New Trends in Mathematical Sciences
http://dx.doi.org/10.20852/ntmsci.2016115601

On the recursive estimation using copula function in the regression model

Djamila Bennafla[1], *Amina Bouchentouf*[2], *Abbes Rabhi*[2] *and Khadidja Sabri*[3]

[1]Stochastic Models, Statistics and Applications Laboratory, University of Saida, Algeria
[2]Laboratory of Mathematics, University of Sidi Bel Abbes, Algeria
[3]Departement of Mathematics, University of Es-Senia Oran, Algeria

Received: 13 February 2015, Revised: 19 November 2015, Accepted: 24 December 2015
Published online: 1 January 2016.

Abstract: The main aim of this paper is to study the recursive estimation of the regression model by the transformed copula, giving its asymptotic properties to improve the performance of predictors nonparametric kernel, reducing their time calculations by using recursive kernels.

Keywords: Copulas, nonparametric estimation, recursive estimation, regression model.

1 Introduction

The craze to the theory of copulas has grown over the past three decades, an ample progress especially with applications in the field of finance: multiple credit risk assessment of structured credit products, replicating the performance of hedge funds, risk measurement multiple market management wallet. Indeed, risk management, evaluation of asset returns, the extreme value theory require modeling of dependence and copula theory is more attractive to finance, it allows to accommodate non-normality of the variables. The copulas are all-powerful mathematical tools to better understand the joint behavior of the markets in which we invest.

The concept of copula was introduced by Abe Sklar in 1959 as a solution to a problem probability statement by Maurice Fréchet in the context of random metric spaces (works with Berthold Schweizer). This concept has long been used very little in statistics; nevertheless, there may be mentioned: Work Kimeldorf and Sampson dependence (1975) and Research Paul Deheuvels (late 70s). The systematic study of copulas begins in the mid-1980s with Christian Genest and his team. The seminal article is: Genest & MacKay (1986). The joy of copulas: Bivariate distributions with uniform marginals. The American Statistician, 40, 280-283. Since then, many statistical developments were led by Genest *et al.*

The main idea of the theory of copulas is to separate the joint distributions, margins and their dependency structures. These functions are simply functions distribution dimension any whose marginal distributions are uniform on $[0,1]$.

Nonparametric estimators of copula densities have been suggested by Gijbels and Mielnicsuk [14] and Fermanian and Scaillet [11], who used kernel methods, Sancetta [21] and Sancetta and Satchell [22], who used techniques based on Bernstein polynomials. Biau and Wegkamp[3] proposed estimating the copula density through a minimum distance criterion. Faugeras [8] in his thesis studied the quantile copula approach to conditional density estimation. Recently

* Corresponding author e-mail: rabhi_abbes@yahoo.fr

Bennafla *et al.*[2] studied study the convergence Almost surely and in Probability (with rate) of regression model via copula function approach.

The objective of this work is devoted to the recursive estimation of the regression model by the transformed copula; At first, we introduce the model, and then we build our recursive estimators, after that we make some notations and assumptions of regularity for our main result contained in the last part of this article.

2 The model

Let $((X_i;Y_i),\ i=1,\dots,n)$ 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.

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

$$f(y|x)=\frac{f_{XY}(x,y)}{f_X(x)},\ f_X(x)\neq 0, \tag{1}$$

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

By introducing Parzen-Rosenblatt [18, 19] 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 $K_h(.)=1/hK(\cdot/h)$ and $K'_{h'}(\cdot)=1/h'K'(\cdot/h')$ 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.

Sklar's theorem below elucidates the role that copulas play in the relationship between bivariate distribution functions and their univariate marginals see Sklar[25].

Theorem 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*

$$F_{X,Y}(x,y)=C(F(x),G(y)),\ -\infty\leq x,y\leq+\infty. \tag{2}$$

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 marginal.

This theorem gives a representation of the bivariate c.d.f. as a function of each univariate c.d.f. That is to say, 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 marginal separately.

