On the multivariate kernel distribution estimator for distribution functions under association

Cecília Azevedo[1] and Paulo E. Oliveira[2]

[1] CMAT- Centro de Matemática
Universidade do Minho
4710-057 Braga, Portugal
(e-mail: cecilia@math.uminho.pt)

[2] Departamento de Matemática
Universidade de Coimbra
3000 Coimbra, Portugal
(e-mail: paulo@mat.uc.pt)

Abstract. In this note we consider the estimation of the multivariate distribution function $\mathbf{F}_p$ of the p-dimensional marginal of a stationary associated sequence. We show, under certain regularity conditions, the almost sure consistency and characterize the asymptotic behavior of the MSE. We also characterize the asymptotic optimal bandwidth. Under some stronger assumptions on the covariance this bandwidth rate is shown to be the same as for the independent case.
Keywords: Association, Kernel estimator, Optimal bandwidth, Mean squared error.

1 Introduction and assumptions

Estimation of distribution functions has been one of the main problems in statistics. Given a stationary sequence of random variables we will consider the estimator of it's p- dimensional marginal distribution function assuming some kind of positive dependence. The various types of positive dependence have received some interest in the literature since the early 1990's. We will consider associated random variables as introduced in Esary et al (1967). For the one-dimensional marginal, the estimator has been studied by Roussas [Roussas, 1993], [Roussas, 2000] and Cai, Roussas [Cai and Roussas, 1998]. Motivated by the need to approximate covariance functions appearing in the study of empirical processes Azevedo, Oliveira [Azevedo and Oliveira, 2000] and Henriques, Oliveira [Henriques and Oliveira, 2002] studied the two dimensional case. This note extends results in [Azevedo and Oliveira, 2000] for the p-dimensional case. We start by recalling the definition of association, as stated in Esary et al (1967).

Definition 1 *For a finite index set I, the random variables (r.v.'s) $\{X_i\}_{i\in I}$ are said to be associated, if for any real-valued coordinatewise increasing* *functions G and H defined on $\mathbb{R}^I$,* $\mathrm{Cov}\,\{G(X_i, i \in I), H(X_j, j \in I)\} \geq 0$, *provided* $\mathbb{E}\left(G^2(X_i, i \in I)\right) < \infty$ *and* $\mathbb{E}\left(H^2(X_j, j \in I)\right) < \infty$. *A sequence of r.v's is said to be associated if any finite subset of the r.v.'s is associated.*

Definition 2 *A smooth estimate of $\mathbf{F}_p$, d.f. of the random vector $\mathbf{X} = (X_1, \ldots, X_p)$, with $p \geq 2$, $\widehat{F}_{n,p}$ is defined, for each $\mathbf{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p$, by*

$$\widehat{F}_{n,p}(\mathbf{x}) = \frac{1}{n-p} \sum_{i=1}^{n-p} \mathbf{U}\left(\frac{\mathbf{x} - \mathbf{X}_{i,p}}{h_n}\right), \tag{1}$$

where $\mathbf{U}$ is a $p-$variate known d.f., the kernel function and, for each fixed p and $i = 1, \ldots, n-p$, $\mathbf{X}_{i,p} = (X_{i+1}, \ldots, X_{i+p})$. The (bandwidths) h_n are positive numbers tending to 0, as $n \to \infty$.

Jin, Shao [Jin and Shao, 1999] have been shown that, under independence, the optimal bandwidth of the p-dimensional kernel distribution estimator of $\mathbf{F}_p$ has order $n^{-1/3}$, for all dimensions. For associated samples, several properties of the univariate estimate $\widehat{F}_n$ of the marginal d.f. F have been investigated by Cai, Roussas [Cai and Roussas, 1998]. These authors proved that the optimal bandwidth rate is of order n^{-1}. The rate $n^{-1/3}$ becomes optimal under some stronger assumptions on the covariance structure. Azevedo, Oliveira [Azevedo and Oliveira, 2000] studied properties of the bivariate estimate $\widehat{F}_{n,k}$ of the d.f. of (X_1, X_{k+1}) with fixed $k = 1, \ldots n-1$, characterizing the optimal bandwidth rate. The results obtained on [Azevedo and Oliveira, 2000] extended the one-dimensional ones.

