(t_{j-1})\leq 2/k$, for n sufficiently large, and the fact that ψ satisfies $\psi(x)\uparrow 1$, as $x\downarrow 0$. This proves the tightness and hence (5.4). If F is continuous we can similarly prove (5.11) with $w'(\mathcal{G}_n;\delta)$ replaced by the ordinary modulus of continuity. This proves, according to Billingsley [5], Theorem 15.5, the last statement of Theorem 5.1. □

Remark 5.1. In order for the local fluctuation inequality (4.4) to work in the tightness proof, the factor m in the denominator of the first part of the exponential expression in (5.15) should be neutralized. This means that a scaling of the Δ_n of order $\sqrt{n/m}$ will be needed, which in turn dictates the order of the covariances as required in (5.2)–(5.3). For the tightness it is of crucial importance that in the present setup we have been able to avoid the occurrence of a factor m in front of the exponential in (4.4). Such a factor naturally arises if in (4.7) one uses a naive upper bound of the type $\mathbf{P}\{\sum_{j=1}^m Y_j>\sum_{j=1}^m\eta_j\}\leq\sum_{j=1}^m\mathbf{P}\{Y_j>\eta_j\}$, as in Chanda and Ruymgaart [9].

Example 5.1. Let us consider the array $\widetilde{X}_{n,i}=\Phi(X_{n,i})$ with the $X_{n,i}$ as in (1.1). We will show that Theorem 5.1 applies to this array. Then by the ordinary central limit theorem (5.1) is fulfilled with $F_n(t)=\overline{F}_n(\Phi^{-1}(t))$, where $\overline{F}_n$ is the c.d.f. of the $X_{n,i}$, and $F(t)=t$, $0\leq t\leq 1$. Let Δ_n be based on the $\widetilde{X}_{n,i}$. It remains to show (5.2). With a little effort it can indeed be shown that, as $n\to\infty$,

$$\begin{aligned}(5.16)\qquad &\frac{1}{n(2m-1)}\sum\sum_{|i-j|<m}\mathbf{P}\{\widetilde{X}_{n,i}\leq s,\widetilde{X}_{n,j}\leq t\}\\ &\to\int_0^1\int_{-\infty}^{\infty}\Phi\Big(\frac{\Phi^{-1}(s)-z}{\sqrt{1-\rho}}\Big)\Phi\Big(\frac{\Phi^{-1}(t)-z}{\sqrt{1-\rho}}\Big)\,d\Phi\Big(\frac{z}{\sqrt{\rho}}\Big)\,d\rho=H(s,t).\end{aligned}$$

It is also worthwhile to note that $H(s,t)-st>0$, for all (s,t), which implies that the limiting process $\mathcal{G}$ in (5.4) is not degenerate or, in other words, that here indeed the excess of the variance matches the order of the dependence.

6. Application to Linear Processes

6.1. Generalities. Returning to model (1.3) let us assume that the coefficients satisfy

$$(6.1)\qquad C_1\frac{1}{1+|k|^\delta}\leq|a_k|\leq C_2\frac{1}{1+|k|^\delta},\qquad \delta>\frac{1}{2},$$

for some numbers $0<C_1<C_2<\infty$. We will not be concerned with exponential decay that can be dealt with in a similar manner. Note that $\sum a_k^2<\infty$, so that the series on the right in (1.3) converges almost surely. Henceforth C_1, C_2, and C will be used as positive generic constants.

To determine for which δ the linear process is long range dependent we will specify the covariance function. This follows from $\gamma_n(h) = \mathbf{E}X_{n,i} \cdot X_{n,i+h}$ (see (1.4)) which will be also needed to see which of the assumptions are verified by the array $X_{n,i}$. We have $\gamma_n(h) = 0$ for $|h| \geq 2m+1$ and, for $1 \leq h \leq 2m$,

$$(6.2)\qquad 0 \leq \gamma_n(h) \leq C \sum_{k=-m}^{m} \frac{1}{(1+|k|^\delta)(1+|k+h|^\delta)}$$
$$\leq C\left\{h^{-\delta} \sum_{k=\lfloor -h/2 \rfloor}^{\lceil h/3 \rceil} \frac{1}{1+|k|^\delta} + \sum_{\lceil h/3 \rceil+1}^{m} \frac{1}{(1+|k|^\delta)(1+|k+h|^\delta)}\right\}.$$

