

Strong convergence rates for the estimation of a covariance operator for associated samples

Carla Henriques · Paulo Eduardo Oliveira

Received: 13 May 2005 / Accepted: 15 September 2006 / Published online: 3 November 2006
© Springer Science+Business Media B.V. 2006

Abstract Let $X_n, n \geq 1$, be a strictly stationary associated sequence of random variables, with common continuous distribution function F . Using histogram type estimators we consider the estimation of the two-dimensional distribution function of (X_1, X_{k+1}) as well as the estimation of the covariance function of the limit empirical process induced by the sequence $X_n, n \geq 1$. Assuming a convenient decrease rate of the covariances $\text{Cov}(X_1, X_{n+1}), n \geq 1$, we derive uniform strong convergence rates for these estimators. The condition on the covariance structure of the variables is satisfied either if $\text{Cov}(X_1, X_{n+1})$ decreases polynomially or if it decreases geometrically, but as we could expect, under the latter condition we are able to establish faster convergence rates. For the two-dimensional distribution function the rate of convergence derived under a geometrical decrease of the covariances is close to the optimal rate for independent samples.

Keywords Association · Empirical process · Histogram estimator · Stationarity

AMS Classification 62G20 · 62G05 · 62G30

1 Introduction, assumptions and definitions

A sequence of random variables $X_n, n \geq 1$, is said to be associated if for any $m \in \mathbb{N}$ and any two real-valued coordinatewise nondecreasing functions f and g it holds

$$\text{Cov}\left(f(X_1, \dots, X_m), g(X_1, \dots, X_m)\right) \geq 0,$$

C. Henriques
Department of Matemática, Escola Sup. Tecnologia de Viseu, Campus Politécnico, 3504-510
Viseu, Portugal
e-mail: cadaheriq@mat.estv.ipv.pt

P. E. Oliveira (✉)
Department of Matemática, University Coimbra, Apartado 3008, 3001-454 Coimbra, Portugal
e-mail: paulo@mat.uc.pt

whenever this covariance exists. This dependence concept, introduced in [Esary et al. \(1967\)](#), occurs naturally in every frame-work that depends on monotone transformations of random variables such as reliability theory, censored or truncated sampling (see, for example, [Barlow and Proschan \(1975\)](#)). The dependence notion appears also in statistical mechanics referred to as FKG inequalities (see [Fortuin et al. \(1971\)](#)).

Let $Z_n(t) = n^{-1/2} \sum_{i=1}^n (I_{(-\infty, t]}(X_i) - F(t))$ be the empirical process induced by the random variables X_n , $n \geq 1$, with common continuous distribution function F , where I_A represents the indicator function of the set A . For associated random variables, under suitable assumptions on their covariance structure (see [Louchichi \(2000\)](#) and [Oliveira and Suquet \(1995, 1998\)](#) for the best known results on the covariance decrease rates for the convergence in $D[0, 1]$, $L^2[0, 1]$ or $L^p[0, 1]$, respectively), Z_n converges weakly to a centered Gaussian process with covariance function

$$\begin{aligned} \Gamma(s, t) &= F(s \wedge t) - F(s)F(t) \\ &+ \sum_{k=1}^{\infty} (\mathbb{P}(X_1 \leq s, X_{k+1} \leq t) - F(s)F(t)) \\ &+ \sum_{k=1}^{\infty} (\mathbb{P}(X_1 \leq t, X_{k+1} \leq s) - F(s)F(t)). \end{aligned} \quad (1)$$

As it is well known, the asymptotic behavior of the empirical process is of great interest in many statistical applications. In several fields of statistics we often find transformations of the empirical process for which it is of interest to characterize their limit in distribution. The results about the asymptotic behavior of the empirical process are a valuable tool to accomplish this. Some classic examples are several goodness of fit tests statistics, such as the Kolmogorov–Smirnov and the Cramér–von Mises ω^2 test statistics, which are, respectively, the sup-norm and the $L^2[0, 1]$ norm of the uniform empirical process (the empirical process of $U[0, 1]$ distributed random variables). Another example of application may be found in [Shao and Yu \(1996\)](#), who are interested in integral functionals of the empirical process and in the mean residual life processes in reliability.

The above comments and remarks motivated the interest in the estimation of the covariance function (1). For this we will estimate the terms appearing in the series and sum a convenient number of these estimates to approximate Γ . We will concentrate on histogram estimators and in proving uniform strong convergence rates.

For each $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, the estimator for $F_k(s, t)$, the distribution function of (X_1, X_{k+1}) , is defined by

$$\hat{F}_{k,n}(s, t) = \frac{1}{n-k} \sum_{i=1}^{n-k} (I_{(-\infty, s]}(X_i) I_{(-\infty, t]}(X_{i+k})). \quad (2)$$

The asymptotic properties of this estimator were studied by [Henriques and Oliveira \(2003\)](#), who derived conditions on the covariance structure of the sequence X_n , $n \geq 1$, for the uniform almost sure consistency of this estimator and for the convergence in distribution of the finite dimensional distributions. Furthermore, the convergence rate of the mean square error was characterized. However, no rates were provided for the uniform strong convergence.