Copulas appear to be naturally linked with the quantile transform: in the case F and G are continuous, formula (2) is simply obtained by defining the copula function as $C(u,v) = F_{X,Y}(F^{-1}(u), G^{-1}(v)),\ 0 \le u \le 1,\ 0 \le v \le 1$. For more details regarding copulas and their properties, one can consult for example the book of Joe [15]. Copulas have endorsed a revived importance in statistics, especially in finance, since the pioneering work of Räuschendorf [20] and Deheuvels [6], who introduced the empirical copula process. Weak convergence of the empirical copula process was investigated by Deheuvels [7], Van der Vaart and Wellner [26], Fermanian, Radulovic and Wegkamp [10]. For the estimation of the copula density, refer to Gijbels and Mielniczuk [14], Fermanian [9] and Fermanian and Scaillet [12].

Here in after, 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), we get for the joint density,

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

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

$$f(y|x) = \frac{f_{XY}(x,y)}{f(x)} = g(y)c(F(x),G(y)),\ \ f(x) \neq 0. \tag{3}$$

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,

$$c_n(u,v) = \frac{1}{nh_n b_n} \sum_{i=1}^{n} K\left(\frac{u-U_i}{h_n}, \frac{v-V_i}{b_n}\right), \tag{4}$$

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 $(U_i,V_i),\ i = 1,\dots,n$ by its empirical counterpart $(F_n(X_i), G_n(Y_i))$, $i = 1,\dots,n$. We therefore obtain a genuine estimator of $c(u,v)$.

$$\widehat{c}_n(u,v) = \frac{1}{nh_n^2} \sum_{i=1}^{n} K_1\left(\frac{u-F_n(X_i)}{h_n}\right) K_2\left(\frac{v-G_n(Y_i)}{h_n}\right). \tag{5}$$

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

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

Our estimated model is given as follows: the regression function $r(x)$ is estimated by the function $\widehat{r}_n(x)$

$$\begin{aligned} r(x) =& \mathbb{E}(Y|X=x) = \int yf(y|x)dy = \int yg(y)c(F(x),G(y))dy \\ =& \mathbb{E}(Yc(F(x),G(y))). \end{aligned} \tag{6}$$

This regression function $r(x)$ is estimated by a function $\widehat{r}_n(x) = \int y\widehat{f}_n(y/x)dy$, thus, we obtain

$$\widehat{r}_n(x) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i \widehat{c}_n(F_n(x),G_n(y))\right) = \mathbb{E}\left(Y\widehat{c}_n(F_n(x),G_n(y))\right).$$

For more detail see [8]. To state our result, we have to make some regularity assumptions on the kernels and the densities.

3 Construction estimators recursive

We propose in this article to study the parametric family of estimators defined by recursive kernel

$$\widehat{r}^{l}_{n,R}(x) = \mathbb{E}\left(Y\widehat{c}^{l}_{n,R}(F_n(x),G_n(y))\right)$$

with for $l \in (0,1)$

$$\widehat{c}^{l}_{n,R}(F_n(x),G_n(y)) = \frac{1}{\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} K_1\left(\frac{F_n(x)-F_n(X_i)}{h_i}\right) K_2\left(\frac{G_n(y)-G_n(Y_i)}{h_i}\right).$$

Recall that $c^{l}_{n,R}$ and $\widehat{c}^{l}_{n,R}$ are density estimators of the copula c, respectively, based on pseudo-data unobservable $(F(X_i),G(Y_i))$, and their approximations $(F_n(X_i),G_n(Y_i))$,

$$c^{l}_{n,R}(u,v) = \frac{n}{\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} K_1\left(\frac{u-U_i}{h_i}\right) K_2\left(\frac{v-V_i}{h_i}\right),$$

$$\widehat{c}^{l}_{n,R}(u,v) = \frac{n}{\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} K_1\left(\frac{u-F_n(X_i)}{h_i}\right) K_2\left(\frac{v-G_n(Y_i)}{h_i}\right),$$