The set of conditions bellow are basically the same as in Cai, Roussas [Cai and Roussas, 1998] together with the conditions used by Jin, Shao [Jin and Shao, 1999] under independence.

Assumptions

(A_1) $\{X_n\}_{n \in \mathbb{N}}$ is a strictly stationary sequence of random variables with bounded density function f and continuous marginal distribution function F;

(A_2) The derivative of f exists and is continuous and bounded;

(A_3) The d.f., $\mathbf{F}_p$, of the random vector $\mathbf{X} = (X_1, \ldots, X_p)$ has bounded and continuous partial derivatives of first and second orders;

(A_4) For each positive integer j, the d.f. of $\mathbf{X}_{p,j} = (X_1, \ldots, X_p, X_{j+1}, \ldots, X_{j+p})$, $\mathbf{F}_{p,j}$, has bounded and continuous partial derivatives of first and second order;

(A_5) The kernel function $\mathbf{U}$ is $p-$differentiable and $\mathbf{u} = \frac{\partial^p \mathbf{U}}{\partial x_1 \ldots \partial x_p}$ is such that:

$$(i) \quad \int_{\mathbb{R}^p} \mathbf{u}(\mathbf{x}) d\mathbf{x} = 1; \; (ii) \quad \int_{\mathbb{R}^p} \mathbf{x}\mathbf{u}(\mathbf{x}) d\mathbf{x} = \mathbf{0}; \; (iii) \quad \int_{\mathbb{R}^p} \mathbf{x}\,\mathbf{x}^T\,\mathbf{u}(\mathbf{x}) d\mathbf{x} < \infty;$$

(A_6) The sequence of bandwidths is such that $n\,h_n^2 \to 0$;

(A_7) $\displaystyle\sum_{n=1}^{\infty} n\, Cov^{1/3}(X_1, X_n) < \infty;$ $(A_7)'$ $\displaystyle\sum_{n=1}^{\infty} Cov^{1/3}(X_1, X_n) < \infty;$

(A_8) $\mathbf{V} = \dfrac{\partial^p \mathbf{U}^2}{\partial x_1 \ldots \partial x_p}$, is such that $\displaystyle\int_{\mathbb{R}^p} \mathbf{x}\,\mathbf{x}^T\, \mathbf{V}(\mathbf{x}) d\mathbf{x} < \infty.$

Remark 1 *Note that* $\displaystyle\int_{\mathbb{R}^p} \mathbf{V}(\mathbf{x}) d\mathbf{x} = \mathbf{U^2}(+\infty, \ldots, +\infty) = 1.$

The conditions $(A_1), (A_2)$ and (A_7) have already been used in Cai and Roussas [Cai and Roussas, 1998] for the treatment of the univariate case. Note further that (A_7) implies $(A_7)'$ which implies the $L^2[0,1]$ weak convergence of empirical process, as proved in Oliveira and Suquet [Oliveira and Suquet, 1999].

Let us define the auxiliar functions $\mathbf{V_1}, \mathbf{V_2}$, $\mathbf{V_3}$ and $\mathbf{V_4}$ from $\mathbb{R}^p$ to $\mathbb{R}$, such that for each $\mathbf{x} = (x_1, \ldots, x_p)$,

