Mathematical Sciences And Applications E-Notes
Volume 2 No. 1 pp. 106–116 (2014) ©MSAEN

ESTIMATION USING COPULA FUNCTION IN REGRESSION MODEL

DJAMILA BENNAFLA, ABBES RABHI AND AMINA ANGELIKA BOUCHENTOUF

(Communicated by Kamile SANLI KULA)

Abstract. Copula models are becoming an increasingly powerful tool for modeling the dependencies between random variables, they have useful applications in many fields such as biostatistics, actuarial science, and finance. In this paper, we investigate the estimate of a regression model, by use of the copula representation. Its asymptotic properties are studied; almost surely convergence and convergence in probability (with rate).

1. Introduction

Copula theory, following the works of Sklar in 1959, allows a flexible modeling of dependence between two or more random variables. In recent years, the growing interest for this theory is phenomenal. In [18] Thomas Mikosch stated that in September 2005, a Google search on the term "copula" produced 650,000 results. Then, in January 2007, this same query generates more than 1.13 million. Given the number of publications in scientific journals and the number of papers available on Internet, it is undeniable that passion to the copula theory is still booming.

The progress of applications of this theory is wide in the field of finance, risk management, performance evaluation of assets, the valuation of derivatives, the extreme value theory, contagion require flexible and practical models of addiction.

The construction and properties of copulas have been studied rather extensively during the last 15 years or so. Hutchinson and Lai (1990) [15] were among the early authors who popularized the study of copulas. Nelsen (1999) [20] presented a comprehensive treatment of bivariate copulas, while Joe (1997) [16] devoted a chapter of his book to multivariate copulas. Further authoritative updates on copulas are given in Nelsen (2006) [19]. Copula methods have many important applications in insurance and finance Cherubini *et al.* (2004) [3] and Embrechts *et al.* (2003) [6].

Briefly speaking, copulas are functions that join multivariate distributions to their one-dimensional marginal distribution functions. Equivalently, copulas are multivariate distributions whose marginals are uniform on the interval $(0, 1)$. In this paper, we restrict our attention to bivariate copulas. Fisher (1997) [13] gave

Date: Received: October 6, 2013; Accepted: April 17, 2014.
2010 *Mathematics Subject Classification.* Firstly 62N02, Secondly 62H12, Thirdly 62J05.
Key words and phrases. Copulas, non parametric estimation, regression model.

two major reasons as to why copulas are of interest to statisticians: firstly, as a way of studying scale-free measures of dependence; and secondly, as a starting point for constructing families of bivariate distributions." Specifically, copulas are an important part of the study of dependence between two variables since they allow us to separate the effect of dependence from the effects of the marginal distributions. This feature is analogous to the bivariate normal distribution where the mean vectors are unlinked to the covariance matrix and jointly determine the distribution. Many authors have studied constructions of bivariate distributions with given marginals: This may be viewed as constructing a copula.

Nonparametric estimators of copula densities have been suggested by Gijbels and Mielnicsuk [14] and Fermanian and Scaillet [10], who used kernel methods, Sancetta [24] and Sancetta and Satchell [25], who used techniques based on Bernstein polynomials. Biau and Wegkamp[1] proposed estimating the copula density through a minimum distance criterion. Faugeras [7] in his thesis studied the quantile copula approach to conditional density estimation.

The aim of this paper is devoted to the estimation of a regression model via a copulae function, the rest of the paper is organized as follows; at first in section 2 we state Sklar's theorem which elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals and at the end of the section we introduce our model, then in section 3 we make some regularity assumptions on the kernels and the densities which, although far from being minimal, are somehow customary in kernel density estimation, the main result and its proof is given in the fourth part of this paper. Then we finish this work by a small conclusion.

2. The model

Let $(X_i; Y_i); i = 1, 2, \ldots, n$ be an independent identically distributed sample from real-valued random variables (X, Y) sitting on a given probability space. For predicting the response Y of the input variable X at a given location x, it is of great interest to estimate not only the conditional mean or regression function $\mathbb{E}(Y/X = x)$, but the full conditional density $f(y/x)$. Indeed, estimating the conditional density is much more informative, since it allows not only to recalculate from the density the conditional expected value $\mathbb{E}(Y/X)$, but also many other characteristics of the distribution such as the conditional variance. In particular, having knowledge of the general shape of the conditional density, is especially important for multi-modal or skewed densities, which often arise from nonlinear or non- Gaussian phenomena, where the expected value might be nowhere near a mode, i.e. the most likely value to appear.