Combining the estimator $\widehat{F}_{k,n}(s, t)$ with the empirical distribution function defined by $\widehat{F}_n(s) = n^{-1} \sum_{i=1}^n I_{(-\infty, s]}(X_i)$, we obtain a natural estimator for the terms $\varphi_k(s, t) = F_k(s, t) - F(s)F(t)$, namely,

$$\widehat{\varphi}_{k,n}(s, t) = \widehat{F}_{k,n}(s, t) - \widehat{F}_n(s)\widehat{F}_n(t). \tag{3}$$

The estimators for the infinite sum in the expression of $\Gamma(s, t)$ and for $\Gamma(s, t)$ itself are, respectively,

$$\sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s, t) \tag{4}$$

and

$$\widehat{\Gamma}_n(s, t) = \widehat{F}_n(s \wedge t) - \widehat{F}_n(s)\widehat{F}_n(t) + \sum_{k=1}^{a_n} (\widehat{\varphi}_{k,n}(s, t) + \widehat{\varphi}_{k,n}(t, s)). \tag{5}$$

where $a_n \rightarrow +\infty$ is such that $\frac{a_n}{n} \rightarrow 0$.

It is well known that the covariance structure of a sequence of associated random variables determines its approximate independence (see, for example, Newman (1984) for a number of results regarding this). As a natural consequence, when dealing with associated samples it is common to have assumptions on the covariance structure of the random variables. Introducing the assumptions to be used in the results of this article we define

$$u(n) = \sum_{j=n+1}^{\infty} \text{Cov}^{1/3}(X_1, X_j). \tag{6}$$

Note the exponent 1/3 in the definition of $u(n)$. In the literature on associated variables this exponent is usually not present. Here it is convenient to include it in the definition of $u(n)$, as we are concerned with indicators of the associated variables rather than in the variables themselves. The need for this exponent is explained by inequality (7) below.

We now introduce two general assumptions to be used throughout the article.

- (S1) $X_n, n \geq 1$, is an associated and strictly stationary sequence of random variables having density function bounded by B_0 ; let $B_1 = 2 \max(2/\pi^2, 45B_0)$.
- (S2) Let $\theta > 0$. There exists a constant $C_0 > 0$ such that $u(n) \leq C_0 n^{-\theta}$, for all $n \geq 1$.

Note that under (S2),

$$u(0) = \sum_{j=1}^{\infty} \text{Cov}^{1/3}(X_1, X_j) < \infty.$$

In order to be able to identify explicit convergence rates for the proposed estimators, we consider the following assumptions on the decay rate of the covariances of the random variables.

- (G) Suppose that there exist $a_0 > 0$ and $a > 1$ such that $\text{Cov}(X_1, X_n) = a_0 a^{-n}$.
- (P) Suppose that $\text{Cov}(X_1, X_n) = a_0 n^{-a}$, with $a_0 > 0$ and $a > 3$.

Note that, under (G) assumption (S2) holds for all $\theta > 0$, while under (P) assumption (S2) holds for all $\theta \in (0, a/3 - 1]$.

Under (S1) we may apply relation (21) in Newman (1980) and Corollary to Theorem 1 in Sadikova (1996) to find (see Lemma 2.6 in Roussas (1995) for details)

$$\text{Cov}(I_{(-\infty, s]}(X_i), I_{(-\infty, t]}(X_j)) \leq B_1 \text{Cov}^{1/3}(X_i, X_j), \quad s, t \in \mathbb{R}, \quad (7)$$

where B_1 is defined in (S1).

In Henriques and Oliveira (2003), the authors have already derived uniform strong convergence rates for the estimators studied here. The method used in Henriques and Oliveira (2003) was based on exponential inequalities, and required quite fast decay of $\text{Cov}(X_1, X_n)$. In fact, the results in Henriques and Oliveira (2003) hold under a geometric decay rate of the covariances (G), but not under a polynomial decay (P). More precisely, under (G), it was established that the estimators $\widehat{F}_{k,n}$ and $\widehat{\varphi}_{k,n}$ are uniformly strong consistent with a convergence rate of order $n^{-1/3} \log^{2/3} n$. Furthermore, also under (G), it is shown that the estimators for the infinite sum $\sum_{k=1}^{\infty} \varphi_k(s, t)$ and for $\Gamma(s, t)$ converge almost surely at a rate of order $n^{-1/3} \log^{5/3} n$. Here, we will improve these results in two ways. First we will be able to characterize the rate of convergence not only under (G) but also under (P); second, under (G) we will derive faster rates of convergence than those mentioned previously. Furthermore, for the estimation of $F_k(s, t)$, under (G), we obtain convergence rates which approach the best possible ones for independent samples (see, in Sect. 3, the comment just before Theorem 3.3). The starting point for the derivation of these rates is a moment inequality for associated random variables by Shao and Yu (1996). Our method was inspired by Masry (2002), who uses moment inequalities to obtain convergence rates for the estimation of the density and its derivatives, considering also associated samples.

In Sect. 2, we will present some auxiliary results needed to establish the above mentioned convergence rates. The moment inequality referred to earlier is included in this section. The results establishing rates of uniform strong convergence are presented in Sects. 3 and 4, Sect. 3 dealing with the estimators $\widehat{F}_{k,n}(s, t)$ and $\widehat{\varphi}_{k,n}(s, t)$ and Sect. 4 with the estimators $\sum_{k=1}^{q_n} \widehat{\varphi}_{k,n}(s, t)$ and $\widehat{\Gamma}_n(s, t)$.

2 Auxiliary results

As mentioned earlier, the convergence rates that will be established in this article follow from a moment inequality, which is stated in the next lemma. This result is a version of Theorem 4.2 of Shao and Yu (1996), for bounded and strictly stationary random variables, where we took (referring to notation of Shao and Yu (1996)) f to be the identity function, $r = \infty$, $\theta = \frac{p-2}{2}$ and $\varepsilon = \frac{p}{2} - 1$.