- $\mathbf{V_1}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_i^2}(\mathbf{x}) \int_{\mathbb{R}^p} a_i^2 \mathbf{u}(\mathbf{a}) d\mathbf{a} + 2 \sum_{j=1}^{p-1} \sum_{i=j+1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_j\, \partial x_i}(\mathbf{x}) \int_{\mathbb{R}^p} a_i a_j \mathbf{u}(\mathbf{a}) d\mathbf{a};$
- $\mathbf{V_2}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial \mathbf{F}_p}{\partial x_i}(\mathbf{x}) \int_{\mathbb{R}^p} a_i \mathbf{V}(\mathbf{a}) d\mathbf{a};$
- $\mathbf{V_3}(\mathbf{x}) = \displaystyle\sum_{i=1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_i^2}(\mathbf{x}) \int_{\mathbb{R}^p} a_i^2 \mathbf{V}(\mathbf{a}) d\mathbf{a} + 2 \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \frac{\partial^2 \mathbf{F}_p}{\partial x_j\, \partial x_i}(\mathbf{x}) \int_{\mathbb{R}^p} a_i a_j \mathbf{V}(\mathbf{a}) d\mathbf{a};$
- $\mathbf{V_4}(\mathbf{x}) = \displaystyle\sum_{i=1}^{2p} \frac{\partial^2 \mathbf{F}_{p,j}}{\partial x_i^2}(\mathbf{x}, \mathbf{x}) \int_{\mathbb{R}^{2p}} a_i^2 \mathbf{u}(\mathbf{a}) d\mathbf{a} + 2 \sum_{i=1}^{2p-1} \sum_{j=i+1}^{2p} \frac{\partial^2 \mathbf{F}_{p,j}}{\partial x_j\, \partial x_i}(\mathbf{x}, \mathbf{x}) \int_{\mathbb{R}^{2p}} a_i a_j \mathbf{u}(\mathbf{a}) d\mathbf{a}.$

2 Consistency of the estimator.

In this section we present some results concerning to consistency of the estimator (1). We first show that $\widehat{F}_{n,p}$ is asymptotic unbiased, characterizing also the convergence rate of $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)$. To derive the asymptotic consistency of $\widehat{F}_{n,p}$, we apply a strong law of large numbers to the random variables $\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)$, $i = 1, \ldots, n-p$. To achieve this last step we shall need to characterize the behavior of each entry of the covariance matrix of the random vector whose entries are the preceding variables.

Theorem 1 *Suppose* $\{X_n\}_{n\in\mathbb{N}}$ *satisfy* $(A_1),(A_3)$ *and* (A_5). *Then, for each* $\mathbf{x}\in\mathbb{R}^p$,

$$\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)=\mathbf{F}_p(\mathbf{x})+\frac{V_1(\mathbf{x})}{2}h_n^2+o(h_n^2).$$

Proof: First note that the kernel estimator (1) can be written as

$$\widehat{F}_{n,p}(\mathbf{x})=\int_{\mathbb{R}^p}\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{s}}{h_n}\right)d\widehat{\phi_n}(\mathbf{s}),\tag{2}$$

where $\widehat{\phi_n}(\mathbf{x})=\dfrac{1}{n-p}\displaystyle\sum_{i=1}^{n-p}\mathbb{I}_{(-\infty,x_1]\times\cdots\times(-\infty,x_p]}(\mathbf{X}_{i,p})$, with $\mathbb{I}_A$ the characteristic function of the set A.

As $\mathbb{E}\left(\widehat{\phi_n}(\mathbf{x})\right)=\mathbf{F}_p(\mathbf{x})$, it follows from (2) applying Fubini's Theorem, that $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)=\displaystyle\int_{\mathbb{R}^p}\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{s}}{h_n}\right)d\mathbf{F}_p(\mathbf{s})=\int_{\mathbb{R}^p}\mathbf{u}(\mathbf{t})\mathbf{F}_p(\mathbf{x}-\mathbf{t}h_n)d\mathbf{t}$. Now, by using a Taylor expansion of order 2 of $\mathbf{F}_p$ and taking account of (A_3) and (A_5), and of the continuity of the second order partial derivatives of $\mathbf{F}_p$, (A_3), the result follows. ■

Note that (A_3) *and* (A_5) *are only required in order to establish a convergence rate. In fact, the convergence of* $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)$ *to* $\mathbf{F}_p(\mathbf{x})$ *follows from an application of the Dominated Convergence Theorem.*

In order to establish the almost sure convergence of (1) we need to control some covariances. Define

- $\mathbf{I}_{nj}(\mathbf{x})=\mathrm{Cov}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{1,p}}{h_n}\right),\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{j,p}}{h_n}\right)\right)$
- $\mathbf{I}_j(\mathbf{x})=\mathrm{Cov}\left(\mathbb{I}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{1,p}),\mathbb{I}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{j,p})\right)$.