A natural approach to estimate the conditional density $f(y/x)$ of Y given $X = x$ would be to exploit the identity

$$f(y/x) = \frac{f_{XY}(x, y)}{f_X(x)}, \quad f_X(x) \neq 0, \tag{2.1}$$

where f_{XY} and f_X denote the joint density of (X, Y) and X, respectively.

By introducing Parzen-Rosenblatt [21, 22] kernel estimators of these densities, namely,

$$\widehat{f}_{n,XY}(x, y) = \frac{1}{n} \sum_{i=1}^{n} K'_{h'}(X_i - x) K_h(Y_i - y),$$

$$\widehat{f}_{n,X}(x) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x),$$

where $K_h(.) = \frac{1}{h}K(./h)$ and $K'_{h'}(.) = \frac{1}{h'}K'(./h')$ are (rescaled) kernels with their associated sequence of bandwidth $h = h_n$ and $h' = h'_n$ going to zero as $n \to 1$, one can construct the quotient

$$\widehat{f}_n(y/x) = \frac{\widehat{f}_{n,XY}(x,y)}{\widehat{f}_{n,X}(x)},$$

and obtain an estimator of the conditional density.

Formally, Sklar's theorem below elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals see Sklar[28].

Theorem 2.1. ***(Sklar 1959)*** *For any bivariate cumulative distribution function $F_{X,Y}$ on $\mathbb{R}^2$, with marginal cumulative distribution functions F of X and G of Y, there exists some function $C : [0,1]^2 \to [0,1]$, called the dependence or copula function, such as*

$$F_{X,Y}(x,y) = C(F(x), G(y)), \quad -\infty \leq x, y \leq +\infty. \tag{2.2}$$

If F and G are continuous, this representation is unique with respect to (F,G). The copula function C is itself a cumulative distribution function on $[0,1]^2$ with uniform marginals.

This theorem gives a representation of the bivariate c.d.f. as a function of each univariate c.d.f. In other words, the copula function captures the dependence structure among the components X and Y of the vector (X,Y), irrespectively of the marginal distribution F and G. Simply put, it allows to deal with the randomness of the dependence structure and the randomness of the marginals separately.

Copulas appear to be naturally linked with the quantile transform: in the case F and G are continuous, formula (2.2) is simply obtained by defining the copula function as $C(u,v) = F_{X,Y}(F^{-1}(u), G^{-1}(v)),\ 0 \leq u \leq 1,\ 0 \leq v \leq 1$. For more details regarding copulas and their properties, one can consult for example the book of Joe [17]. Copulas have witnessed a renewed interest in statistics, especially in finance, since the pioneering work of Räuschendorf [23] and Deheuvels [4], who introduced the empirical copula pro- cess. Weak convergence of the empirical copula process was investigated by Deheuvels [5], Van der Vaart and Wellner [29], Fermanian, Radulovic and Wegkamp [9]. For the estimation of the copula density, refer to Gijbels and Mielniczuk [14], Fermanian [8] and Fermanian and Scaillet [11].

From now on, we assume that the copula function $C(u,v)$ has a density $c(u,v)$ with respect to the Lebesgue measure on $[0,1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u,v)$ and $c(u,v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U,V) = (F(x), G(y))$. By differentiating formula (2.2), we get for the joint density,

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = f(x)g(y)c(F(x), G(y)),$$

where $c(u,v) := \frac{\partial^2 C(u,v)}{\partial u \partial v}$ is the above mentioned copula density. Eventually, we can obtain the following explicit formula of the conditional density

$$f(y/x) = \frac{f_{XY}(x,y)}{f(x)} = g(y)c(F(x),G(y)), \quad f(x) \neq 0. \tag{2.3}$$

So, let

$$f_n(y/x) = \widehat{g}_n(y)\widehat{c}_n(F_n(x),G_n(y)),$$

be an estimator which builds on the idea of using synthetic data. where $\widehat{g}_n(y)$, $\widehat{c}_n$, $F_n(x)$, $G_n(y)$ are estimators of the density g of Y, the copula density c, the c.d.f. F of X and G of Y respectively. Its study then reveals to be particularly simple: it reduces to the ones already done on nonparametric density estimation.

