Some results for empirical processes of locally dependent time series

Eindhoven University
of Technology

Department of Mathematics
and Computing Sciences

Memorandum COSOR 98-01

Some results for empirical processes of locally dependent time series

J.H.J. Einmahl and F. Ruymgaart

Eindhoven, January 1998
The Netherlands

SOME RESULTS FOR EMPIRICAL PROCESSES OF LOCALLY DEPENDENT TIME SERIES

John H.J. Einmahl*†
Eindhoven University of Technology

Frits H. Ruymgaart*
Texas Tech University

Abstract

In this paper we derive some fundamental properties of locally dependent time series 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 time series are of independent interest, but they may also serve as useful approximations to other stochastic processes. Some applications are briefly indicated.

*) Research partially supported by NATO Collaborative Research Grant CRG 960097.

†) Research partially supported by European Union HCM grant ERB CHRX-CT 940693.

AMS 1991 subject classifications: primary 62M10, 62G30, secondary 60F05.

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

1 Introduction

A time series with local dependence of order $m(n)$ is a triangular array $\{X_{n,i},\ i = 1, ..., n \in \mathbb{N}\}$ of real valued random variables such that the variables in the n-th row

$$X_{n,1}, \ldots, X_{n,n} \text{ are } \{m(n) - 1\}\text{-dependent}, \tag{1.1}$$

for some $m(n) \in \{1, \ldots, n\}$. Local dependence of order $m(n) = 1$ reduces to independence. Although for our considerations a necessary condition is that $m(n)/n \to 0$, as $n \to \infty$, the interesting case where $m(n) \to \infty$ at not too high a rate is included. Locally dependent time series are both of independent interest and useful as approximations to other stochastic processes. They are regularly studied in the literature: see for instance, Hoeffding (1963), Billingley (1968), Berk (1973), Zetterqvist (1988), Chanda and Ruymgaart (1990), and Nieuwenhuis (1992). Recently, Barbour (1990) and Reinert (1995, 1996) considered the empirical process of such a time series, including the case $m(n) \to \infty$. Portnoy (1991) exploited approximation by means of locally dependent processes studying regression quantiles in non-stationary dependent cases. It should be noted that local dependence is a flexible dependence concept that is different from the usual mixing assumptions. For mixing and associated sequences the empirical process has been recently considered in Shao & Yu (1996).

In this paper we will focus on weak convergence of the empirical process and, in passing, derive some tools that are of interest in their own right. In order to establish the weak convergence of the finite dimensional distributions we need a generalization of Berk's (1973) central limit theorem for locally dependent time series (Section 2). To verify tightness we will use a powerful exponential inequality for the local fluctuations of the empirical process. Naive blocking would produce such an inequality with a factor of order $m(n)$ before the exponential expression. Due to this factor the inequality would not be strong enough to derive tightness, and elimination of this factor is a nontrivial problem addressed in Section 3. The weak convergence of the empirical process is presented in Section 4. Finally, Section 5 is devoted to an example and some brief remarks about how our results could be used in the study of stochastic processes that allow a suitable approximation by locally dependent arrays. Typically such processes exhibit a strong short term and a relatively weak long term dependence.

Because of its intrinsic interest, in this paper we derive some fundamental tools for the triangular array itself. As we observed above, further properties of the array may be obtained from these tools and carried over to other processes by approximation. In this respect our approach is complementary to the one in, e.g., Chanda & Ruymgaart (1990) where an approximation is used to derive the tools directly for a certain class of time series. The technicalities inherent to this approximation, however, have an averse effect on the formulation of the main tool: the local fluctuation inequality of the empirical process of the time series. The corresponding inequality in their paper, moreover, suffers from the serious drawback of containing the factor $m(n)$ before the exponential expression which we discussed in the previous paragraph.

Throughout the sequel $\{X_{n,i}\}$ will always denote a locally dependent array of order $m(n)$, such that

$$\frac{m(n)}{n} \to 0, \text{ as } n \to \infty. \tag{1.2}$$

In each of the subsequent sections specific conditions will be added depending on the situation at hand.