Lemma 1 *Suppose that* $\{X_n\}_{n\in\mathbb{N}}$ *satisfy* (A_1), (A_3), (A_4) *and* (A_5). *Then, for each* $j>1$, *and* $\mathbf{x}\in\mathbb{R}^p$,

(i) $\mathbf{I}_{nj}(\mathbf{x})=\mathbf{I}_j(\mathbf{x})+O(h_n^2)=\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x})-\mathbf{F}_p^2(\mathbf{x})+O(h_n^2)$;

(ii) *For* $j>p-1,\mathbf{I}_j(\mathbf{x})\le\displaystyle\sum_{k=1}^{p}(p-k+1)\mathrm{Cov}^{1/3}(X_1,X_{j+k})+$
$+\sum_{k=1}^{p-1}(p-k)\mathrm{Cov}^{1/3}(X_1,X_{j-k+1})$.

Proof: Condition (i) follows from rewriting the covariance $I_{nj}=\displaystyle\int_{\mathbb{R}^{2p}}\mathbf{U}(\frac{\mathbf{x}-\mathbf{s}}{h_n})\mathbf{U}(\frac{\mathbf{x}-\mathbf{t}}{h_n})d\mathbf{F}_{p,j}(\mathbf{s},\mathbf{t})-\left(\int_{\mathbb{R}^p}\mathbf{U}(\frac{\mathbf{x}-\mathbf{s}}{h_n})d\mathbf{F}_p(s)\right)^2$. For the first term, writing the function $\mathbf{U}$ as an integral and by using Fubini's Theorem, we have $\displaystyle\int_{\mathbb{R}^{2p}}\mathbf{u}(\mathbf{a})\mathbf{u}(\mathbf{b})\mathbf{F}_{p,j}(\mathbf{x}-\mathbf{a})(\mathbf{x}-\mathbf{b})d\mathbf{a}d\mathbf{b}$. So, expanding $\mathbf{F}_{p,j}$ to the second order and using (A_4) and (A_5), this integral is equal to

$\mathbf{F}_{pj}(\mathbf{x},\mathbf{x}) + O(h_n^2)$, which together with the behavior of $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)$, completes the proof of (i). To prove condition (ii) we need use the inequality,

$$\text{Cov}\,\left(1\!\!\mathrm{I}_{(-\infty,s]}(Y_1), 1\!\!\mathrm{I}_{(-\infty,t]}(Y_2)\right) \leq M\text{Cov}^{\,1/3}(Y_1,Y_2), \tag{3}$$

where Y_1, Y_2 are associated random variables with common distribution function with a bounded density and $M > 0$ is constant (see Sadikova [Sadikova, 1966]), and the following lemma (Lebowitz [Lebowitz, 1972]),

Lemma 2 *Let A and B be subsets of $\{1,\ldots,n\}$ and x_i real with $i \in A \cup B$. Let $H_{A,B} = P(X_i > x_i, i \in A \cup B) - P(X_j > x_j, j \in A)P(X_k > x_k, k \in B)$.*
If $(X_1,\ldots,X_n)$ is associated then, $0 \leq H_{A,B} \leq \sum_{i\in A, j\in B} H_{\{i\},\{j\}}$.