Lemma 2.1 (Shao and Yu (1996)) *Let Y_n , $n \geq 1$, be a strictly stationary sequence of associated random variables satisfying $E(Y_j) = 0$ and $|Y_j| \leq M < \infty$, for $j \in \mathbb{N}$. Let $p > 2$ and assume that there exists a constant $C_1 > 0$ such that, for each $n \in \mathbb{N}$,*

$$\sum_{j=n+1}^{\infty} \text{Cov}(Y_1, Y_j) \leq C_1 n^{-\frac{p-2}{2}}.$$

Then, there exists a constant $K = K(p)$ such that, for all $n \in \mathbb{N}$,

$$E \left| \sum_{i=1}^n Y_i \right|^p \leq K n^{p/2} \left(E |Y_1|^p + \left(\max_{i \leq n} \sum_{j=1}^n \text{Cov}(Y_i, Y_j) \right)^{p/2} + M^{p-2} C_1 \right).$$

Based on the previous result we establish an inequality that will be essential in proving convergence rates.

Lemma 2.2 *Let $k \in \mathbb{N}_0$ be fixed and let ε_n be a sequence of positive numbers. Suppose (S1) and (S2) are satisfied with $\theta = \frac{p-2}{2}$, for some $p > 2$, for the latter. Then, there exists a constant $K = K(p)$ such that, for each $n > k$ and $s, t \in \mathbb{R}$,*

$$P \left(\left| \widehat{F}_{k,n}(s, t) - F_k(s, t) \right| > \varepsilon_n \right) \leq \frac{D}{\varepsilon_n^p (n - k)^{p/2}} \tag{8}$$

with

$$D = K \left(1 + (10B_1u(0))^{p/2} + c_k \right)$$

and

$$c_k = B_1 \max \{ 4C_0(k + 1)^\theta, 3C_0 + 2u(0)k^\theta \}.$$

Proof For each $n \in \mathbb{N}$ and fixed $s, t \in \mathbb{R}$ define

$$W_{k,n} = I_{(-\infty, s]}(X_n)I_{(-\infty, t]}(X_{k+n}) - F_k(s, t)$$

so that we can write

$$\widehat{F}_{k,n}(s, t) - F_k(s, t) = \frac{1}{n - k} \sum_{i=1}^{n-k} W_{k,i}.$$

Given (S1), since the $W_{k,n}$ are decreasing functions of the variables X_n , the sequence $W_{k,n}$, $n \geq 1$, is associated and strictly stationary. Furthermore, $|W_{k,n}| \leq 1$ and $E(W_{k,n}) = 0$, for each $n \geq 1$.

We want to apply Lemma 2.1 to the sequence $W_{k,n}$, $n \geq 1$, so that we need to check that there exists a positive constant C_1 such that, for all $m \geq 1$,

$$\sum_{j=m+1}^{\infty} \text{Cov}(W_{k,1}, W_{k,j}) \leq C_1 m^{-\theta}. \tag{9}$$

Applying a classical inequality by [Lebowitz \(1972\)](#) and (7) we obtain

$$\begin{aligned} \text{Cov}(W_{k,1}, W_{k,j}) &\leq B_1 \left[\text{Cov}^{1/3}(X_1, X_j) + \text{Cov}^{1/3}(X_1, X_{k+j}) \right. \\ &\quad \left. + \text{Cov}^{1/3}(X_{k+1}, X_j) + \text{Cov}^{1/3}(X_{k+1}, X_{k+j}) \right]. \end{aligned} \tag{10}$$

Thus, on account of (S2), we get, for every $m > k$,

$$\begin{aligned} \sum_{j=m+1}^{\infty} \text{Cov}(W_{k,1}, W_{k,j}) &\leq 2B_1u(m) + B_1u(m + k) + B_1u(m - k) \\ &\leq 2B_1C_0m^{-\theta} + B_1C_0(m + k)^{-\theta} + B_1C_0(m - k)^{-\theta} \\ &\leq 4B_1C_0(m - k)^{-\theta} \\ &\leq 4B_1C_0(k + 1)^\theta m^{-\theta}, \end{aligned}$$

where, for the last step we used the fact that $\left\{ \left(\frac{m}{m-k} \right)^\theta, m > k \right\}$ is a nonincreasing sequence in m , so bounded above by the first-term.

Moreover, for any $m \leq k$, we obtain from (10),

$$\begin{aligned} \sum_{j=m+1}^{\infty} \text{Cov}(W_{k,1}, W_{k,j}) &\leq 2B_1u(m) + B_1u(m+k) + B_1 \sum_{j=m+1}^{\infty} \text{Cov}^{1/3}(X_{k+1}, X_j) \\ &\leq 2B_1u(m) + B_1u(m+k) \\ &\quad + B_1 \left(\sum_{j=1}^{k-m+1} \text{Cov}^{1/3}(X_1, X_j) + \sum_{j=2}^{\infty} \text{Cov}^{1/3}(X_1, X_j) \right) \\ &\leq 2B_1u(m) + B_1u(m+k) + 2B_1u(0). \end{aligned}$$

Again using (S2) we get, for every $m \leq k$,

$$\begin{aligned} \sum_{j=m+1}^{\infty} \text{Cov}(W_{k,1}, W_{k,j}) &\leq 2B_1C_0m^{-\theta} + B_1C_0(m+k)^{-\theta} + 2B_1u(0) \\ &\leq 3B_1C_0m^{-\theta} + 2B_1u(0) = (3B_1C_0 + 2B_1u(0)m^\theta) m^{-\theta} \\ &\leq (3B_1C_0 + 2B_1u(0)k^\theta) m^{-\theta}. \end{aligned}$$

Setting $c_k = B_1 \max\{4C_0(k+1)^\theta, 3C_0 + 2u(0)k^\theta\}$, (9) then follows with $C_1 = c_k$.