It follows that for $\frac{1}{2} < \delta < 1$

$$(6.3)\qquad 0 \leq \gamma_n(h) \leq C\left\{h^{1-2\delta} + \int_{\lceil h/3 \rceil}^{m} \frac{1}{1+x^\delta(x+h)^\delta}\, dx\right\}$$
$$\leq C\left\{h^{1-2\delta} + h^{1-2\delta}\int_{1/3}^{\infty} y^{-2\delta}\, dy\right\} \leq Ch^{1-2\delta}.$$

For $\delta > 1$ we have

$$(6.4)\qquad 0 \leq \gamma_n(h) \leq C\left\{h^{-\delta}\int_0^\infty \frac{1}{1+x^\delta}\, dx + h^{1-2\delta}\right\} \leq Ch^{-\delta}.$$

For the precise order we also need a lower bound:

$$(6.5)\qquad \gamma_n(h) \geq C \sum_{k=\lceil -h/2 \rceil}^{m} \frac{1}{(1+|k|^\delta)(1+|k+h|^\delta)}$$
$$\geq C\left\{h^{-\delta} \sum_{k=\lceil -h/2 \rceil}^{\lceil h/3 \rceil} \frac{1}{1+|k|^\delta} + \sum_{k=\lceil h/3 \rceil+1}^{m} \frac{1}{(1+|k|^\delta)(1+|k+h|^\delta)}\right\}.$$

For $\frac{1}{2} < \delta < 1$ this yields

$$(6.6)\qquad \gamma_n(h) \geq Ch^{-\delta}\int_0^{h/3} \frac{1}{1+x^\delta}\, dx \geq Ch^{1-2\delta},$$

and for $\delta > 1$

$$(6.7)\qquad \gamma_n(h) \geq Ch^{-\delta}\int_{1/6}^{h/3} x^{-\delta}\, dx \geq Ch^{-\delta}.$$

Summarizing, we have obtained that

$$(6.8)\qquad \begin{cases} C_1 h^{1-2\delta} \leq \gamma_n(h) \leq C_2 h^{1-2\delta}, & \frac{1}{2} < \delta < 1 \\ C_1 h^{-\delta} \leq \gamma_n(h) \leq C_2 h^{-\delta}, & \delta > 1, \end{cases}$$

for $|h| \le 2m$, $m \in \mathbb{N}$. This entails that, as $h \to \infty$,

$$\gamma(h) = \mathbf{E}X_i X_{i+h} \quad \text{is of order} \begin{cases} h^{1-2\delta}, & \frac{1}{2} < \delta < 1, \\ h^{-\delta}, & \delta > 1. \end{cases} \tag{6.9}$$

Since obviously

$$\sum_{h=-\infty}^{\infty} \gamma(h) = \infty \qquad \text{for} \quad \frac{1}{2} < \delta < 1, \tag{6.10}$$

the linear process is long range dependent (Beran [3]) for δ in that range. For $\delta > 1$ this is no longer the case.

6.2. APPROXIMATION. At sampling stage n we are given the sample $\mathbf{X}_n = (X_1, \dots, X_n)$ that we want to approximate with $\mathbf{X}_{n,n} = (X_{n,1}, \dots, X_{n,n})$. A good approximation would be one for which

$$\max_{1 \le i \le n} |X_i - X_{n,i}| = o_p(n^{-\varepsilon}) \qquad \text{as} \quad n \to \infty \tag{6.11}$$

for some $\varepsilon > 0$. This would entail that $\|\mathbf{X}_n - \mathbf{X}_{n,n}\| = o_p(n^{\frac{1}{2}-\varepsilon})$, where $\|\cdot\|$ is the Euclidean distance in $\mathbb{R}^n$. Let us try to achieve this for m of order n^ρ, for some $0 < \rho < 1$, so that (2.2) is fulfilled. Since $\mathbf{Var}\,(X_i - X_{n,i}) = O(\int_m^\infty (1 + x^\delta)^{-2} dx) = O(m^{1-2\delta})$ it follows from Chebyshev's inequality that (6.11) is satisfied for $2\varepsilon + 1 + \rho(1-2\delta) < 0$. Hence both (2.2) and (6.11) are satisfied for $(2\varepsilon+1)/(2\delta-1) < \rho < 1$, but this is only possible if $\delta > 1+\varepsilon$. Consequently in this way we only obtain useful approximations for certain linear processes that are not long range dependent.