In fact, according lemma 2,

$$\text{Cov}\,\left(1\!\!\mathrm{I}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{1,p}), 1\!\!\mathrm{I}_{(-\infty,\mathbf{x}]}(\mathbf{X}_{j,p})\right) \leq$$

$\leq \displaystyle\sum_{k=1}^{p}\sum_{i=1}^{p}\text{Cov}\,\left(1\!\!\mathrm{I}_{(-\infty,x_k]}(X_k), 1\!\!\mathrm{I}_{(-\infty,x_{j+i}]}(X_{j+i})\right).$

Now applying innequality (3), we have

$\text{Cov}\,\left(1\!\!\mathrm{I}_{(-\infty,x_k]}(X_k), 1\!\!\mathrm{I}_{(-\infty,x_{j+i}]}(X_{j+i})\right) \leq M\,\text{Cov}^{\,1/3}(X_k, X_{j+i})$, so

$\mathbf{I}_j(\mathbf{x}) \leq M \displaystyle\sum_{k=1}^{p}\sum_{i=1}^{p}\text{Cov}^{\,1/3}(X_k, X_{j+i})$. The sequence $\{X_n\}_{n\in\mathbb{N}}$ being stationary,

$\mathbf{I}_j(\mathbf{x}) \leq M\sum_{k=1}^{p}(p-k+1)\text{Cov}^{\,1/3}(X_1, X_{j+k}) + \sum_{k=1}^{p-1}(p-k)\text{Cov}^{\,1/3}(X_1, X_{j-k+1})$. ■

Remark 2 *Note that if the covariance sequence*

$$\{\text{Cov}\,(X_1, X_{j+1})\}_{j\in\mathbb{N}} \tag{4}$$

is decreasing, $\quad \mathbf{I}_j(\mathbf{x}) \leq p^2\text{Cov}^{\,1/3}(X_1, X_{j+1}).$

Theorem 2 *Suppose the variables $X_n, n \geq 1$, satisfy (A_1), (A_2), (A_3), (A_4), (A_5), (A_7) and (A_8). Then, for every $\mathbf{x} \in \mathbb{R}^p$, $\widehat{F}_{n,p}(\mathbf{x}) \to \mathbf{F}_p(\mathbf{x})$ almost surely.*

Proof: As proved in Theorem 1, $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right) \to \mathbf{F}_p(\mathbf{x})$, so it's enough to prove that the variables $\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{m,p}}{h_n}\right)$, $m \geq 1$ satisfy a strong law of large numbers. These variables are stationary and associated, as $\mathbf{U}$ is coordinatewise nondecreasing. Then, according to Newman [Newman, 1980] they satisfy a strong law of large numbers if

$$\lim_{n\to\infty}\frac{1}{n-p}\sum_{j=1}^{n-p} I_{n,j}(\mathbf{x}) = 0. \tag{5}$$

From conditions (i) and (ii) of the preceding lemma,
$I_{n,j}(\mathbf{x}) \leq M\sum_{k=1}^{p}(p-k+1)\mathrm{Cov}^{1/3}(X_1,X_{j+k}) + \sum_{k=1}^{p-1}(p-k)\mathrm{Cov}^{1/3}(X_1,X_{j-k+1}) + O(h_n^2)$. Now condition (5) is a consequence of (A_7) and association, so the theorem follows. ■

3 The behavior of the mean square error.

In this section we study the asymptotics and convergence rate of the mean square error (MSE). This characterization will then be used to derive the optimal bandwidth convergence rate. This convergence rate for the bandwidth is, as will be explained later, of order n^{-1}, thus a different convergence rate than the one in the independent case. But if we consider a decreasing rate on the sequence of the covariances (see Cai, Roussas [Cai and Roussas, 1998]) we obtain a convergence rate of order $n^{-1/3}$, as in the independent case (see Jin, Shao [Jin and Shao, 1999]), for all dimensions p.
As usual write $\mathrm{MSE}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) + \left(\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right) - \mathbf{F}_p(\mathbf{x})\right)^2$.
The behavior of $\mathbb{E}\left(\widehat{F}_{n,p}(\mathbf{x})\right)$ being known (cf.Theorem 1), we need to describe the asymptotics and convergence rate for the variance term.