2 A central limit theorem

In this section it will be required that

$$\mathrm{E}|X_{n,i}|^{2+\delta} \leq M, \quad \mathrm{E}X_{n,i} = 0, \quad i = 1, \ldots, n, \quad n \in \mathbb{N}, \tag{2.1}$$

for some δ, $M \in (0, \infty)$. We will, moreover, assume that for some $K \in (0, \infty)$ and sequence $\{\ell(n)\}$ satisfying $1 \leq \ell(n) \leq 2m(n)$ we have

$$\mathbf{Var}(X_{n,i+1} + \cdots + X_{n,j}) \leq K(j-i)\ell(n), \tag{2.2}$$

$$\mathbf{Var}(X_{n,1} + \cdots + X_{n,n})/\{n\ell(n)\} \to \sigma^2 \in [0, \infty), \text{ as } n \to \infty, \tag{2.3}$$

and, finally, that

$$\frac{\{m(n)\}^{2+2/\delta}}{n\{\ell(n)\}^{1+2/\delta}} \to 0, \text{ as } n \to \infty. \tag{2.4}$$

Theorem 2.1. *Assuming (2.1) - (2.4) we have*

$$\frac{1}{\sqrt{n\ell(n)}} \sum_{i=1}^{n} X_{n,i} \to \mathcal{N}(0, \sigma^2), \text{ as } n \to \infty. \tag{2.5}$$

Proof. Let us choose a sequence of integers $\{q(n)\}$ such that $q(n) > 2m(n)$, $m(n)/q(n) \to 0$, and (2.4) holds true with $m(n)$ replaced by $q(n)$. This can always be achieved. In the remainder part of this proof we will simply write m, ℓ, q rather than $m(n), \ell(n), q(n)$. Define integers ν and r by $n = q\nu + r$ and $0 \le r < q$.

Next let us introduce

$$U_j := \sum_{i=(j-1)q+1}^{jq-m+1} X_{n,i}, \quad V_j := \sum_{i=jq-m+2}^{jq} X_{n,i}, \tag{2.6}$$

for $j = 1, \ldots, \nu$, and

$$R := \sum_{i=\nu q+1}^{n} X_{n,i}. \tag{2.7}$$

Since $q > 2m$ the $U_1, \ldots, U_\nu$ are independent random variables and so are the $V_1, \ldots, V_\nu$. We see from (2.2) and (2.4) that

$$\begin{aligned} & \mathbf{Var}\frac{1}{\sqrt{n\ell}}\left(\sum_{i=1}^{n} X_{n,i} - \sum_{j=1}^{\nu} U_j\right) = \qquad (2.8) \\ = \; & \frac{1}{n\ell}\mathbf{Var}\left(\sum_{j=1}^{\nu} V_j + R\right) \le \\ \le \; & \frac{1}{n\ell} K\{\nu(m-1)\ell + r\ell\} \to 0, \text{ as } n \to \infty. \end{aligned}$$

Apparently this entails that

$$\mathbf{Var}\frac{1}{\sqrt{n\ell}}\sum_{j=1}^{\nu} U_j \to \sigma^2 \in [0, \infty), \text{ as } n \to \infty, \tag{2.9}$$

and that it suffices to show that the random variables in (2.9) have the limiting distribution in (2.5).

The desired asymptotic normality follows from the Lyapunov central limit theorem if we can show that

$$\lim_{n\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{j=1}^{\nu} \mathbf{E}|U_j|^{2+\delta} = 0, \tag{2.10}$$

where $B_n^2 := \mathbf{Var}\left(\sum_{j=1}^{\nu} U_j\right)$. By (2.1), (2.6), and Minkovski's inequality we have

$$\mathbf{E}|U_j|^{2+\delta} \le M(q+1-m)^{2+\delta}. \tag{2.11}$$

Furthermore (2.9) implies that, for large n, $B_n^2 \geq \frac{1}{2}\sigma^2 n\ell$. Hence, if $\sigma^2 > 0$, it follows that for some number $C \in (0, \infty)$ we have