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

$$F_n(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{X_i \le x}$$

and

$$G_n(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{Y_i \le y}.$$

Our families estimators which can be calculated recursively by

$$\begin{aligned}\widehat{c}^{\,l}_{n+1,R}(F_n(x),G_n(y)) =& \frac{n+1\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2}{n\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}\widehat{c}^{\,l}_{n,R}(F_n(x),G_n(y))\\ &+\frac{n+1}{h_{n+1}^{2l}\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}K_1\left(\frac{F_{n+1}(x)-F_{n+1}(X_{n+1})}{h_{n+1}}\right)K_2\left(\frac{G_{n+1}(y)-G_{n+1}(Y_{n+1})}{h_{n+1}}\right),\end{aligned}$$

$$\begin{aligned}c^l_{n+1,R}(u,v) =& \frac{n+1\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2}{n\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}\widehat{c}^{\,l}_{n,R}(u,v)\\ &+\frac{n+1}{h_{n+1}^{2l}\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}K_1\left(\frac{u-U_{n+1}}{h_{n+1}}\right)K_2\left(\frac{v-V_{n+1}}{h_{n+1}}\right),\end{aligned}$$

$$\begin{aligned}\widehat{c}^{\,l}_{n+1,R}(u,v) =& \frac{n+1\left(\sum_{i=1}^{n} h_i^{1-l}\right)^2}{n\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}\widehat{c}^{\,l}_{n,R}(u,v)\\ &+\frac{n+1}{h_{n+1}^{2l}\left(\sum_{i=1}^{n+1} h_i^{1-l}\right)^2}K_1\left(\frac{u-F_{n+1}(X_n+1)}{h_{n+1}}\right)K_2\left(\frac{v-G_{n+1}(Y_n+1)}{h_{n+1}}\right).\end{aligned}$$

For estimators asymptotically unbiased and full equal to 1, we normalize by the quantity

$$B_{n,(1-l)} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^{1-l}.$$

Thus, we obtain:

$$\widehat{c}^{\,l}_{n,R}(F_n(x),G_n(y)) = \frac{B^{-2}_{n,(1-l)}}{nh_n^{2(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}K_1\left(\frac{F_n(x)-F_n(X_i)}{h_i}\right)K_2\left(\frac{G_n(y)-G_n(Y_i)}{h_i}\right),$$

$$c^l_{n,R}(u,v) = \frac{B^{-2}_{n,(1-l)}}{nh_n^{2(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}K_1\left(\frac{u-U_i}{h_i}\right)K_2\left(\frac{v-V_i}{h_i}\right)$$

and

$$\widehat{c}^{l}_{n,R}(u,v)=\frac{B^{-2}_{n,(1-l)}}{nh_n^{2(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}K_1\left(\frac{u-F_n(X_i)}{h_i}\right)K_2\left(\frac{v-G_n(Y_i)}{h_i}\right).$$

4 Notations and assumptions

We note the $i-th$ moment of a generic kernel (possibly multivariate) K as

$$m_i(K)=\int u^iK(u)du$$

and the L_p norm of a function h by $\|s\|_p=\int s^p$. We use the sign $\simeq$ to denote the order of the bandwidths. Set (u,v) fixed point in the interior of $supp(c)$. The support of the density function c is noted by $supp(c)=\overline{\{(u,v)\in\mathbb{R}^2;c(u,v)>0\}}$ where $\overline{A}$ stands for the closure of a set A. Finally, $\mathcal{O}_{\mathbb{P}}(\cdot)$ and $o_{\mathbb{P}}(\cdot)$ (respectively $\mathcal{O}_{a.s}(\cdot)$ and $o_{a.s}(\cdot)$) will stands for convergence and boundedness in probability (respectively almost surely).

Assumption (A)

(i) F of functions X and G of Y are strictly increasing and differentiable.
(ii) Density c is twice continuously differentiable with bounded second derivatives on its support.
(iii) Density c is uniformly continuous and non-zero almost everywhere on a compact set $D\subset(0,1)\times(0.1)$ included inside the bracket of c.

Assumption (B)

(i) The kernels K_i are bounded and bounded variation, for $i=\{1,2\}$.
(ii) $0<K_i<\alpha$ for a constant α; for $i=\{1,2\}$.
(iii) The kernels K_i are second order: $m_0(K_i)=\int K_i(x)dx=1$, $m_1(K_i)=\int xK_i(x)dx=0$, $m_2(K_i)=\int x^2K_i(x)dx<+\infty$, for $i=\{1,2\}$.
(iv) K_i is twice differentiable with second partial derivatives bounded, for $i=\{1,2\}$.