Lemma 3 *Suppose the sequence* $\{X_n\}_{n\in\mathbb{N}}$ *satisfy* (A_1), (A_3), (A_4), (A_5) *and* (A_8). *Then for all* $\mathbf{x}$ *in* $\mathbb{R}^p$,

$$\textbf{(i)}\quad \mathbb{E}\big(\mathbf{U}^2(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n})\big) = \mathbf{F}_p(\mathbf{x}) - h_n\mathbf{V}_2(\mathbf{x}) + \frac{h_n^2}{2}\mathbf{V}_3(\mathbf{x}) + o(h_n^2)$$
$$\textbf{(ii)}\quad \left|\mathrm{Var}\left(\mathbf{U}\left(\tfrac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right) - \mathbf{F}_p(\mathbf{x})(1-\mathbf{F}_p(\mathbf{x})) + h_n\mathbf{V}_2(\mathbf{x})\right| = h_n^2(\mathbf{V}_3(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x})) + o(h_n^2).$$

Proof: In what concerns to (**i**), we have, by definition,
$\mathbb{E}\left(\mathbf{U}^2\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right) = \int_{\mathbb{R}^p}\mathbf{U}^2\left(\frac{\mathbf{x}-\mathbf{s}}{h_n}\right)d\mathbf{F}_p(\mathbf{s}) \int_{\mathbb{R}^p}\left(\int_{(-\infty,\mathbf{x}]}\mathbf{V}(\mathbf{a})d\mathbf{a}\right)d\mathbf{F}_p(\mathbf{s})$
By using Fubini Theorem and changing variables,
$\mathbb{E}\left(\mathbf{U}^2\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right) = \int_{\mathbb{R}^p}\mathbf{V}(\mathbf{a})\mathbf{F}_p(\mathbf{x}-\mathbf{a}h_n)d\mathbf{a}$. Using a Taylor expansion of order 2 of $\mathbf{F}_p$ and taking account of (A_5) and the definitions of $\mathbf{V}_2$ and $\mathbf{V}_3$, we have (**i**). In order to obtain (**ii**), knowing that
$\mathrm{Var}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right) = \mathbb{E}\left(\mathbf{U}^2\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right) - \left(\mathbb{E}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right)\right)\right)^2$, it is suffices to apply (**i**) and Theorem 1. ■

Definition 3 *Let* $\sigma^{\mathbf{2}}(\mathbf{x}) = \mathbf{F}_p(\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) + 2\sum_{j=2}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$ *and*
$\mathbf{c}_n(\mathbf{x}) = 2\sum_{j=n-p+1}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right) + \frac{2}{n-p}\sum_{j=2}^{n-p}(j-1)\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$

Theorem 3 *Suppose that* $\{X_n\}_{n\in\mathbb{N}}$ *satisfy* (A_1), (A_3), (A_4), (A_5), (A_6), (A_7) *and* (A_8). *Then*
$(n-p)\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \sigma^2(\mathbf{x}) - h_n\mathbf{V}_2(\mathbf{x}) + (n-p-1)h_n^2\left(\mathbf{V}_4(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x})\right) + O(h_n^2) - c_n(\mathbf{x})$.

Proof: $\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \dfrac{1}{(n-p)^2}\displaystyle\sum_{i,j=1}^{n-p}\mathrm{Cov}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{i,p}}{h_n}\right), \mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{j,p}}{h_n}\right)\right)$.
By stationarity, $\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \dfrac{1}{n-p}\mathrm{Var}\left(\mathbf{U}\left(\dfrac{\mathbf{x}-\mathbf{X}_{1,p}}{h_n}\right)\right) + \dfrac{2}{(n-p)^2}$
$\displaystyle\sum_{j=2}^{n-p}(n-p-j+1)\mathbf{I}_{n,p}(\mathbf{x})$ By using the preceding lemma and lemma 1,

$$(n-p)\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \mathbf{F}_p(\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) - \mathbf{V}_2(\mathbf{x})h_n + (\mathbf{V}_3(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x}))h_n^2 + $$
$$+\frac{2}{n-p}\sum_{j=2}^{n-p}(n-p-j+1)\times\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) + \frac{h_n^2}{2}(\mathbf{V}_4(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x}))\right).$$