$$\frac{1}{B_n^{2+\delta}} \sum_{j=1}^{\nu} \mathbf{E}|U_j|^{2+\delta} \leq \frac{\nu M(q+1-m)^{2+\delta}}{(\frac{1}{2}\sigma^2 n\ell)^{(2+\delta)/2}} \leq \tag{2.12}$$

$$\leq C \frac{q^{1+\delta}}{n^{\delta/2}\ell^{(2+\delta)/2}} = C\left(\frac{q^{2+2/\delta}}{n\ell^{1+2/\delta}}\right)^{\delta/2} \to 0, \text{ as } n \to \infty,$$

by (2.4) with m replaced by q. If $\sigma^2 = 0$ weak convergence to the degenerate normal law is immediate from Chebyshev's inequality.

Remark 2.1. For $\ell(n) = 1$ our result reduces to the Theorem in Berk (1973).

Remark 2.2. For $\ell(n)$ and $m(n)$ of the same order, condition (2.4) simplifies to $m(n)/n \to 0$, as $n \to \infty$, which is part of the very definition of local dependence. If $\sigma^2 > 0$ and ℓ and m of the same order it follows that the order of the variance matches the order of the dependence. Without this assumption tightness of the empirical process cannot be proved by our method if it is true at all. Therefore in Section 4 we will exclusively deal with this case so that condition (2.4) will not be needed there.

3 A local fluctuation inequality

In this section we assume that the $X_{n,i}$ are identically distributed in each row, i.e. that

$$\mathbf{P}\{X_{n,i} \leq t\} = F_n(t), \ t \in \mathbb{R}, \ i = 1, \ldots, n, \ n \in \mathbb{N}. \tag{3.1}$$

The result of this section will be for an arbitrary but fixed sample size n, so that there is no need to specify the asymptotic behavior of $m(n)$ or F_n. To be in keeping with the notation of Section 4, however, the index n will not be omitted. For the empirical c.d.f. we will employ the usual notation

$$\hat{F}_n(t) := \frac{1}{n} \sum_{i=1}^{n} 1_{(-\infty,t]}(X_{n,i}), \ t \in \mathbb{R}. \tag{3.2}$$

We write

$$\Delta_n := \hat{F}_n - F_n, \tag{3.3}$$

for the discrepancy between the empirical c.d.f. and the common c.d.f. of the random variables in the n-th row.

Theorem 3.1. *Let $a_0 < b_0$ with $F_n(b_0) - F_n(a_0) \leq \frac{1}{2}$. Then we have, for any $\varepsilon \in (0,1)$,*

$$\begin{aligned} &\mathbf{P}\left\{\sup_{a_0 \leq a < b \leq b_0} |\Delta_n(b) - \Delta_n(a)| \geq \lambda\right\} \leq \\ &\leq C(\varepsilon) \exp\left(\frac{-(1-\varepsilon)n\lambda^2}{2m(n)\{F_n(b_0) - F_n(a_0)\}}\right. \\ &\left.\times \psi\left(\frac{\sqrt{n}\lambda}{\sqrt{m(n)\left\lfloor \frac{n}{m(n)} \right\rfloor \{F_n(b_0) - F_n(a_0)\}}}\right)\right), \\ &\lambda \geq 0, \end{aligned} \tag{3.4}$$

where $\psi(x) = 2x^{-2}\{(1+x)\log(1+x) - x\}, x \in (0,\infty)$, and $\psi(0) = 1$.

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

$$\begin{aligned} &\mathbf{P}\{\sup_{I \subset I_0} |\Delta_n\{I\}| \geq \lambda\} \leq \\ \leq\ &\mathbf{P}\{\sup_{I \subset I_0} \Delta_n\{I\} \geq \lambda\} + \mathbf{P}\{\sup_{I \subset I_0} -\Delta_n\{I\} \geq \lambda\}. \end{aligned} \tag{3.5}$$

We will only consider the first term on the right in (3.5): the second one can be treated similarly. It is easy to see that it is sufficient to prove the inequality with "$\geq \lambda$" replaced by "$> \lambda$". Let us, moreover, first assume that

$$\frac{n}{m(n)} =: \nu \in \mathbb{N}. \tag{3.6}$$

At the end of the proof we will briefly indicate some modifications that will enable us to accommodate noninteger $n/m(n)$. Again, let us write m rather than $m(n)$ in the remainder part of the proof.