Assumption (C)

(i) $h_n\to 0$ and $nh_n^3\to\infty$, as n goes infinity.
(ii) $\forall r\leq 3$, $B_{n,r}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^r\to\beta_r<\infty$.

5 Main results

In this section, we establish the almost sure convergence, convergence in mean square and asymptotic normality of the family of recursive regression estimators $\widehat{r}^{l}_{n,R}(x)$.

Theorem 2. *Let the regularity assumptions (A), (B) et (C) be satisfied and if the bandwidth h_n tends to zero as $n\to\infty$ in such a way that*

$$h_n^{l-3}\sqrt{\ln n\ln\ln n}\to 0,\ h_n^{l-3}\sqrt{\frac{\ln\ln n}{n}}\to 0\ and\ \frac{\ln\ln n}{nh_n^{2(2-l)}}\to 0,$$

then

$$\begin{aligned}\widehat{r}_n(x) =& r(x) + \mathscr{O}_{a.s}\left(\frac{\beta_{(1-l)}^{-2}\beta_{(-2l)}}{h_n^{3-l}}\sqrt{\ln n \ln\ln n}\right) + \mathscr{O}_{a.s}\left(\frac{\beta_{(1-l)}^{-2}\beta_{(-2l)}}{h_n^{3-l}}\sqrt{\frac{\ln\ln n}{n}}\right)\\ &+\mathscr{O}_{a.s}\left(\frac{\beta_{(1-l)}^{-2}\beta_{(-2l)}\ln\ln n}{nh_n^{2(2-l)}}\right).\end{aligned}$$

Proof. Let

$$r(x) = \mathbb{E}\left(YC(F(x), G(y))\right)$$

and

$$\widehat{r}_{n,R}^{l}(x) = \mathbb{E}\left(Y\widehat{c}_{n,R}^{l}(F_n(x), G_n(x))\right).$$

The main ingredient to prove that $\widehat{r}(x)$ converges into $r(x)$ results in the following decomposition:

$$\begin{aligned}\widehat{r}_{n,R}^{l}(x) - r(x) =& \mathbb{E}\left(Y\widehat{c}_{n,R}^{l}(F_n(x), G_n(y)) - Yc(F(x), G(y))\right)\\ =& \mathbb{E}\Big(Y\Big[\widehat{c}_{n,R}^{l}(F_n(x), G_n(y)) - c_{n,R}^{l}(F(x), G(y))\\ &+c_{n,R}^{l}(F(x), G(y)) - c(F(x), G(y))\Big]\Big).\end{aligned}$$

Then, it is sufficient to prove that $\widehat{c}_{n,R}^{l}(U,V)$ converges to $c_{n,R}^{l}(U,V)$, and $c_{n,R}^{l}(U,V)$ converges to $c(U,V)$ with $U = F(x)$, $V = G(y)$.

Now, le's recall a preliminary result that will be needed; the convergence of Kolmogorov-Smirnov statistic:

For $(X_i,\ i = 1, 2, \ldots, n)$ an i.i.d. sample of a real random variable X with common *cdf* 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. [24] and [25] for recent accounts). For our purpose, we only need to formulate these results in the following rough form:

Lemma 1. *For an i.i.d. sample from a continuous* cdf F,

$$\|F_n - F\|_\infty = \mathscr{O}_{a.s}\left(\frac{\ln\ln n}{n}\right)\quad i = 1, 2, \ldots, n. \tag{7}$$

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,

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

Let

$$c_{n,R}^{l}(u,v) = \frac{B_{n,(1-l)}^{-2}}{nh_n^{2(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}K_1\left(\frac{u - F(X_i)}{h_i}\right)K_2\left(\frac{v - G(Y_i)}{h_i}\right),$$

and

$$\widehat{c}_{n,R}^{l}(u,v) = \frac{B_{n,(1-l)}^{-2}}{nh_n^{2(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}K_1\left(\frac{u - F_n(X_i)}{h_i}\right)K_2\left(\frac{v - G_n(Y_i)}{h_i}\right).$$