6.3. ESTIMATION OF THE AUTOREGRESSION FUNCTION. The question of estimating an autoregression function like

$$\mathbf{E}\big(\psi(X_{i+d}) \mid X_i = x_1, \dots, X_{i+d-1} = x_d\big), \tag{6.12}$$

for some measurable function ψ is a nonparametric curve estimation problem. Estimation might be performed via approximation by the locally dependent array of the $X_{n,i}$, and by estimating the autoregression function of the array for a suitable choice of $m(n)$. Estimators may be derived from the compound empirical process, and fluctuation inequalities like (4.4) are very useful to obtain rates of a.s. convergence of such estimators (see Chanda and Ruymgaart [9]). Rates, however, will not be optimal, because the fluctuation inequality is based on a kind of "worst case" analysis of the situation, where precise dependence among the m components of the empirical process is not taken into account because this is mathematically very hard.

Acknowledgements. The authors are grateful to a referee for constructive criticism that led to an improvement of an earlier draft of this paper.

References

[1] D. W. K. Andrews, *Nonstrong mixing autoregressive processes*, J. Appl. Probab., 21 (1984), 930–934.

[2] A. D. Barbour, *Stein's method for diffusion approximation*, Probab. Theory Rel. Fields, 84 (1990), 297–322.

[3] J. Beran, *Statistics for Long-Memory Processes*, Chapman and Hall, New York, 1994.

[4] K. N. Berk, *A central limit theorem for m-dependent random variables with unbounded m*, Ann. Probab., 1 (1973), 352–354.

[4] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[4] D. Bosq, *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statist., Vol. 110, Springer, New York, 1996.

[4] R. C. Bradley, *On the growth of variances in a central limit theorem for strongly mixing sequences*, Bernoulli, 5 (1999), 67–80.

[8] K. C. Chanda and F. H. Ruymgaart, *General linear processes: a property of the empirical process applied to density and mode estimation*, J. Time Ser. Anal., 11 (1990), 185–199.

[9] K. C. Chanda and F. H. Ruymgaart, *Curve estimation for m_n-decomposable time series including bilinear processes*, J. Multivar. Anal., 38 (1991), 149–166.

[10] J. H. J. Einmahl, *Multivariate Empirical Processes*, CWI Tract, Vol. 32, Amsterdam, 1987.

[11] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), 13–30.

[12] G. Nieuwenhuis, *Central limit theorems for sequences with m(n)-dependent main part*, J. Statist. Plann. Inference, 32 (1992), 229–241.

[13] G. Nieuwenhuis and F. H. Ruymgaart, *Some stochastic inequalities and asymptotic normality of serial rank statistics in general linear processes*, J. Statist. Plann. Inference, 25 (1990), 53–79.

[14] M. Peligrad, *On the asymptotic normality of sequences of weakly dependent random variables*, J. Theor. Probab., 9 (1996), 703–715.

[15] W. Philipp and W. Stout, *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables*, Memoirs Amer. Math. Soc., Vol. 2, AMS, Providence, Rhode Island, 1975.

[16] S. Portnoy, *Asymptotic behavior of regression quantiles in non-stationary dependent cases*, J. Multivar. Anal., 38 (1991), 100–113.

[17] G. Reinert, *A weak law of large numbers for empirical measures via Stein's method*, Ann. Probab., 23 (1995), 334–354.

[18] Q.-M. Shao and H. Yu, *Weak convergence for weighted empirical processes of dependent sequences*, Ann. Probab., 24 (1996), 2098–2127.

[19] G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics*, Wiley, New York, 1986.

[20] L. Zetterqvist, *Asymptotic distribution of Mann's test for trend for m-dependent seasonal observations*, Scand. J. Statist., 15 (1988), 81–95.

[*Received December* 1997; *revised December* 2000]