Writing $\Delta_{\nu,j} = \hat{F}_{\nu,j} - F_n$ it is obvious that

$$\Delta_n = \frac{1}{m} \sum_{j=1}^{m} \Delta_{\nu,j}, \tag{3.7}$$

where the empirical c.d.f. $\hat{F}_{\nu,j}$ is based on the ν random variables X_j, X_{j+m}, $\ldots, X_{j+(\nu-1)m}$ that are i.i.d. because of the $(m-1)$-dependence (see also Chanda & Ruymgaart (1990)). Let us set

$$T_j := \sup_{I \subset I_0} \Delta_{\nu,j}\{I\}. \tag{3.8}$$

Now we have by the Markov inequality and the Jensen inequality (c.f. Hoeffding (1963)), for $t > 0$,

$$\mathbf{P}\{\sup_{I\subset I_0}\Delta_n\{I\} > \lambda\} \le \mathbf{P}\left\{\frac{1}{m}\sum_{j=1}^m T_j > \lambda\right\} \le \tag{3.9}$$

$$\le e^{-t\lambda}\mathbf{E}\exp\left(\frac{t}{m}\sum_{j=1}^m T_j\right) \le e^{-t\lambda}\left(\frac{1}{m}\sum_{j=1}^m \mathbf{E}\exp(tT_j)\right) =$$

$$= e^{-t\lambda}\mathbf{E}\exp(tT_1) = \mathbf{E}\exp(t(T_1-\lambda)).$$

Writing $s = t/\sqrt{\nu}$ and $Y = \sqrt{\nu}T_1$ we have

$$\mathbf{E}\exp(t(T_1-\lambda)) = \mathbf{E}\exp(s\sqrt{\nu}(Y/\sqrt{\nu}-\lambda)) = \tag{3.10}$$

$$= e^{-s\lambda/\sqrt{\nu}}\mathbf{E}e^{sY}.$$

Recall that $Y = \sup_{I\subset I_0}\sqrt{\nu}\Delta_{\nu,1}\{I\} =: \sup_{I\subset I_0}U_{\nu,1}\{I\}$, where $U_{\nu,1}$ is the empirical process of $X_1, X_{1+m}, \ldots, X_{1+(\nu-1)m}$. It has been shown in Einmahl (1987, Chapter 2) that

$$\mathbf{P}\{Y > \lambda\} \le 8\,\mathbf{P}\left\{Z + \sqrt{8\,F_n\{I_0\}} > \lambda\right\},\ \lambda\in\mathbb{R}, \tag{3.11}$$

where $Z = (V - \nu F_n\{I_0\})/\sqrt{\nu}$ with V a Poisson $(\nu F_n\{I_0\})$ random variable. This entails that

$$\mathbf{P}\{e^{sY} > e^{s\lambda}\} \le 8\,\mathbf{P}\left\{e^{s(Z+\sqrt{8F_n\{I_0\}})} > e^{s\lambda}\right\}, \tag{3.12}$$

and hence

$$\mathbf{E}e^{sY} = \int_0^\infty \mathbf{P}\{e^{sY} > x\}dx \le \tag{3.13}$$

$$\le 8\int_0^\infty \mathbf{P}\left\{e^{s(Z+\sqrt{8F_n\{I_0\}})} > x\right\}dx = 8\,\mathbf{E}e^{s(Z+\sqrt{8\,F_n\{I_0\}})}.$$