We have now,

$$(n-p)\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \mathbf{F}_p(\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x}) - \mathbf{V}_2(\mathbf{x})h_n + (\mathbf{V}_3(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x}))h_n^2 + $$
$$+\sum_{j=2}^{n-p}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right) + (n-p-j+1)h_n^2\left(\mathbf{V}_4(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x})\right) - $$
$$\frac{2}{n-p}\sum_{j=2}^{n-p}(n-p-j+1)\times\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right) + O(h_n^2).$$

Replacing $\displaystyle\sum_{j=2}^{n-p}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$ by $\displaystyle\sum_{j=2}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$ and subtracting to later result $\displaystyle\sum_{j=n-p+1}^{\infty}\left(\mathbf{F}_{p,j}(\mathbf{x},\mathbf{x}) - \mathbf{F}_p^2(\mathbf{x})\right)$, we obtain now the expression for the variance of $\widehat{F}_{n,p}(\mathbf{x})$. ■

We may present now the behavior of the MSE.

Theorem 4 *Suppose* $\{X_n\}_{n\in\mathbb{N}}$, *satisfy* (A_1), (A_3), (A_4), (A_5), (A_6), (A_7) *and* (A_8). *Then,*
$(n-p)\mathrm{MSE}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \sigma^{\mathbf{2}}(\mathbf{x}) - h_n\,\mathbf{V_2}(\mathbf{x}) + O(n\,h_n^2) + o(h_n + n\,h_n^2) - \mathbf{c}_n(\mathbf{x})$.

Note that $c_n \to 0$, according to the assumptions made, and that c_n is independent of the bandwidth choice. It is now evident that an optimization of the convergence rate of the MSE is achieved by choosing $h_n = O(n^{-1})$ for all dimensions p. In fact, $h_n(\mathbf{x}) = \frac{\mathbf{V}_2(\mathbf{x})}{2(n-p-1)(\mathbf{V}_4(\mathbf{x}) - \mathbf{F}_p(\mathbf{x})\mathbf{V}_1(\mathbf{x}))}$.

To obtain, as in the independent case, the asymptotic optimal bandwidth of order $n^{-1/3}$, we replace assumptions (A_6) and (A_7) by,

(A_6^*) $nh_n^4 \to 0$ $\quad(A_7^*)$ $\sum_{j=1}^{\infty}\left(\mathrm{Cov}\,(X_1, X_{j+1})\right)^{\frac{1-\tau}{3}} < \infty,\ 0 < \tau < 1,$

as Cai and Roussas, 1998, did in the univariate case and providing that the sequence of covariances (4) is decreasing.

Theorem 5 *Suppose* $\{X_n\}_{n\in\mathbb{N}}$, *satisfy* (A_1), (A_3), (A_4), (A_5), (A_8), (A_6^*) *and* (A_7^*). *Then,*
$(n-p)\mathrm{MSE}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \sigma^2(\mathbf{x}) - h_n\,\mathbf{V_2}(\mathbf{x}) + O(n\,h_n^4) + o(h_n + nh_n^4) - \mathbf{c}_n(\mathbf{x})$.

Proof: To prove this result we use the identity $\mathbf{I}_{nj}(\mathbf{x}) = \mathbf{I}_j(\mathbf{x}) + O(h_n^2)$ (cf. Lema 1). As we noted in Remark 2, if we obtain an upper bound for $\mathbf{I}_j$ and, consequently, for $\mathbf{I}_{nj}$ we may use the following identity

$$|\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})| = |\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})|^{\tau}|\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})|^{1-\tau} \leq c^{\tau}\,h_n^{2\tau}\,p^{2(1-\tau)} \cdot\left|\left(\mathrm{Cov}^{1/3}(X_1, X_{j+1})\right)^{1-\tau}\right| = \tilde{c}\,h_n^{2\tau}\left|\left(\mathrm{Cov}^{1/3}(X_1, X_{j+1})\right)^{1-\tau}\right|,$$
where $\tilde{c} = c^{\tau}\,p^{2(1-\tau)}$ is constant.