Thus, by Lemma 2.1, there exists a constant $K = K(p)$ such that, for all $m \geq 1$,

$$\mathbb{E} \left| \sum_{i=1}^m W_{k,i} \right|^p \leq K m^{p/2} \left(1 + \left(2 \sum_{j=1}^m \text{Cov}(W_{k,1}, W_{k,j}) \right)^{p/2} + c_k \right), \tag{11}$$

given the stationarity of $W_{k,n}$, $n \geq 1$, and the fact that $|W_{k,n}| \leq 1$.

On account of (10) we obtain,

$$\begin{aligned} \sum_{j=1}^m \text{Cov}(W_{k,1}, W_{k,j}) &\leq 2B_1 \sum_{j=1}^m \text{Cov}^{1/3}(X_1, X_j) + B_1 \sum_{j=1}^m \text{Cov}^{1/3}(X_1, X_{k+j}) \\ &\quad + B_1 \sum_{j=1}^m \text{Cov}^{1/3}(X_{k+1}, X_j) \\ &\leq 2B_1u(0) + B_1 \sum_{j=k+1}^{\infty} \text{Cov}^{1/3}(X_1, X_j) \\ &\quad + B_1 \left(\sum_{j=1}^k \text{Cov}^{1/3}(X_{k+1}, X_j) + \sum_{j=k+1}^{\infty} \text{Cov}^{1/3}(X_{k+1}, X_j) \right) \\ &\leq 2B_1u(0) + B_1u(0) + 2B_1u(0) \\ &= 5B_1u(0). \end{aligned}$$

From (11) it now follows that, for all $m \geq 1$,

$$E \left| \sum_{i=1}^m W_{k,i} \right|^p \leq K \left(1 + (10B_1u(0))^{p/2} + c_k \right) m^{p/2}.$$

Finally, using the Markov inequality we find, for all $n > k$,

$$\begin{aligned} P \left(\left| \widehat{F}_{k,n}(s,t) - F_k(s,t) \right| \geq \varepsilon_n \right) &\leq \left(\frac{1}{\varepsilon_n(n-k)} \right)^p E \left| \sum_{i=1}^{n-k} W_{k,i} \right|^p \\ &\leq \frac{D}{\varepsilon_n^p (n-k)^{p/2}} \end{aligned}$$

with $D = K \left(1 + (10B_1u(0))^{p/2} + c_k \right)$. □

For the formulation of the next results we need to introduce some additional notation. Let t_n be a sequence of positive integers such that $t_n \rightarrow +\infty$. For each $n \in \mathbb{N}$ and each $i = 1, \dots, t_n$, put $x_{n,i} = Q(i/t_n)$, where Q is the quantile function of F . Define then, for $n, k \in \mathbb{N}$,

$$D_{n,k} = \sup_{s,t \in \mathbb{R}} \left| \widehat{F}_{k,n}(s,t) - F_k(s,t) \right|$$

and

$$D_{n,k}^* = \max_{i,j=1,\dots,t_n} \left| \widehat{F}_{k,n}(x_{n,i}, x_{n,j}) - F_k(x_{n,i}, x_{n,j}) \right|.$$

To prove a uniform version of the preceding lemma we will apply the following result which is proved in Theorem 2 of [Henriques and Oliveira \(2003\)](#).

Lemma 2.3 *If the sequence $X_n, n \geq 1$, satisfies (S1), then, for each $n \in \mathbb{N}$ and each $k \in \mathbb{N}_0$,*

$$D_{n,k} \leq D_{n,k}^* + \frac{2}{t_n} \quad a.s.$$

Lemma 2.4 *Let ε_n and t_n be two sequences of positive numbers such that $t_n \rightarrow +\infty$ and $\varepsilon_n t_n \rightarrow +\infty$, and $k \in \mathbb{N}_0$ be fixed. Suppose (S1) holds and (S2) is satisfied with $\theta = \frac{p-2}{2}$, for some $p > 2$. Then, for every n large enough,*

$$P \left(\sup_{s,t \in \mathbb{R}} \left| \widehat{F}_{k,n}(s,t) - F_k(s,t) \right| > \varepsilon_n \right) \leq \frac{2^p t_n^2 D}{\varepsilon_n^p (n-k)^{p/2}}, \tag{12}$$

with D defined as in Lemma 2.2.

Proof From Lemma 2.3 and taking into account that $\varepsilon_n t_n \rightarrow +\infty$ we obtain, for every n large enough,

$$\begin{aligned} P(D_{n,k} > \varepsilon_n) &\leq P\left(D_{n,k}^* + \frac{2}{t_n} > \varepsilon_n\right) \\ &\leq P\left(D_{n,k}^* > \frac{\varepsilon_n}{2}\right) + P\left(\frac{2}{t_n} > \frac{\varepsilon_n}{2}\right) \\ &\leq \sum_{i,j=1,\dots,t_n} P\left(|\widehat{F}_{k,n}(x_{n,i}, x_{n,j}) - F_k(x_{n,i}, x_{n,j})| > \frac{\varepsilon_n}{2}\right) \\ &\leq t_n^2 \max_{i,j=1,\dots,t_n} P\left(|\widehat{F}_{k,n}(x_{n,i}, x_{n,j}) - F_k(x_{n,i}, x_{n,j})| > \frac{\varepsilon_n}{2}\right). \end{aligned}$$

Now, apply Lemma 2.2 to complete the proof. □

3 Uniform strong convergence rates for $\widehat{F}_{k,n}$

Using Lemma 2.4 of the last section we will now obtain uniform strong convergence rates for the estimator $\widehat{F}_{k,n}$. In this section the letter C stands for a positive constant, which may take different values in each appearance. In each case the value of the constant is independent of n , but may depend on k and p .