From now on, we assume that the copula function $C(u,v)$ has a density $c(u,v)$ with respect to the Lebesgue measure on $[0,1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u,v)$ and $c(u,v)$ are then the cumulative distribution function (c.d.f.) and density respectively of the transformed variables $(U,V) = (F(X),G(Y))$.

Now, To build an estimator of the conditional density we have to use a Parzen-Rosenblatt kernel type non parametric estimator of the marginal density g of Y.

$$\widehat{g}_n(y) := \frac{1}{nh_n}\sum_{i=1}^{n} K_0\left(\frac{y-Y_i}{h_n}\right),$$

the empirical distribution functions $F_n(x)$ and $G_n(y)$ for $F(x)$ and $G(y)$ respectively,

$$F_n(x) = \sum_{j=1}^{n} 1_{X_j \leq x} \text{ and } G_n(y) = \sum_{j=1}^{n} 1_{Y_j \leq y}.$$

Concerning the copula density $c(u,v)$, we noted that $c(u,v)$ is the joint density of the transformed variables $(U,V) := (F(x),G(y))$. Therefore, $c(u,v)$ can be estimated by the bivariate Parzen-Rosenblatt kernel type non parametric density (pseudo) estimator,

$$c_n(u,v) := \frac{1}{nh_nb_n}\sum_{i=1}^{n} K\left(\frac{u-U_i}{h_n}, \frac{v-V_i}{b_n}\right), \tag{2.4}$$

where K is a bivariate kernel and h_n, b_n its associated bandwidth. For simplicity, we restrict ourselves to product kernels, i.e. $K(u,v) = K_1(u)K_2(v)$ with the same bandwidths $h_n = b_n$.

Nonetheless, since F and G are unknown, the random variables (U_i, V_i) are not observable, i.e. c_n is not a true statistic. Therefore, we approximate the pseudo-sample (U_i, V_i), $i = 1,2,\ldots,n$ by its empirical counterpart $(F_n(X_i), G_n(Y_i))$, $i = 1,2,\ldots,n$. We therefore obtain a genuine estimator of $c(u,v)$.

$$\widehat{c}_n(u,v) := \frac{1}{nh_n^2}\sum_{i=1}^{n} K_1\left(\frac{u-F_n(X_i)}{h_n}\right) K_2\left(\frac{v-G_n(Y_i)}{b_n}\right). \tag{2.5}$$

Now, let us present Our estimated model, the regression function $r(x)$, is given as follows:

$$r(x) = Yc_n(F(x),G(y)), \quad |Y| \leq M, \quad Y, m \in \mathbb{R}.$$

This regression function $r(x)$ is estimated by a function $\widehat{r} = Y\widehat{c}_n(F(x),G(y))$.

To state our main result, we will have to make some regularity assumptions on the kernels and the densities which, although far from being minimal, are somehow customary in kernel density estimation.

3. Notations and Assumptions

Set x and y two fixed points in the interior of $supp(f)$ and $supp(g)$ respectively. The support of the densities function f and g are noted by

$$supp(f) = \overline{\{x \in \mathbb{R}; f(x) > 0\}} \quad \text{and} \quad supp(g) = \overline{\{y \in \mathbb{R}; g(y) > 0\}},$$

where A stands for the closure of a set A.

N.B. $o_P(.)$ and $O_p(.)$ (respectively $o_{a.s}(.)$ and $O_{a.s}(.)$)will stands for convergence and boundedness in probability (respectively almost surely).