Combining (3.10) and (3.13) we see that we have to bound

$$8e^{-s(\lambda\sqrt{\nu}-\sqrt{8F_n\{I_0\}})}\mathbf{E}e^{sZ}. \tag{3.14}$$

Since this holds true for every $s > 0$, the best result for (3.9) is obtained by minimizing (3.14) over s. Exploiting a well-known result for the moment generating function for Poisson variables we see that minimization yields

$$8\,\exp\left(-\frac{\left(\lambda\sqrt{\nu}-\sqrt{8\,F_n\{I_0\}}\right)^2}{2F_n\{I_0\}}\psi\left(\frac{\lambda\sqrt{\nu}-\sqrt{8\,F_n\{I_0\}}}{\sqrt{\nu}F_n\{I_0\}}\right)\right). \tag{3.15}$$

For $\lambda \geq 2\sqrt{8\ F_n\{I_0\}/\nu}/\varepsilon$ this expression is bounded by

$$8\exp\left(-\frac{\left(\lambda\sqrt{\nu}-\sqrt{8\ F_n\{I_0\}}\right)^2}{2\ F_n\{I_0\}}\,\psi\left(\frac{\lambda}{F_n\{I_0\}}\right)\right) \leq$$

$$\leq\ 8\exp\left(-\frac{(1-\varepsilon)\nu\lambda^2}{2\ F_n\{I_0\}}\psi\left(\frac{\lambda}{F_n\{I_0\}}\right)\right) =$$

$$=\ 8\exp\left(-\frac{(1-\varepsilon)n\lambda^2}{2mF_n\{I_0\}}\psi\left(\frac{\sqrt{n}\lambda}{\sqrt{m\frac{n}{m}F_n\{I_0\}}}\right)\right).$$

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

$$\mathbf{P}\{\sup_{I\subset I_0}\Delta_n\{I\} > \lambda\} \leq 1 \leq C'(\varepsilon)\exp\left(\frac{-n\lambda^2}{2mF_n\{I_0\}}\right), \tag{3.16}$$

and (3.4) follows with $C(\varepsilon) = 8 \vee C'(\varepsilon)$.

If n/m is not an integer set $\nu := \lfloor n/m \rfloor$ and $\delta := n/m - \nu \in (0,1)$. Then we have

$$\Delta_n = \frac{\nu+1}{n}\sum\nolimits_{j=1}^{\delta m}\Delta_{\nu+1,j} + \frac{\nu}{n}\sum\nolimits_{j=1}^{(1-\delta)m}\Delta_{\nu,j}, \tag{3.17}$$

where the $\Delta_{\nu+1,j}$ are based on i.i.d. samples of size $\nu+1$ and the $\Delta_{\nu,j}$ on samples of size ν. Now let

$$T_j := \sup_{I\subset I_0}\Delta_{\nu,j}\{I\},\ j = 1,\ldots,(1-\delta)m, \tag{3.18}$$

$$\tilde{T}_j := \sup_{I\subset I_0}\Delta_{\nu+1,j}\{I\},\ j = 1,\ldots,\delta m, \tag{3.19}$$

and application of the Markov inequality yields

$$\mathbf{P}\{\sup_{I\subset I_0}\Delta_n\{I\} > \lambda\} \leq \tag{3.20}$$

$$\leq\ \mathbf{P}\left\{\frac{\nu+1}{n}\sum\nolimits_{j=1}^{\delta m}\tilde{T}_j + \frac{\nu}{n}\sum\nolimits_{j=1}^{(1-\delta)m}T_j > \lambda\right\} \leq$$

$$\leq\ e^{-t\lambda}\mathbf{E}\exp\left(t\left(\frac{\nu+1}{n}\sum\nolimits_{j=1}^{\delta m}\tilde{T}_j + \frac{\nu}{n}\sum\nolimits_{j=1}^{(1-\delta)m}T_j\right)\right).$$

We next apply Jensen's inequality and arrive at the same conclusion with only minor modifications.

Remark 3.1. Following Einmahl (1987) it is immediate that inequality (3.4) extends to multivariate random vectors with intervals replaced by rectangles.