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

$$\widehat{c}^{l}_{n,R}(u,v) - c^{l}_{n,R}(u,v) = \frac{B^{-2}_{n,(1-l)}}{nh_n^{2(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} \Gamma_{i,n},$$

with

$$\Gamma_{i,n} = K_1\left(\frac{u-F_n(X_i)}{h_i}\right)K_2\left(\frac{v-G_n(Y_i)}{h_i}\right) - K_1\left(\frac{u-F(X_i)}{h_i}\right)K_2\left(\frac{v-G(Y_i)}{h_i}\right).$$

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

$$\|Z_{i,n}\| = \mathscr{O}_{a.s}\left(\sqrt{\frac{lnlnn}{n}}\right) \quad i = 1, \ldots, n. \tag{8}$$

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,

$$\begin{aligned}\Gamma =& \frac{B^{-2}_{n,(1-l)}}{nh_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} Z^T_{i,n} \nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right) + \frac{B^{-2}_{n,(1-l)}}{nh_n^{4(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} Z^T_{i,n} \nabla^2 K\left(\frac{u-\tilde{U}_{i,n}}{h_i}, \frac{v-\tilde{V}_{i,n}}{h_i}\right) Z_{i,n} \\ =& \Gamma_1 + \Gamma_2\end{aligned}$$

where $Z^T_{i,n}$ 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 Γ_1 as,

$$\begin{aligned}\Gamma_1 &= \frac{B^{-2}_{n,(1-l)}}{nh_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} Z^T_{i,n} \left(\nabla K\left(\frac{u-F(X_i)}{h_i}, \ldots\right) - \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_i}, \ldots\right)\right) \\ &+ \frac{B^{-2}_{n,(1-l)}}{nh_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} Z^T_{i,n} \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right) \\ &= \Gamma_{11} + \Gamma_{12}\end{aligned}$$

We again decompose one step further Γ_{11}, set

$$A_i = \nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right) - \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right).$$

Then

$$\begin{aligned}|\Gamma_{11}| &\leq \frac{B^{-2}_{n,(1-l)}\|Z_{i,n}\|}{nh_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} (\|A_i\| - \mathbb{E}\|A_i\|) + \frac{B^{-2}_{n,(1-l)}\|Z_{i,n}\|}{nh_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} \mathbb{E}\|A_i\| \\ &= \Gamma_{111} + \Gamma_{112}.\end{aligned}$$

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

To prove our results we often use the following lemma,

Lemma 2. *Let $(w_n)_{n\geq 1}$ be a sequence of real numbers tending to w. If Assumptions (H.2) are satisfied for all $r =]-\infty, 3]$ then*

$$\frac{1}{n}\sum_{i=1}^{n}\frac{h_i}{h_n}^{r} w_i \to \beta_r w, \text{ when } n \to \infty.$$

Lemma 2 is an immediate consequence of Lemma Toeplitz (see Appendix). Indeed, if we set,

$$a_{n,i} = \begin{cases} \frac{1}{n}\left(\frac{h_i}{h_n}\right)^r & if \quad i \leq n; \\ 0 & if \quad i > n, \end{cases}$$

then:

(i) for all $i > 1, a_{n,i} \leq \frac{h_i^r}{nh_n^r} \to 0$ as $n \to \infty$, through the assumption (H.2)-(i).
(ii) With the assumption (H.2)-(ii), we also have

$$\lim_{n\to\infty}\sum_{i=1}^{n} a_{n,i} = \lim_{n\to\infty} B_{n,r} = \beta_r < \infty.$$

(iii) There exists $C > 0$ such that for all $n > 1$, $\sum_{i=1}^{n}|a_{n,i}| < C < \infty$ through the convergence $B_{n,r}$.