Assumptions

- (i) the c.d.f F of X and G of Y are strictly increasing and differentiable.
- (ii) the densities g and c are twice continuously differentiable with bounded second derivatives on their support.
- (iii) the densities g and c are uniformly continuous and non-vanishing almost everywhere on a compact set $J := [a, b]$ and $D \subset (0, 1) \times (0, 1)$ included in the interior of $supp(g)$ and $supp(c)$, respectively.
- (iv) K and K_0 are of bounded support and of bounded variation.
- (v) $0 \leq K \leq C$ and $0 \leq K_0 \leq C$ for some constant C.
- (vi) K and K_0 are second order kernels.
- (vii) K it is twice differentiable with bounded second partial derivatives.

Recall that $c_n(u, v)$ is the kernel copula (pseudo) density estimator from the unobservable, but fixed with respect to n, pseudo data $(F(X_i), G(Y_i))$, and that $\widehat{c}_n(u, v)$ is its analogue made from the approximate data $(F_n(X_i), G_n(Y_i)$. The heuristic of the reason why our estimator works is that the $n^{-1/2}$ in probability rate of convergence in uniform norm of F_n and G_n to F and G is faster than the $1/\sqrt{na_n^2}$ rate of the non parametric kernel estimator c_n of the copula density c. Therefore, the approximation step of the unknown transformations F and G by their empirical counterparts F_n and G_n does not have any impact asymptotically on the estimation step of c by c_n. Put in another way, one can approximate $\widehat{c}_n(F_n(x), G_n(y))$ by $c_n(F(x), G(y))$ at a faster rate than the convergence rate of $c_n(F(x), G(y))$ to $c(F(x), G(y))$.

4. Main Result

This part of the paper is devoted to the study of almost surely convergence and convergence in probability (with rate) of our estimator introduced above.

Theorem 4.1. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied, if h_n tends to zero as $n \to \infty$ in such a way that*

$$na_n^4 \to \infty, \frac{\sqrt{\ln \ln n}}{na_n^3} \to 0,$$

then,

$$\widehat{r}_n(x) = r(x) + O_P\left(h_n^2 + \frac{1}{\sqrt{nh_n^2}} + \frac{1}{nh_n^4} + \frac{\sqrt{\ln \ln n}}{nh_n^3}\right)$$

Proof. Let $\widehat{r}(x) = Yc_n(F(x), G(x))$, to demonstrate that $\widehat{r}(x)$ converge to $r(x)$ it is sufficient to prove that $\widehat{c}_n(U, V) \to c_n(U, V)$, with $U = F(x)$, $V = G(x)$.

For $(X_i, i = 1, 2, \ldots, n)$ an i.i.d. sample of a real random variable X with common c.d.f. F, the Kolmogorov-Smirnov statistic is defined as $D_n := \|F_n - F\|$. Glivenko-Cantelli, Kolmogorov and Smirnov, Chung, Donsker among others have studied its convergence properties in increasing generality (See e.g. [27] and [28] for recent accounts). For our purpose, we only need to formulate these results in the following rough form:

Lemma 4.1. *For an i.i.d. sample from a continuous c.d.f. F,*

$$\|F_n - F\|_\infty = O_P\left(\frac{1}{\sqrt{n}}\right), \quad i = 1, 2, \ldots, n, \tag{4.1}$$

$$\|F_n - F\|_\infty = O_{a.s}\left(\frac{\ln \ln n}{n}\right) \quad i = 1, 2, \ldots, n. \tag{4.2}$$

Since F is unknown, the random variables $U_i = F(X_i)$ are not observed. As a consequence of the preceding lemma, one can naturally approximate these variables by the statistics $F_n(X_i)$. Indeed,

$$\|F(X_i) - F_n(X_i)\| \leq \sup_{x\in\mathbb{R}} \|F(x) - F_n(x)\| = \|F_n - F\|_\infty a.s.$$

Let

$$c_n(U, V) = \frac{1}{na_n^2} \sum_{i=1}^{n} K_1\left(\frac{U - F_n(x_i)}{a_n}\right) K_2\left(\frac{V - G_n(y_i)}{a_n}\right),$$

$$\widehat{c}_n(U, V) = \frac{1}{na_n^2} \sum_{i=1}^{n} K_1\left(\frac{U - F_n(x_i)}{a_n}\right) K_2\left(\frac{V - G_n(y_i)}{a_n}\right).$$

So, we must show that $F_n(x_i)$ converge to $F(x_i)$ and $G_n(y_i)$ converge to $G(y_i)$.