4 Weak convergence of the empirical process

It will be convenient, but not necessary, to assume that the $X_{n,i}$ take values in $[0,1]$. In addition to (3.1) we require that there exists a c.d.f. F such that

$$\sup_{0\le t\le 1}|F_n(t)-F(t)|\to 0, \text{as } n\to\infty. \tag{4.1}$$

Furthermore we assume that

$$\frac{1}{n(2m(n)-1)}\sum\sum_{|i-j|<m(n)}\mathbf{P}\{X_{n,i}\le s, X_{n,j}\le t\}\to H(s,t), \tag{4.2}$$
as $n\to\infty, (s,t)\in[0,1]\times[0,1]$.

Theorem 4.1. *There exists a centered Gaussian process $\mathcal{G}$ with covariance function*

$$\mathbf{E}\mathcal{G}(s)\mathcal{G}(t)=H(s,t)-F(s)F(t),\ (s,t)\in[0,1]\times[0,1], \tag{4.3}$$

such that

$$\sqrt{\frac{n}{2m(n)-1}}\Delta_n\Rightarrow\mathcal{G}, \text{as } n\to\infty. \tag{4.4}$$

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 (4.4) it suffices to prove suitable weak convergence of the finite dimensional distributions (fidi's) and tightness (Billingsley (1968, Theorem 15.1)).

Let us start with the fidi's and choose $0\le t_1<\cdots<t_k\le 1$. Let us write again ℓ and m instead of $\ell(n)$, and $m(n)$. We need to prove that

$$\sqrt{\frac{n}{2m-1}}(\Delta_n(t_1),\ldots,\Delta_n(t_k))\Rightarrow(\mathcal{G}(t_1),\ldots,\mathcal{G}(t_k)), \text{ as } n\to\infty. \tag{4.5}$$

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

$$\sqrt{\frac{n}{2m-1}}\sum_{v=1}^{k}\lambda_v\Delta_n(t_v)\Rightarrow\sum_{v=1}^{k}\lambda_v\mathcal{G}(t_v). \tag{4.6}$$

For this purpose we will apply Theorem 2.1 where for the $X_{n,i}$ we now take

$$\tilde{X}_{n,i}:=\sum_{v=1}^{k}\lambda_v\xi_{n,i}(t_v), \xi_{n,i}(t_v):=1_{(-\infty,t_v]}(X_{n,i})-F_n(t_v). \tag{4.7}$$

Since the $\tilde{X}_{n,i}$ are bounded and centered, condition (2.1) is automatically fulfilled. Taking $\ell = 2m-1$, condition (2.2) is trivially satisfied. To verify condition (2.3), note that

$$\begin{aligned}
&\frac{1}{n(2m-1)}\mathbf{Var}(\tilde{X}_{n,1}+\cdots+\tilde{X}_{n,n}) = \qquad (4.8)\\
&= \frac{1}{n(2m-1)}\mathbf{E}\left\{\sum_{i=1}^{n}\sum_{v=1}^{k}\lambda_v\xi_{n,i}(t_v)\right\}^2 =\\
&= \frac{1}{n(2m-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\mathbf{E}\xi_{n,i}(t_v)\xi_{n,j}(t_w) =\\
&= \frac{1}{n(2m-1)}\sum\sum_{|i-j|<m}\sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\mathbf{E}\xi_{n,i}(t_v)\xi_{n,j}(t_w) \to\\
&\to \sum_{v=1}^{k}\sum_{w=1}^{k}\lambda_v\lambda_w\{H(t_v,t_w)-F(t_v)F(t_w)\}, \text{ as } n\to\infty.
\end{aligned}$$

This limit is obviously nonnegative and condition (2.3) with $\ell = 2m-1$ follows. Because we are dealing with the case $\ell = 2m-1$, condition (2.4) is redundant and Theorem 2.1 yields the asymptotic normality with limiting variance equal to the number at the end of (4.8). This also settles (4.5).

For the tightness in $D([0,1])$ we invoke Billingsley (1968, Theorem 15.2). For convenience let us write

$$\mathcal{G}_n(t) := \sqrt{\frac{n}{2m-1}}\Delta_n(t), t\in[0,1]. \qquad (4.9)$$

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

$$\mathbf{P}\{\sup_t|\mathcal{G}_n(t)| > a\} \le \eta, \text{ for all } n\in\mathbb{N}. \qquad (4.10)$$

This follows easily from Theorem 3.1.