• **Negligibility of Γ_2.**

$$\begin{aligned}\Gamma_2 &= \frac{B_{n,(1-l)}^{-2}}{nh_n^{4(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}} Z_{i,n}^T \nabla^2 K\left(\frac{u-\tilde{U}_{i,n}}{h_i}, \frac{v-\tilde{V}_{i,n}}{h_i}\right) Z_{i,n} \\ &= \frac{B_{n,(1-l)}^{-2}}{nh_n^{2(2-l)}}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^{-2l} Z_{i,n}^T \nabla^2 K\left(\frac{u-\tilde{U}_{i,n}}{h_i}, \frac{v-\tilde{V}_{i,n}}{h_i}\right) Z_{i,n}.\end{aligned}$$

By the boundedness assumption on the second-order derivatives of the kernel, lemma 2 and equation (8),

$$\Gamma_2 = \mathscr{O}_{a.s}\left(\frac{\beta_{(1-l)}^{-2}\beta_{(-2l)}\ln\ln n}{nh_n^{2(2-l)}}\right)$$

• **Negligibility of Γ_{12}.**

$$\begin{aligned}\Gamma_{12} &= \frac{B_{n,(1-l)}^{-2}}{nh_n^{3(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}} Z_{i,n}^T \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right) \\ &= \frac{B_{n,(1-l)}^{-2}}{nh_n^{3-l}}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^{2l} Z_{i,n}^T \mathbb{E}\nabla K\left(\frac{u-F(X_i)}{h_i}, \frac{v-G(Y_i)}{h_i}\right).\end{aligned}$$

Bias results on the bivariate gradient kernel estimator (See Scott [23] chapter 6) entail that

$$\mathbb{E}\nabla\left(K_1\left(\frac{u-F(X_i)}{h_i}\right)K_2\left(\frac{v-G(Y_i)}{h_i}\right)\right) = h_n^3\nabla c(u,v) + O(h_n^5).$$

Cauchy-Schwartz inequality yields that

$$|\Gamma_{12}| \leq \frac{nB^{-2}_{n,(1-l)}B_{-2l}\|Z_{i,n}\|}{nh_n^{3-l}} \left\| \mathbb{E}\nabla \left(K_1\left(\frac{u-F(X_i)}{h_i}\right) K_2\left(\frac{v-G(Y_i)}{h_i}\right)\right)\right\|.$$

In turn, with equation (8) and lemma 2

$$\Gamma_{12} = O_{a.s}\left(\frac{\beta^{-2}_{(1-l)}\beta_{(-2l)}}{h_n^{3-l}}\sqrt{\frac{\ln\ln n}{n}}\right).$$

• **Negligibility of Γ_{111}.**

$$\begin{aligned}\Gamma_{111} &= \frac{B^{-2}_{n,(1-l)}\|Z_{i,n}\|}{nh_n^{3(1-l)}}\sum_{i=1}^{n}\frac{1}{h_i^{2l}}\left(\|A_i\|-\mathbb{E}\|A_i\|\right)\\ &= \frac{B^{-2}_{n,(1-l)}\|Z_{i,n}\|}{nh_n^{3-l}}\sum_{i=1}^{n}\left(\frac{h_i}{h_n}\right)^{-2l}\left(\|A_i\|-\mathbb{E}\|A_i\|\right).\end{aligned}$$

Boundedness assumption on the derivative of the kernel imply that $\|A_i\| \leq 2\alpha$ a.s. We apply Hoeffding inequality for independent, centered, bounded by M, but non identically distributed random variables (η_j) (e.g. see [5]),

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

Here, for every $\varepsilon > 0$, with $M = 2\alpha$, $\eta_j = \|A_i\| - \mathbb{E}\|A_i\|$, $t = \varepsilon\sqrt{\frac{1}{n}\ln\ln n}$, therefore, which is the definition of almost complete convergence (a.co.), see e.g. [13] definition A.3. p. 230. In turn, it means that

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

and by the Borell-Cantelli lemma,

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

Therefore, by lemma 2 and using equation(8), we have that

$$\Gamma_{111} = O_{a.s}\left(\frac{\beta^{-2}_{(1-l)}\beta_{(-2l)}}{h_n^{3-l}}\sqrt{\ln n\ln(\ln n)}\right).$$