$$\widehat{c}_n(U, V) - c_n(U, V) = \frac{1}{na_n^2}\left(\sum_{i=1}^{n} K_1\left(\frac{U - F_n(x_i)}{a_n}\right) K_2\left(\frac{V - G_n(y_i)}{a_n}\right)\right.$$
$$\left. - \sum_{i=1}^{n} K_1\left(\frac{U - F(x_i)}{a_n}\right) K_2\left(\frac{V - G(y_i)}{a_n}\right)\right),$$

with

$$\Pi_{i,n} = K_1\left(\frac{U - F_n(x_i)}{a_n}\right) K_2\left(\frac{V - G_n(y_i)}{a_n}\right) - K_1\left(\frac{U - F(x_i)}{a_n}\right) K_2\left(\frac{V - G(y_i)}{a_n}\right).$$

Let

$$Z_{i,n} = \begin{pmatrix} F_n(x_i) - F(x_i) \\ G_n(y_i) - G(y_i) \end{pmatrix}$$

$\|F_n(X_i) - F(X_i)\| \leq \|F_n - F\|_\infty$ and $\|G_n(Y_i) - G(Y_i)\| \leq \|Gn - G\|$ a.s. for every i = 1,2,..., n. Preceding Lemma thus entails that the norm of $Z_{i,n}$ is independent of i and such that

$$\|Z_{i,n}\| = O_P\left(\frac{1}{\sqrt{n}}\right), \quad i = 1, 2, \dots, n, \tag{4.3}$$

$$\|Z_{i,n}\| = O_{a.s}\left(\frac{\ln\ln n}{n}\right) \quad i = 1, 2, \dots, n. \tag{4.4}$$

Now, for every fixed $(u, v) \in [0, 1]^2$, since the kernel K is twice differentiable, there exists, by Taylor expansion, random variables $\tilde{U}_{i,n}$ and $\tilde{V}_{i,n}$ such that, almost surely,

$$\Pi = \frac{1}{na_n^3}\sum_{i=1}^{n} Z_{i,n}^T \nabla\left(K_1\left(\frac{U - F_n(x_i)}{a_n}\right)K_2\left(\frac{V - G_n(y_i)}{a_n}\right)\right)$$
$$+\frac{1}{2na_n^4}\sum_{i=1}^{n} Z_{i,n}^T \nabla^2\left(K_1\left(\frac{U - \tilde{U}_{i,n}}{a_n}\right)K_2\left(\frac{V - \tilde{V}_{i,n}}{a_n}\right)\right)Z_{i,n} = \Pi_1 + \Pi_2,$$

where $Z_{i,n}^T$ denotes the transpose of the vector $Z_{i,n}$ and ∇K and $\nabla^2 K$ the gradient and the Hessian respectively of the multivariate kernel function K.

By centering at expectations, decompose further the first term Π_1 as,

$$\Pi_1 = \frac{1}{na_n^3}\sum_{i=1}^{n} Z_{i,n}\nabla\left(K_1\left(\frac{U - F(x_i)}{a_n}\right)K_2\left(\frac{V - G(y_i)}{a_n}\right)\right) - \mathbb{E}\nabla\left(K_1\left(\frac{U - F(x_i)}{a_n}\right)K_2\left(\frac{V - G(y_i)}{a_n}\right)\right)$$
$$+\frac{1}{na_n^3}\sum_{i=1}^{n} Z_{i,n}^T \mathbb{E}\nabla\left(K_1\left(\frac{U - F(x_i)}{a_n}\right)K_2\left(\frac{V - G(y_i)}{a_n}\right)\right) = \Pi_{11} + \Pi_{12}$$

We again decompose one step further Π_{11}, Set

$$A_i = \nabla\left(K_1\left(\frac{U - F(x_i)}{a_n}\right)K_2\left(\frac{V - G(y_i)}{a_n}\right)\right) - \mathbb{E}\nabla\left(K_1\left(\frac{U - F(x_i)}{a_n}\right)K_2\left(\frac{V - G(y_i)}{a_n}\right)\right).$$

Then

$$|\Pi_{11}| \leq \frac{\|Z_{i,n}\|}{na_n^3}\sum_{i=1}^{n}(\|A_i\| - \mathbb{E}\|A_i\|) + \frac{\|Z_{i,n}\|}{na_n^3}\sum_{i=1}^{n}\mathbb{E}\|A_i\| = \Pi_{111} + \Pi_{112}.$$

We now proceed to the study of the order of each terms in the previous decompositions.