Lemma 3.1 *Let $k \in \mathbb{N}_0$ be fixed. Suppose (S1) holds and (S2) is satisfied with $\theta = \frac{p-2}{2}$, for some $p > 2$. Then we have, for every $0 < \delta < \frac{p-2}{2}$,*

$$\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| = O\left((\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}}\right) \quad a.s.$$

Proof Fix $0 < \delta < \frac{p-2}{2}$ and put $t_n = \frac{\log n}{\varepsilon_n}$ in order to have $\varepsilon_n t_n \rightarrow +\infty$. Now, choosing $\varepsilon_n = (\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}}$, we obtain from Lemma 2.4, for n large enough,

$$P\left(\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| > \varepsilon_n\right) \leq C \frac{(\log n)^2}{\varepsilon_n^{p+2} (n-k)^{p/2}} \leq C n^{-(1+\delta)}.$$

The sequence on the right-hand-side above being summable, the result follows by the Borel–Cantelli Lemma. □

Note that, $\frac{p-2-2\delta}{2p+4}$ approaches $1/2$ as p grows to ∞ , so the convergence rate established in the previous lemma can be arbitrarily close to $n^{-1/2}$, if a sufficiently large p can be chosen. As stated in the next theorem, this is always possible under assumption (G), but not under (P). In fact, with polynomially decreasing covariances, in order to have a convergence rate close to $n^{-1/2}$, we need to impose a further assumption on the covariances, namely, that this polynomial decrease is fast enough. This leads to large values of the exponent appearing in (P).

Theorem 3.2 *Let $k \in \mathbb{N}$ be fixed and suppose (S1) holds.*

(a) *Under (G) we have, for every $0 < \gamma < 1/2$,*

$$\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| = O(n^{-\gamma}) \quad a.s.$$

(b) Under (P) we have, for every $0 < \gamma < \frac{1}{2} - \frac{3}{a+3}$,

$$\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| = O\left((\log n)^{\frac{3}{3+a}} n^{-\gamma}\right) \quad a.s.$$

Proof To prove a), fix $0 < \gamma < 1/2$. Now, choose $p > 2$ and $0 < \delta < \frac{p}{2} - 1$ so that $\frac{p-2-2\delta}{2p+4} > \gamma$. As already mentioned in Sect. 1, under (G) the assumption (S2) is satisfied regardless of the value of θ . In particular (S2) is satisfied for $\theta = \frac{p-2}{2}$. From Lemma 3.1, it follows that

$$\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| \leq C (\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}} \leq C n^{-\gamma} \quad a.s.,$$

proving part (a).

Under (P), as remarked in Sect. 1, assumption (S2) holds if $\theta \leq a/3 - 1$. According to Lemma 3.1, to obtain the best possible convergence rate, p must be chosen as large as possible. As $\theta = \frac{p-2}{2} \leq \frac{a}{3} - 1$, it follows that we must choose $p = 2a/3$. Then, for any fixed $\gamma \in \left(0, \frac{1}{2} - \frac{3}{a+3}\right)$, part (b) of the theorem follows directly from Lemma 3.1, replacing p by $2a/3$ and setting $\delta = \frac{a}{3} - 1 - 2\gamma \frac{a+3}{3}$.

Note that with $k = 0$ and $s = t$ the estimator $\widehat{F}_{k,n}(s,t)$ reduces to the one-dimensional empirical distribution function $\widehat{F}_n(s)$. So, the convergence rates of the previous theorem apply also to \widehat{F}_n . However, under (P) it is possible to obtain a slightly faster convergence rate for \widehat{F}_n . In fact, under the conditions of Lemma 2.4, with $k = 0$, we would obtain, for every n large enough,

$$\mathbb{P}\left(\sup_{s \in \mathbb{R}} |\widehat{F}_n(s) - F(s)| > \varepsilon_n\right) \leq \frac{2^p t_n D}{\varepsilon_n^p n^{p/2}}. \tag{13}$$

Then, following the arguments of the proofs of Lemma 3.1 and Theorem 3.2, we would find that, for every $0 < \gamma < \frac{1}{2} - \frac{9}{2(2a+3)}$,

$$\sup_{s \in \mathbb{R}} |\widehat{F}_n(s) - F(s)| = O\left((\log n)^{\frac{3}{3+2a}} n^{-\gamma}\right) \quad a.s.$$

This rate is somewhat faster than the rate given in part (b) of the previous theorem.

We note also that the convergence rate for the case of geometrically decreasing covariances is arbitrarily near the optimal rate for \widehat{F}_n , in the independent setting. In fact, for independent samples, the Law of the Iterated Logarithm implies that the best possible convergence rate for the one-dimensional empirical distribution function is $O\left(\left(\frac{\log \log n}{n}\right)^{1/2}\right)$, which is just slightly faster than the rate given in the previous theorem for the case of geometrically decreasing covariances.

The next theorem is the analogue of Theorem 3.2 for the estimator $\widehat{\varphi}_{k,n}$.