• **Negligibility of Γ_{112}.**

The right hand side of the previous inequality is, after an integration by parts, of order a_n^3 by the results on the kernel estimator of the gradient of the density (See Scott [23] chapter 6). Therefore,

$$\sum_{i=1}^{n}\mathbb{E}\|A_i\| = \mathscr{O}(nh_n^3)$$

and

$$\Gamma_{112} = \frac{B_{n,(1-l)}^{-2} \|Z_{i,n}\|}{n h_n^{3(1-l)}} \sum_{i=1}^{n} \frac{1}{h_i^{2l}} \mathbb{E}\|A_i\| = \frac{B_{n,(1-l)}^{-2} \|Z_{i,n}\|}{n h_n^{3-l}} \sum_{i=1}^{n} \left(\frac{h_i}{h_n}\right)^{2l} \mathbb{E}\|A_i\|.$$

Thus,

$$\Gamma_{112} = \mathscr{O}_{a.s} \left(\frac{\beta_{(1-l)}^{-2} \beta_{-2l}}{h_n^{3-l}} \sqrt{\frac{\ln(\ln n)}{n}} \right),$$

by equation (8) and lemma 2, recollecting all elements, we eventually obtain that

$$\begin{aligned} \Gamma =& \Gamma_{111} + \Gamma_{112} + \Gamma_{12} + \Gamma_2 \\ =& O_{a.s} \left(\frac{\beta_{(1-l)}^{-2} \beta_{(-2l)}}{h_n^{3-l}} \sqrt{\ln n \ln(\ln n)} \right) + \mathscr{O}_{a.s} \left(\frac{\beta_{(1-l)}^{-2} \beta_{(-2l)}}{h_n^{3-l}} \sqrt{\frac{\ln(\ln n)}{n}} \right) + \mathscr{O}_{a.s} \left(\frac{\beta_{(1-l)}^{-2} \beta_{(-2l)} \ln(\ln n)}{n h_n^{2(2-l)}} \right). \end{aligned}$$

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

It is now interested in the convergence in mean square $\widehat{r}_{n,R}^{l}(x)$. The following corollary establishes the asymptotic MSE exact family($\widehat{r}_{n,R}^{l}(x)$).

Corollary 1. *[1] Suppose that assumption of regularity are satisfied. the choice:*

$$h_n = M_n \left(\frac{\ln n}{n}\right)^{1/5}, M_n \downarrow m > 0,$$

implies that: For all $l \in [0,1]$

$$n^{4/5} \mathbb{E}\left[\widehat{r}_{n,R}^{l}(x) - r(x)\right] \to M_2(x, K, \varphi, f, l) \text{ when } n \to \infty,$$

with:

$$M_2(x, K, \varphi, f, l) = \left\{ \frac{m^2(4+l)[r(x) b_f(x) + b_\varphi(x)]}{f(x)(2+l)} \right\}^2 + \frac{(4+l)^2 \|K\|_2^2 V(x)}{10c(2+l) f(x)},$$

where

$$b_\varphi(x) = \frac{1}{2} \frac{\partial \varphi}{\partial x}(x) \int_{\mathbb{R}} \nu K(\nu) d\nu,$$

and

$$V(x) = \int K^2(t) dt \frac{\varphi(x) - r^2(x)}{f(x)}, \tag{9}$$

The following corollary establishes the asymptotic normality of $\widehat{r}_{n,R}^{l}(x)$

Corollary 2. *[1] Suppose that assumption of regularity are satisfied and that the sequence h_n is as:*

$$n h_n^5 \longrightarrow 0 \quad \text{when } n \to \infty,$$

and all sequences of integers u_n and v_n, we have: $u_n \sim v_n \Rightarrow h_{u_n} \sim h_{v_n}$. There is a real positive $\zeta_0 > 4$ as

$$\frac{n h_n}{(\ln n)^{\zeta_0}} \to +\infty \text{ when } n \to \infty,$$

then

$$\sqrt{nh_n}[\widetilde{r}_{n,R}^{l}(x) - r(x)] \hookrightarrow \mathcal{N}\left[0, \frac{\beta_{1-2l}\|K\|_2^2 V(x)}{\beta_{1-l}^2 f(x)}\right] \text{ when } n \to \infty,$$

where $V(x)$ is given in (9).

Appendix

The theorem due to Bochner on the limit of $(f * K_{h_n})$ is presented thereafter in the form of lemma.