• **Negligibility of** Π_2.

By the boundedness assumption on the second-order derivatives of the kernel, and equations (4.3) and (4.4),

$$\Pi_2 = O_P\left(\frac{1}{na_n^4}\right), \text{ and } \Pi_2 = O_{a.s}\left(\frac{\ln\ln n}{na_n^4}\right)$$

• **Negligibility of** Π_{12}.

Bias results on the bivariate gradient kernel estimator (See Scott [26] chapter 6) entail that

$$\mathbb{E}\nabla\left(K_1\left(\frac{U-F(x_i)}{a_n}\right)K_2\left(\frac{V-G(y_i)}{a_n}\right)\right)=a_n^3\nabla c(u,v)+O(a_n^5)$$

Cauchy-Schwarz inequality yields that

$$|\Pi_{12}|\leq\frac{\|nZ_{i,n}\|}{na_n^3}\left\|\mathbb{E}\nabla\left(K_1\left(\frac{U-F(x_i)}{a_n}\right)K_2\left(\frac{V-G(y_i)}{a_n}\right)\right)\right\|$$

In turn, with equations (4.3) and (4.4),

$$\Pi_{12}=O_P\left(\frac{1}{\sqrt{n}}\right),\ \text{ and }\ \Pi_{12}=O_{a.s}\left(\frac{\ln\ln n}{n}\right)$$

- **Negligibility of** Π_{11}

  - **Negligibility of** Π_{111}.

Boundedness assumption on the derivative of the kernel imply that $\|A_i\|\leq 2C$ a.s. We apply Hoeffding inequality for independent, centered, bounded by M, but non identically distributed random variables (η_j) (e.g. see [2]),

$$\mathbb{P}\left(\sum_{j=1}^{n}\eta_j>t\right)\leq exp\left(\frac{-t^2}{2nM^2}\right)$$

Here, for every $\epsilon>0$, with $M=2C$, $\eta_j=\|A_i\|-\mathbb{E}\|A_i\|$, $t=\epsilon\sqrt{\frac{1}{n}\ln\ln n}$, therefore,

$$\sum_{i=1}^{n}(\|A_i\|-\mathbb{E}\|A_i\|)=O_p(\sqrt{n\ln\ln n})$$

which is the definition of almost complete convergence (a.co.), see e.g. [12] definition A.3. p. 230. In turn, it means that

$$\sum_{i=1}^{n}(\|A_i\|-\mathbb{E}\|A_i\|)=O_{a.co}\left(\sqrt{n\ln n}\right)$$

and by the Borell-Cantelli lemma,

$$\sum_{i=1}^{n}(\|A_i\|-\mathbb{E}\|A_i\|)=O_{a.s}(\sqrt{n\ln n})$$

Therefore, using equations equations (4.3) and (4.4), we have that

$$\Pi_{111}=O_P\left(\frac{\sqrt{\ln\ln n}}{na_n^3}\right)=O_{a.s}\left(\sqrt{\ln n}\frac{\sqrt{\ln\ln n}}{na_n^3}\right)$$

- **Negligibility of** Π_{112}

The r.h.s. of the previous inequality is, after an integration by parts, of order a3 n by the results on the kernel estimator of the gradient of the density (See Scott [26] chapter 6). Therefore,

$$\sum_{i=1}^{n}\mathbb{E}\|A_i\|=O(na_n^2)$$

$$\Pi_{112} = \frac{\|nZ_{i,n}\|}{na_n^3}\sum_{i=1}^{n}\mathbb{E}\|A_i\| = O_P\left(\frac{1}{\sqrt{n}}\right) = O_{a.s}\left(\frac{\sqrt{\ln\ln n}}{n}\right)$$

by equations (4.3) and (4.4).