Theorem 3.3 *Let $k \in \mathbb{N}$ be fixed and suppose (S1) holds.*

(a) Under (G) we have, for every $0 < \gamma < 1/2$,

$$\sup_{s,t \in \mathbb{R}} |\widehat{\varphi}_{k,n}(s,t) - \varphi_k(s,t)| = O(n^{-\gamma}) \quad a.s.$$

(b) Under (P) we have, for every $0 < \gamma < \frac{1}{2} - \frac{3}{a+3}$,

$$\sup_{s,t \in \mathbb{R}} |\widehat{\varphi}_{k,n}(s,t) - \varphi_k(s,t)| = O\left((\log n)^{\frac{3}{3+a}} n^{-\gamma}\right) \quad a.s.$$

Proof As

$$\begin{aligned} \sup_{s,t \in \mathbb{R}} |\widehat{\varphi}_{k,n}(s,t) - \varphi_k(s,t)| &\leq \sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| + \sup_{s,t \in \mathbb{R}} |F(s)F(t) - \widehat{F}_n(s)\widehat{F}_n(t)| \\ &\leq \sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| + \sup_{s,t \in \mathbb{R}} F(s) |F(t) - \widehat{F}_n(t)| \\ &\quad + \sup_{s,t \in \mathbb{R}} \widehat{F}_n(t) |F(s) - \widehat{F}_n(s)| \\ &\leq \sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| + 2 \sup_{s \in \mathbb{R}} |F(s) - \widehat{F}_n(s)| \end{aligned}$$

the result follows immediately from Theorem 3.2. □

4 Uniform strong convergence rates for $\widehat{\Gamma}_n$

In this section, we will derive uniform strong convergence rates for the estimators of the sum $\sum_{k=1}^{\infty} \varphi_k(s,t)$ and of the covariance function $\Gamma(s,t)$. As in the previous section, C stands for a generic positive constant, but now independent of k and n (it may depend only on p).

Lemma 4.1 *Suppose (S1) holds and (S2) is satisfied with $\theta = \frac{p-2}{2}$, for some $p > 2$. Then, for each $0 < \delta < \frac{p-2}{2}$, if $a_n = n^{\frac{p-2-2\delta}{p^2+3p}}$ we have*

$$\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| = O\left((\log n)^{\frac{2}{p+2}} n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}}\right) \quad a.s.$$

Proof Let $0 < \delta < \frac{p-2}{2}$ and take $\varepsilon_n = (\log n)^{\frac{2}{2+p}} n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}}$ and $t_n = \frac{a_n}{\varepsilon_n} \log n$. Now, write

$$\mathbb{P}\left(\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| > \varepsilon_n\right) \leq \sum_{k=1}^{a_n} \mathbb{P}\left(\sup_{s,t \in \mathbb{R}} |\widehat{F}_{k,n}(s,t) - F_k(s,t)| > \frac{\varepsilon_n}{a_n}\right). \tag{14}$$

Note that, as $0 < \delta < \frac{p-2}{2}$, we have $\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)} > 0$ and $\frac{p-2-2\delta}{p^2+3p} > 0$, so that $\varepsilon_n \rightarrow 0, a_n \rightarrow +\infty, t_n \rightarrow +\infty$ and $\frac{\varepsilon_n}{a_n} t_n \rightarrow \infty$. Also, as $\frac{p-2-2\delta}{p^2+3p} < 1, \frac{a_n}{n} \rightarrow 0$.

From (14), applying Lemma 2.4 with $\frac{\varepsilon_n}{a_n}$ replacing ε_n , we conclude that there exists a constant $K = K(p)$ such that, for all n large enough,

$$\begin{aligned} &\mathbb{P}\left(\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| > \varepsilon_n\right) \\ &\leq \sum_{k=1}^{a_n} K \left(1 + (10B_1u(0))^{p/2} + c_k\right) \frac{2^p t_n^2 a_n^p}{\varepsilon_n^p (n-k)^{p/2}}, \end{aligned} \tag{15}$$

where $c_k = B_1 \max \{4C_0(k + 1)^\theta, 3C_0 + 2u(0) k^\theta\}$.

Since $k \leq a_n$, it follows that

$$c_k \leq B_1 \max \{4C_0(a_n + 1)^\theta, 3C_0 + 2u(0) a_n^\theta\} \leq C a_n^\theta,$$

so that from (15) we obtain, for n large enough,

$$\begin{aligned} & \mathbb{P} \left(\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| > \varepsilon_n \right) \\ & \leq \sum_{k=1}^{a_n} K \left(1 + (10B_1u(0))^{p/2} + C a_n^\theta \right) \frac{2^p t_n^2 a_n^p}{\varepsilon_n^p (n - k)^{p/2}} \\ & \leq \sum_{k=1}^{a_n} C a_n^\theta \frac{t_n^2 a_n^p}{\varepsilon_n^p (n - k)^{p/2}} \\ & \leq C \frac{t_n^2 a_n^{3p/2}}{\varepsilon_n^p (n - a_n)^{p/2}}, \end{aligned} \tag{16}$$

remembering that $\theta = \frac{p-2}{2}$.

By elementary manipulations it is easy to check that

$$-\frac{p^2 - p(4 + 2\delta) + 4(1 + \delta)}{2(p^2 + 3p)} = \frac{p - 2 - 2\delta}{p^2 + 3p} \cdot \frac{3p + 4}{2(p + 2)} - \frac{p - 2 - 2\delta}{2(p + 2)}, \tag{17}$$

so, we may write $\varepsilon_n = (\log n)^{\frac{2}{p+2}} a_n^{\frac{3p+4}{2(p+2)}} n^{-\frac{p-2-2\delta}{2(p+2)}}$. Inserting this and the choice made for t_n on the right-hand-side of (16) it follows that

$$\mathbb{P} \left(\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| > \varepsilon_n \right) \leq C \frac{n^{\frac{p-2-2\delta}{2}}}{(n - a_n)^{p/2}}.$$