Lemma 3. *[4] If f is integrable and if K is a kernel limited and integrable such that $|xK(x)| \longrightarrow 0$ when $x \to \infty$ and with integral 1, we have*

$$\lim_{h_n \to 0} (f * K_{h_n})(x) = f(x).$$

In a general way if g is continuous, we can write $K_{h_n} * g \underset{h_n \to 0}{\longrightarrow} g$ a.s.

The following Lemma says Toeplitz and recalled by Masry [16].

Lemma 4. *Let $(a_n)_{b \geq 1, k \geq 1}$ a real sequence and $(w_n)_{n \geq 1}$ a sequence which converge into w. Suppose that*

(i) For all $k \geq 1$, $\lim_{n \to \infty} a_{n,k} = 0$.

(ii) $\lim_{n \to \infty} \sum_{k=1}^{\infty} a_{n,k} = A < \infty$.

(iii) There exists a constant $C > 0$ such that for any $n \geq 1$, $\sum_{k=1}^{\infty} |a_{n,k}| < CA < \infty$.

Then we have

$$\sum_{k=1}^{\infty} a_{n,k} w_k \underset{n \to \infty}{\longrightarrow} Aw.$$

Lemma 5. *(Borel-Cantelli Lemma) [17] Let $(A_n)_{n \geq 1}$ a sequence of events. If $\sum_{i=1}^{\infty} \mathbb{P}(A_n) < \infty$ (resp. $= \infty$ and if A_n are independent), then*

$$\mathbb{P}\left(\overline{\lim_{n \to \infty}} A_n\right) = 0 \ (resp. \ = 1).$$

Recall that for a sequence of events A_n,

$$\overline{\lim_{n \to \infty}} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k = \lim_{n \to \infty} \downarrow \bigcup_{k \geq n} A_k.$$

References

[1] A. Amiri, *Rates of almost sure convergence of a family of recursive kernel estimators*. Annal de l'Institut de statistique de l'Université de Paris, ISUP., 54(3) (2010), 3-24.

[2] D. Bennafla, A. Rabhi and A. A. Bouchentouf, *Etimation using copula function in regression model*, Mathematical Sciences And Applications E-Notes, 2(1) (2014), 106-116.

[3] G. Biau and M. H. Wegkamp, *A note on minimum distance estimation of copulas densities*, Statist. Probab. Lett., 73 (2006), 105-114.

[4] D. Bosq and J. P. Lecoutre, *Théorie de l'estimation fonctionnelle*, ECONOMICA (eds), Paris, (1987).

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

[6] P. Deheuvels, *La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance*. Acad. Roy. Belg. Bull. Cl. Sci., 65(5-6) (1979), 274-292.

[7] P. Deheuvels, *A Kolmogorov-Smirnov type test for independence and multivariate samples*, Rev. Roumaine Math. Pures & Appl., 26(2) (1981), 213-226.

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

[9] J. D. Fermanian, *Goodness-of-fit tests for copulas*, J. Multivariate Anal., 95(1) (2005), 119-152.

[10] J. D. Fermanian, D. Radulovïc and M. Wegkamp, *Weak convergence of empirical copula processes*. Bernoulli, 10(5) (2004), 847-860.

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

[12] J. D. Fermanian and O. Scaillet, *Nonparametric estimation of copulas for time series*. Journal of Risk., 5(4) (2003), 25-54.

[13] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis: Theory and Practice*. Springer Series in Statistics, Springer, New York, 2006.

[14] I. Gijbels and J. Mielniczuk, *Estimating the density of a copula function*. Comm. Statist. Theory and Methods., 19(2) (1990), 445-464.

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

[16] E. Masry, *Recursive probability density estimation for weakly dependent stationary processes*, IEEE Trans. Inform. Theory., 32(2) (1986), 254-267.

[17] J. Neveu, *Bases mathématiques du calcul des probabilités*, 2nd. ed. Masson et Cie, Editeurs Paris, 1970.

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

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

[20] L. Rüschendorf, *Asymptotic distributions of multivariate rank order statistics*. Ann. Stat., 4(5) (1976), 912-923.

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

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

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

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

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

[26] A. W. Van Der Vaart and J. A. Wellner, *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996.