If we consider the following expression for the variance,
$(n-p)\mathrm{Var}\left(\widehat{F}_{n,p}(\mathbf{x})\right) = \mathrm{Var}\left(\mathbf{U}\left(\frac{\mathbf{x}-\mathbf{X}_{1,p}}{h_n}\right)\right) +$
$+\dfrac{2}{n-p}\displaystyle\sum_{j=2}^{n-p}(n-p-j+1)\left|\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})\right| + \sum_{j=2}^{n-p}(n-p-j+1)\mathbf{I}_j(\mathbf{x})$, then,
$\dfrac{1}{n-p}\displaystyle\sum_{j=2}^{n-p}(n-p-j+1)\left|\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})\right| \leq \sum_{j=2}^{n-p}\left|\mathbf{I}_{nj}(\mathbf{x}) - \mathbf{I}_j(\mathbf{x})\right| \leq$
$\tilde{c}h_n^{2\tau}\sum_{j=2}^{\infty}\left(\mathrm{Cov}^{1/3}(X_1, X_{j+1})\right)^{1-\tau} = O(h_n^{2\tau})$, by using (A_7^*). The result now follows readily. ■

Once again, is now evident that an optimization of the convergence rate of the MSE is achieved by choosing $h_n = O(n^{-1/3})$, for all dimensions p.

Corollary 1 *Suppose* $\{X_n\}_{n\in\mathbb{N}}$, *satisfy* $(A_1), (A_3), (A_4), (A_5)$, (A_6^*), (A_7^*) *and* (A_8). *Suppose further that the covariance sequence (4) is decreasing. Then, the asymptotic optimal bandwidth* $\{h_n\}_{n\in\mathbb{N}}$ *of kernel estimator of* $\mathbf{F}_p$ *is, for all dimensions p, in the MSE sense, of order* $O(n^{-1/3})$.

This work has been partially supported by CMAT and FCT under the program POCI 2010.

References

[Azevedo and Oliveira, 2000]C. Azevedo and P.E. Oliveira. Kernel-type estimation of bivariate distribution function for associated random variables. In *New Trends in Probability and Statistics*, pages 17–25, 2000.

[Cai and Roussas, 1998]Z. Cai and G. Roussas. Efficient estimation of a distribution function under quadrant dependence. *Scand. J. Statist.*, pages 211–224, 1998.

[Henriques and Oliveira, 2002]C. Henriques and P. Oliveira. Covariance estimation for associated random variables. *préprint*, 2002.

[Jin and Shao, 1999]Z. Jin and Y. Shao. On kernel estimation of a multivariate distribution function. *Statist. Probab. Lett.*, pages 163–168, 1999.

[Lebowitz, 1972]J. Lebowitz. Bounds on the correlations and analyticity properties of ferromagnetic ising spin systems. *Comm. Math. Phys.*, pages 313–321, 1972.

[Newman, 1980]C.M. Newman. Normal fluctuations and the fkg inequalities. *Comm. Math. Phys.*, pages 119–128, 1980.

[Oliveira and Suquet, 1999]P. Oliveira and Ch. Suquet. An $l^2[0,1]$ invariance principle for lpqd random variables. *Port. Mathematica*, pages 367–379, 1999.

[Roussas, 1993]G. Roussas. Smooth estimation of the distribution function under association: Asymptotic normality. *Tech. Report*, 1993.

[Roussas, 2000]G. Roussas. Asymptotic normality of the function under *Statist. Probab. Lett.*, pages 1–12, 2000.

[Sadikova, 1966]S.M. Sadikova. Two dimensional analogues of an inequality of essén with applications to the central limit theorem. *Theory Probab. Appl.*, pages 325–335, 1966.