Recollecting all elements, we eventually obtain that

$$\begin{aligned}\Pi = \Pi_{111} + \Pi_{112} + \Pi_{12} + \Pi_2 &= O_P\left(\frac{1}{\sqrt{n}}\right) + O_P\left(\frac{\ln\ln n}{na_n^3}\right) + O_P\left(\frac{1}{na_n^4}\right)\\ &= O_{a.s}\left(\sqrt{\frac{\ln\ln n}{n}}\right) + O_{a.s}\left(\frac{\sqrt{\ln n}\sqrt{\ln\ln n}}{na_n^3}\right) + O_{a.s}\left(\frac{\ln\ln n}{na_n^4}\right).\end{aligned}$$

By this last step we conclude the proof of our theorem.

□

After giving the proof of the convergence in probability, let us present the rate of convergence in the following corollary.

Corollary 4.1. *we get the rate of convergence, by choosing the bandwidth which balance the bias and variance trade-off: for an optimal choice of $h_n \simeq n^{-1/6}$, we get*

$$\widehat{r}_n(x) = r(x) + O_P(n^{-1/3})$$

Therefore, our estimator is rate optimal in the sense that it reaches the minimax rate $n^{-1/3}$ of convergence.

Now, Almost sure results can be proved in the same way: we have the following strong consistency result,

Theorem 4.2. *Let the regularity assumptions (i)-(vii) on the densitie and the kernel be satisfied.If the bandwidth h_n tends to zero as $n \to \infty$ in such a way that*

$$\frac{\sqrt{\ln n \ln\ln n}}{nh_n^3} \to 0, \frac{\ln\ln n}{nh_n^4} \to 0,$$

then,

$$\widehat{r}_n(x) = r(x) + O_{a.s}\left(h_n^2 + \sqrt{\frac{\ln\ln n}{nh_n^2}} + \frac{\ln\ln n}{nh_n^4} + \frac{\sqrt{\ln n\ln\ln n}}{nh_n^3}\right)$$

For the proof of this theorem, It is sufficient to follow the same lines as the preceding theorem , but uses the a.s. results of the consistency of the kernel density estimators of lemmas 3.13 and 3.15 and of the approximation propositions 3.16 and 3.17. It is therefore similar and omitted [7].

Corollary 4.2. *For $h_n \simeq (\ln\ln n/n)^{1/6}$ which is the optimal trade-off between the bias and the stochastic term, one gets the optimal rate*

$$\widehat{r}_n(x) = r(x) + O_{a.s}\left(\frac{\ln\ln n}{n}\right)^{1/3}.$$

For the he proof, we follow the same way given in in [7]

5. Conclusion

In this paper we established the asymptotic properties of a regression model via copula function approach, it will be interesting for further work to study the asymptotic normality of such model, to investigate the recursive estimation, it is also important to study the asymptotic properties of a conditional copula model.

References

[1] Biau,G., AND Wegkamp,M.H.," A note on minimum distance estimation of copulas densities"., *Statist. Probab. Lett.*, No.73, (2006), pp. 105-114.

[2] Bosq, D., *Nonparametric statistics for stochastic processes*, second ed., Vol. 110 of Lecture Notes in Statistics. Springer-Verlag, New York, (1998). Estimation and prediction.

[3] Cherubini, U., Luciano, E., AND Vecchiato, W., *Copula Methods in Finance*. John Wiley and Sons, Chichester (2004).

[4] Deheuvels, P., "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance". *Acad. Roy. Belg. Bull. Cl. Sci.*, Vol.65,No.5 And 6, (1979),pp. 274-292.

[5] Deheuvels, P., "A Kolmogorov-Smirnov type test for independence and multivariate samples", *Rev. Roumaine Math. Pures Appl.*Vol 26,No.2, (1981), pp. 213-226.

[6] Embrechts, P., Lindskog, F., McNeil, A., *Modelling dependence with copulas and applications to risk management.*, Handbook of Heavey Tailed Distributions in Finance, S.T. Rachev (ed.). Elsevier, Amsterdam (2003).

[7] Faugeras, Olivier Paul., *Contributions à la prévision statistique*, Thèses de doctorat de l'université Pierre er Marie Curie, 28 Novembre 2008.