As $\frac{a_n}{n} \rightarrow 0$, we have $\frac{n^{\frac{p-2-2\delta}{2}}}{(n - a_n)^{p/2}} \sim n^{-(1+\delta)}$, thus the sequence on the upper bound above is summable. Then, from the Borel–Cantelli Lemma it follows that

$$\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| = O \left((\log n)^{\frac{2}{p+2}} n^{-\frac{p^2 - p(4+2\delta) + 4(1+\delta)}{2(p^2 + 3p)}} \right) \quad a.s. \tag{18}$$

Now, we may write

$$\begin{aligned} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| & \leq \left| \sum_{k=1}^{a_n} (\widehat{\varphi}_{k,n}(s,t) - \varphi_k(s,t)) \right| + \left| \sum_{k=a_n+1}^{\infty} \varphi_k(s,t) \right| \\ & \leq \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| + a_n |F(t) - \widehat{F}_n(t)| \\ & \quad + a_n |F(s) - \widehat{F}_n(s)| + \left| \sum_{k=a_n+1}^{\infty} \varphi_k(s,t) \right|. \end{aligned}$$

Thus,

$$\begin{aligned} & \sup_{s,t \in IR} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| \\ & \leq \sup_{s,t \in IR} \left| \sum_{k=1}^{a_n} (\widehat{F}_{k,n}(s,t) - F_k(s,t)) \right| \\ & \quad + 2a_n \sup_{s,t \in IR} |F(t) - \widehat{F}_n(t)| + \sup_{s,t \in IR} \left| \sum_{k=a_n+1}^{\infty} \varphi_k(s,t) \right|. \end{aligned} \tag{19}$$

The convergence rate of the first term on the right-hand-side above is given in (18). From Lemma 3.1 the second term is almost surely $O\left(a_n (\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}}\right)$. Since $\frac{3p+4}{2p+4} > 1$ and taking into account (17), we have

$$\begin{aligned} a_n (\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}} & < a_n^{\frac{3p+4}{2p+4}} (\log n)^{\frac{2}{p+2}} n^{-\frac{p-2-2\delta}{2p+4}} \\ & = (\log n)^{\frac{2}{p+2}} n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}}. \end{aligned}$$

Thus,

$$a_n \sup_{s,t \in IR} |F(t) - \widehat{F}_n(t)| = O\left((\log n)^{\frac{2}{p+2}} n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}} \right), \quad a.s.$$

Finally, we will check that the third term on the right-hand-side of (19) is of the same order. According to (7), we have

$$\begin{aligned} \sup_{s,t \in IR} \left| \sum_{k=a_n+1}^{\infty} \varphi_k(s,t) \right| & = \sup_{s,t \in IR} \sum_{k=a_n+1}^{\infty} \text{Cov}(I_{(-\infty,s]}(X_1), I_{(-\infty,t]}(X_{k+1})) \\ & \leq B_1 \sum_{k=a_n+1}^{\infty} \text{Cov}^{1/3}(X_1, X_{k+1}) \\ & = B_1 u(a_n) \\ & \leq C a_n^{-\frac{p-2}{2}}, \end{aligned}$$

since assumption (S2) is satisfied for $\theta = \frac{p-2}{2}$. Now, it is easy to check that

$$a_n^{-\frac{p-2}{2}} = n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}},$$

hence,

$$\sup_{s,t \in IR} \left| \sum_{k=a_n+1}^{\infty} \varphi_k(s,t) \right| = O\left(n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}} \right)$$

so the proof is concluded. □

It is worth noticing that choosing, in the previous lemma, $a_n = (\log n)^{-\frac{4}{p^2+3p}} n^{\frac{p-2-2\delta}{p^2+3p}}$, we would have obtained the slightly improved convergence rate,

$$\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| = O \left((\log n)^{\frac{2(p-2)}{p^2+3p}} n^{-\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)}} \right), \quad a.s.$$

As the gain in the convergence rate is marginal, we will continue with the choice for a_n as in the lemma, just remarking the slightly improved version at the end.

Theorem 4.2 *Suppose (S1) holds.*

(a) *Under (G) we have, for every $0 < \gamma < 1/2$,*

$$\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| = O(n^{-\gamma}) \quad a.s.,$$

if $a_n = n^{\frac{p-2-2\delta}{p^2+3p}}$, with $\delta > 0$ and $p > 2$ chosen such that $\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)} > \gamma$.

(b) *Under (P) we have, for every $0 < \gamma < \frac{1}{2} - \frac{21a-18}{2a(2a+9)}$,*

$$\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| = O \left((\log n)^{\frac{3}{3+a}} n^{-\gamma} \right) \quad a.s.,$$

if $a_n = n^{\frac{3}{a-3}\gamma}$.

Proof Follow the arguments of the proof of Theorem 3.2, invoking Lemma 4.1 instead of Lemma 3.1 (for the proof of part b) take $\delta = \frac{a-3}{3} - \gamma \frac{a(2a+9)}{3(a-3)}$. \square

The next result states the convergence rates for the $\widehat{\Gamma}_n$.

Theorem 4.3 *Suppose (S1) holds.*

(a) *Under (G) we have, for every $0 < \gamma < 1/2$,*

$$\sup_{s,t \in \mathbb{R}} \left| \widehat{\Gamma}_n(s,t) - \Gamma(s,t) \right| = O(n^{-\gamma}) \quad a.s.,$$

if $a_n = n^{\frac{p-2-2\delta}{p^2+3p}}$, with $\delta > 0$ and $p > 2$ chosen such that $\frac{p^2-p(4+2\delta)+4(1+\delta)}{2(p^2+3p)} > \gamma$.