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

$$\mathbf{P}\{w'(\mathcal{G}_n;\delta)\ge\varepsilon\} \le \eta, \text{ for all } n\ge n_0, \qquad (4.11)$$

where

$$w'(\mathcal{G}_n;\delta) := \inf_{\substack{\text{all finite sets}\\ 0=t_0<t_1<\cdots<t_k=1\\ \text{with } t_j-t_{j-1}>\delta}} \max_{1\le j\le k} w_j(\mathcal{G}_n), \qquad (4.12)$$

and

$$w_j(\mathcal{G}_n) = \sup\{|\mathcal{G}_n(s) - \mathcal{G}_n(t)| : s,t \in [t_{j-1}, t_j)\}. \tag{4.13}$$

Now choose the t_j in such a way that

$$F(t_j-) - F(t_{j-1}) \leq \frac{1}{k}. \tag{4.14}$$

Application of Theorem 3.1 yields

$$\begin{aligned} & \mathbf{P}\{\max_j w_j(\mathcal{G}_n) \geq \varepsilon\} \leq \sum_{j=1}^{k} \mathbf{P}\{w_j(\mathcal{G}_n) \geq \varepsilon\} \leq \qquad (4.15) \\ \geq\ & C(\frac{1}{2})k \exp\left(\frac{-\varepsilon^2(2m-1)}{2\cdot 2m\cdot(2/k)}\psi\left(\frac{\varepsilon\sqrt{2m-1}}{\sqrt{m\lfloor n/m\rfloor}(2/k)}\right)\right) \leq \\ \leq\ & \eta, \text{ for all } n \text{ sufficiently large,} \end{aligned}$$

for a sufficiently large k. For (4.15) we use (4.1) and (4.14) 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 (4.4). If F is continuous we can similarly prove (4.11) with $w'(\mathcal{G}_n;\delta)$ replaced by the ordinary modulus of continuity. This proves, according to Billingsley (1968, Theorem 15.5), the last statement of Theorem 4.1.

Remark 4.1. Under natural additional conditions it is possible to refine Theorem 4.1 to the case where the empirical process is weighted by a q-function (Chibisov-O'Reilly-type theorems). When dealing with mean residual life and related processes, it is important to know the behaviour of the integral functionals $\int_{-\infty}^{x} \sqrt{\frac{n}{2m(n)-1}}\Delta_n(y)dy$ of the empirical process. These integral functionals can be easily treated with the just mentioned refinements of Theorem 4.1 to the weighted empirical process. See, e.g., Shao and Yu (1996) for this type of applications under a different dependence structure.

5 Some applications

A. Application to locally dependent arrays. The theory of empirical processes derived above finds its usual application to the problem of finding the limiting distribution of linear rank statistics and linear combinations of order statistics of the array elements themselves. Another application is indicated in Remark

4.1. These applications are well-known and will not be considered here. We rather want to consider an example where the local dependence structure is explicitly constructed.

Suppose that $Y_1, Y_2, \ldots$ is an infinite sequence of i.i.d. random variables with mean 0 and variance 1. At stage n let us define

$$X_{n,i} := \frac{1}{\sqrt{m(n)}} \sum_{j=1}^{m(n)} Y_{i+j-1}. \tag{5.1}$$

This obviously is a locally dependent array of order $m(n)$. We will assume as before that $m(n)/n \to 0$ and we will also restrict our attention to the most interesting case where $m(n) \to \infty$. (For $m(n)$ equal to a fixed m, the $X_{n,i}$ are already thoroughly studied in the literature, see, e.g., Billingsley (1968, p. 167).) The $X_{n,i}$ are identically distributed with c.d.f. F_n, say, where

$$\sup_{t \in \mathbb{R}} |F_n(t) - \Phi(t)| \to 0, \text{ as } n \to \infty, \tag{5.2}$$

with Φ the standard normal c.d.f., by the central limit theorem.