[8] Fermanian, J.D., "Goodness-of-fit tests for copulas.", *J. Multivariate Anal.* Vol.95, No.1, (2005), pp. 119-152.

[9] Fermanian, J.D., Radulovïc, D., AND Wegkamp, M., "Weak convergence of empirical copula processes". Bernoulli Vol.10,No.5, (2004), pp. 847-860.

[10] Fermanian, J.D., AND Scaillet, O., *Some statistical pitfalls in copula modelling for financical applications*, E. Klein (Ed.), Capital formation, Gouvernance and Banking. Nova Science Publishing, New York, (2005).

[11] Fermanian, J.D.,AND Scaillet, O., "Nonparametric estimation of copulas for time series", *Journal of Risk*, Vol.5,No. 4, (2003), pp. 25-54.

[12] Ferraty, F., AND Vieu, P., *Nonparametric functional data analysis.* Springer Series in Statistics. Springer, New York, (2006). Theory and practice.

[13] Fisher, N.I., *Copulas*, Encyclopedia of Statistical Sciences, Updated Volume 1, S. Kotz, C. B. Read, D. L. Banks (eds.), pp. 159164. John Wiley and Sons, New York (1997).

[14] Gijbels, I., AND Mielniczuk, J., "Estimating the density of a copula function", *Comm. Statist. Theory Methods* ,Vol.19,No.2 (1990), pp.445-464.

[15] Hutchinson, T.P., Lai, C.D., *Continuous Bivariate Distributions: Emphasising Applications*,Rumsby Scientific Publishing, Adelaide (1990).

[16] Joe, H., *Multivariate Models and Dependence Concepts*, Chapman and Hall, London (1997).

[17] Joe, H., *Multivariate models and dependence concepts*, vol. 73 of Monographs on Statistics and Applied Probability. Chapman & Hall, London, (1997).

[18] Mikosch, T., "Copulas: Taies and facts", *Extrmes*,Vol. 9,No.18,(2006), pp.3-22 .

[19] Nelsen, R.B., *An Introduction to Copulas*, 2nd edition. Springer-Verlag, New York (2006).

[20] Nelsen, R.B., *An Introduction to Copulas*, Springer-Verlag, New York (1999).

[21] Parzen, E., "On estimation of a probability density function and mode", *Ann. Math. Statist.33*, (1962) , pp.1065-1076.

[22] Pickands, III, J., "Statistical inference using extreme order statistics", *Ann. Statist.3*,(1975), pp.119-131.

[23] Rüschendorf, L.," Asymptotic distributions of multivariate rank order statistics", *Ann. Statist. 4*, 5 (1976), pp.912-923.

[24] Sancetta,A., "Nonparametric estimation of multivariate distributions with given marginals: L_2 theory",*Cambridge Working papers in Economics* No.0320, (2003).

[25] Sancetta, A., Satchell,S., "The Barnestien coplua and its application to modelling and approximation of multivariate distributions", *Econometric theory*, 20, (2004), pp.535-562.

[26] Scott, D. W., *Multivariate density estimation. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley & Sons Inc., New York, (1992). Theory, practice, and visualization, A Wiley-Interscience Publication.

[27] Shorack, G. R., AND Wellner, J. A., *Empirical processes with applications tostatistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons Inc., New York, (1986).

[28] Sklar, M., "Fonctions de répartition à n dimensions et leurs marges", *Publ. Inst. Statist. Univ. Paris 8*, (1959), pp.229-231.

[29] Van Der Vaart, A. W., AND Wellner, J. A., *Weak convergence and empirical processes*, Springer Series in Statistics. Springer-Verlag, New York, (1996). With applications to statistics.

Stochastic Models, Statistics and Applications Moulay Tahar University, Saida P.O.Box. 138 En-Nasr Saida 20 000 Algeria.

E-mail address: dbennafla@yahoo.fr

Department of Mathematics, Djillali Liabes University, Sidi Bel Abbes, P.O.Box. 89, Sidi Bel Abbes 22000, Algeria.

E-mail address: rabhi_abbes@yahoo.fr

E-mail address: bouchentouf_amina@yahoo.fr