(b) *Under (P) we have, for every $0 < \gamma < \frac{1}{2} - \frac{21a-18}{2a(2a+9)}$,*

$$\sup_{s,t \in \mathbb{R}} \left| \widehat{\Gamma}_n(s,t) - \Gamma(s,t) \right| = O \left((\log n)^{\frac{3}{3+a}} n^{-\gamma} \right) \quad a.s.,$$

if $a_n = n^{\frac{3}{a-3}\gamma}$.

Proof First write

$$\begin{aligned} \sup_{s,t \in \mathbb{R}} \left| \widehat{\Gamma}_n(s,t) - \Gamma(s,t) \right| &\leq \sup_{s,t \in \mathbb{N}} \left| \widehat{F}_n(s \wedge t) - F(s \wedge t) \right| + 2 \sup_{s \in \mathbb{R}} \left| \widehat{F}_n(s) - F(s) \right| \\ &\quad + \sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right| \\ &\quad + \sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(t,s) - \sum_{k=1}^{\infty} \varphi_k(t,s) \right|. \end{aligned} \tag{20}$$

Thus, under (G) the result follows directly from Theorems 3.2 and 4.2.

For the proof of part (b), fix $\gamma \in \left(0, \frac{1}{2} - \frac{21a-18}{2a(2a+9)}\right)$. Since $\frac{21a-18}{2a(2a+9)} > \frac{3}{a+3}$, obviously $\gamma \in \left(0, \frac{1}{2} - \frac{3}{a+3}\right)$. It follows then from Theorem 3.2 that the two first terms of the right-hand-side of (20) are almost surely $O\left((\log n)^{\frac{3}{3+a}} n^{-\gamma}\right)$. Finally, if $a_n = n^{\frac{3}{a-3}\gamma}$, by Theorem 4.2 the fourth and fifth-terms are of the same order. \square

As mentioned earlier, choosing $a_n = (\log n)^{-\frac{4}{p^2+3p}} n^{\frac{p-2-2\delta}{p^2+3p}}$ we would obtain a slight improvement on our convergence rates, for the case of polynomially decreasing covariances. In fact, in the previous two theorems, with this choice for a_n , we would obtain the same convergence rate under (G), whereas under (P) (with p and δ as in the proof of Theorem 4.2), both $\sup_{s,t \in \mathbb{R}} \left| \sum_{k=1}^{a_n} \widehat{\varphi}_{k,n}(s,t) - \sum_{k=1}^{\infty} \varphi_k(s,t) \right|$ and $\sup_{s,t \in \mathbb{R}} \left| \widehat{\Gamma}_n(s,t) - \Gamma(s,t) \right|$ would be, almost surely,

$$O\left((\log n)^{\frac{6a-18}{2a^2+9a}} n^{-\gamma}\right)$$

for any $0 < \gamma < \frac{1}{2} - \frac{21a-18}{2a(2a+9)}$.

Acknowledgments Carla Henriques supported by Centro de Matemática da Universidade de Coimbra and PRODEP (Action 5.3). Paulo Eduardo Oliveira supported by Centro de Matemática da Universidade de Coimbra, Fundação para a Ciência e Tecnologia (FCT) and POCTI.

References

- Barlow R, Proschan F (1975) Statistical theory of reliability and life testing: Probability models. Holt, Rinehart and Winston, New York
- Esary JD, Proschan F, Walkup DW (1967) Association of random variables, with applications. *Ann Math Stat* 38:1466–1474
- Fortuin CM, Kasteleyn PW, Ginibre J (1971) Correlation inequalities on some partially ordered sets. *Commun Math Phys* 22:89–103
- Henriques C, Oliveira PE (2003) Estimation of a two-dimensional distribution function under association. *J Stat Plann Inf* 113:137–150
- Henriques C, Oliveira PE (2006) Convergence rates for the estimation of two-dimensional distribution functions under association and estimation of the covariance of the limit empirical process. *J Nonparametr Stat* 18:119–128
- Lebowitz J (1972) Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems. *Comm Math Phys* 28:313–321
- Louhichi S (2000) Weak convergence for empirical processes of associated sequences. *Ann Inst H Poincaré Probab Stat* 36:547–567
- Masry E (2002) Multivariate probability density estimation for associated processes: strong consistency and rates. *Stat Probab Lett* 58:205–219
- Newman CM (1980) Normal fluctuations and the FKG inequalities. *Comm Math Phys* 74:119–128
- Newman CM (1984) Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: Tong YL (ed) *Inequalities in Statistics and Probability*. IMS Lecture notes—monograph series, vol 5, pp 127–140 Inst Math Statist, Hayward, CA
- Oliveira PE, Suquet C (1995) $L^2[0,1]$ weak convergence of the empirical process for dependent variables. In: Antoniadis A, Oppenheim G (eds) *Lecture notes in statistics vol 103, wavelets and statistics, actes des XV rencontres Franco-Belges de Statisticiens (Ondelettes et Statistique)* pp 331–344, Springer, New York
- Oliveira PE, Suquet C (1998) Weak convergence in $L^p[0,1]$ of the uniform empirical process under dependence. *Stat Probab Lett* 39:363–370

- Roussas GG (1995) Asymptotic normality of a smooth estimate of a random field distribution function under association. *Stat Probab Lett* 24:77–90
- Sadikova SM (1996) Two-dimensional analogies of an inequality of Esseen with applications to the central limit theorem. *Theory Probab Appl* 11:325–335
- Shao QM, Yu H (1996) Weak convergence for weighted empirical process of dependent sequences. *Ann Probab* 24:2098–2127