We will show that Theorem 4.1 applies to this array. In order to establish this, it remains to show (4.2), which indeed appears to hold true. Computing the left hand side of (4.2) and passing to the limit is somewhat cumbersome, however, but elementary. We therefore omit it and only specify the limit function:

$$H(s,t) = \int_0^1 \int_{-\infty}^{\infty} \Phi\left(\frac{s-y}{\sqrt{1-\rho}}\right) \Phi\left(\frac{t-y}{\sqrt{1-\rho}}\right) d\Phi\left(\frac{y}{\sqrt{\rho}}\right) d\rho, \tag{5.3}$$

$(s,t) \in \mathbb{R} \times \mathbb{R}$. (Here and in (4.1) we tacitly replaced $[0,1]$ by $\mathbb{R}$.) It is also worthwhile to note that $H(s,t) - \Phi(s)\Phi(t) > 0$, for all (s,t), which implies that the limiting process $\mathcal{G}$ in (4.4) is not degenerate or, in other words, that here indeed the order of the variance matches the order of the local dependence.

B. Approximating time series. Let $\tilde{X}_1, \tilde{X}_2, \ldots$ be a time series of centered random variables and $\{X_{n,i}\}$ a centered array of random variables that are locally dependent of order $m(n)$, where all random variables are defined on the same sample space. Suppose that at sampling stage n

$$\mathbf{Var}(X_{n,i} - \tilde{X}_i) \leq Cn^{-\delta}, \ \delta > 2, \ i = 1, \ldots, n, \tag{5.4}$$

for some number $C \in (0, \infty)$ that does not depend on n. For $a \in \mathbb{R}^n$ we write $\|a\|_\infty := \max_{1 \leq i \leq n} |a_i|$, and $\tilde{X}_{(n)} := (\tilde{X}_1, \ldots, \tilde{X}_n)$, $X_{(n)} := (X_{n,1}, \ldots, X_{n,n})$.

It easily follows that

$$\|\tilde{X}_{(n)} - X_{(n)}\|_\infty =_{a.s.} \mathcal{O}(n^{-\varepsilon}), \text{ for any } \varepsilon < (\delta - 2)/2. \tag{5.5}$$

For sufficiently large δ this allows us to carry over properties of the locally dependent array, that can be proved using the results obtained in the previous sections, to the time series.

An example of a class of time series where this approach might work is the class of linear processes.

$$\tilde{X}_i := \sum\nolimits_{k\in\mathbb{Z}} a_k \varepsilon_{i-k}, \tag{5.6}$$

where the ε_k are i.i.d. error variables with zero mean and finite variance, and the numbers a_k decrease at a suitable rate. The approximating $X_{n,i}$ are given by

$$X_{n,i} := \sum\nolimits_{|k|\le\nu(n)} a_k \varepsilon_{i-k}, \tag{5.7}$$

This array is locally dependent of order $m(n) := 2\nu(n) + 1$. Some strong convergence results in Chanda & Ruymgaart (1996) for time series that allow this kind of decomposition could also be obtained by deriving them first for the triangular array and then using (5.5). Obtaining weak convergence of the empirical process of the time series in (5.6) is possible along these lines in case the coefficients a_k decay sufficiently fast to ensure 'short range' dependence, but seems impossible to obtain for the 'long range' dependence case, since then $m(n)/n \to 0$ can not be satisfied.

C. Sampling from a production process. Production processes are processes where strong dependence may occur at clusters of neighboring products, and where the dependence is weak or absent when products are far away in time. Samples from such processes may naturally lead to locally dependent arrays when during each period of time a number of neighboring products is checked. Here it seems reasonable to increase the number of neighboring products, with the total sample size.

References

[1] A.D. Barbour (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84**, 297-322.

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

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

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

[5] J.H.J. Einmahl (1987). *Multivariate Empirical Processes.* CWI Tract **32**, Amsterdam.

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

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

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

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

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

[11] G. Reinert (1996). Gaussian approximations for random measures via Stein's method. Preprint.

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