

2008/45



Asymptotic properties of the Bernstein density copula
for dependent data

Taoufik Bouezmarni, Jeroen V.K. Rombouts
and Abderrahim Taamouti

CORE

Voie du Roman Pays 34

B-1348 Louvain-la-Neuve, Belgium.

Tel (32 10) 47 43 04

Fax (32 10) 47 43 01

E-mail: corestat-library@uclouvain.be

<http://www.uclouvain.be/en-44508.html>

CORE DISCUSSION PAPER
2008/45

**Asymptotic properties of the Bernstein density copula
for dependent data**

Taoufik BOUEZMARNI¹, Jeroen V.K. ROMBOUTS²
and Abderrahim TAAMOUTI³

July 2008

Abstract

Copulas are extensively used for dependence modeling. In many cases the data does not reveal how the dependence can be modeled using a particular parametric copula. Nonparametric copulas do not share this problem since they are entirely data based. This paper proposes nonparametric estimation of the density copula for α -mixing data using Bernstein polynomials. We study the asymptotic properties of the Bernstein density copula, i.e., we provide the exact asymptotic bias and variance, we establish the uniform strong consistency and the asymptotic normality.

Keywords: nonparametric estimation, copula, Bernstein polynomial, α -mixing, asymptotic properties, boundary bias

JEL Classification: C13, C14

¹ Département de mathématiques et de statistique, Université de Montréal, Canada and Institute of Statistics, Université catholique de Louvain, Belgium.

² Institute of Applied Economics at HEC Montréal, CIRANO, CIRPEE,, Canada and CORE, Université catholique de Louvain, Belgium. E-mail: jeroen.rombouts@hec.ca

³ Economics Department, Universidad Carlos III de Madrid, Spain. E-mail: ataamout@eco.uc3m.es

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

The correlation coefficient of Pearson, the Kendall's tau, and Spearman's rho are widely used to measure the dependence between variables. Despite their simplicity to implement and interpret, they are not able to capture all forms of dependence. The copula function has the advantage to model completely the dependence among variables. Further, the above coefficients of dependence can be deduced from the copula. Thanks to Sklar (1959), the copula function is directly linked to the distribution function. Indeed, the distribution function can be controlled by marginal distributions, which give the information on each component, and the copula that captures the dependence between components.

There are several ways to estimate copulas. First, the parametric approach imposes a specific model for both the copula and marginal distributions. We estimate the parameters using the maximum likelihood or inference function for margins. This approach is widely used in practice because of its simplicity, see Oakes (1982) and Joe (2005) for details. Second, the semiparametric approach assumes a parametric model for the copula and a nonparametric model for the marginal distributions. This approach is studied by Oakes (1982), Genest and Rivest (1993), and Genest, Ghoudi, and Rivest (1995). These methods are not efficient since they involve two-step estimation. To overcome this problem, Chen, Fan, and Tsyrennikov (2006) use the sieve maximum likelihood estimation. Chen and Fan (2006) investigate the semiparametric approach for the estimation of copula in the context of dependent data. In order to reduce dimensionality and to remove the problem of boundary bias when the support of the variables is bounded, Bouezmarni and Rombouts (2008) propose a semiparametric estimation procedure for a multivariate density with a parametric copula and asymmetric kernels density estimators for the marginal densities. In this paper, we are interested in a third way of estimating copula functions, which is nonparametric estimation. This approach considers nonparametric models for both the copula and marginal distributions. Deheuvels (1979) suggests the multivariate empirical distribution to estimate the copula function. Gijbels and Mielniczuk (1990) estimate a bivariate copula using smoothing kernel methods. Also, they suggest the reflection method in order to solve the boundary bias problem of the kernel methods. Chen and Huang (2007) propose a bivariate estimator based on the local linear estimator, which is consistent everywhere in the support of the copula function. Rödel (1987) uses the orthogonal series method. For independent and identically distributed (*i.i.d.*) data, Sancetta and Satchell (2004) develop a Bernstein polynomial estimator of the copula function and find an upper bound of the asymptotic bias and variance and the asymptotic normality of the Bernstein density copula estimator.

In this paper, we consider α -mixing dependent data and propose nonparametric estimation of

the density copula based on the Bernstein polynomials. We study the asymptotic properties of the Bernstein density copula, i.e., we provide the exact asymptotic bias and variance, and we establish the uniform strong consistency and the asymptotic normality of Bernstein estimator for the density copula.

Motivated by Weierstrass theorem, Bernstein polynomials are considered by Lorentz (1953) who proves that any continuous function can be approximated by Bernstein polynomials. For density functions, estimation using the Bernstein polynomial is suggested by Vitale (1975) and with a slight modification by Grawronski and Stadtmüller (1981). Tenbusch (1994) investigates the Bernstein estimator for bivariate density functions and Bouezmarni and Rolin (2007) prove the consistency of Bernstein estimator for unbounded density functions. Kakizawa (2004) and Kakizawa (2006) consider the Bernstein polynomial to estimate density and spectral density functions, respectively. Tenbusch (1997) and Brown and Chen (1999) propose estimators of the regression functions based on the Bernstein polynomial. In the Bayesian context, Bernstein polynomials are explored by Petrone (1999a), Petrone (1999b), Petrone and Wasserman (2002), and Ghosal (2001).

This paper is organized as follows. The Bernstein copula estimator is introduced in Section 2. Section 3 provides the asymptotic properties of the Bernstein density copula estimator, that is the asymptotic bias and variance, the uniform strong consistency, and the asymptotic normality for α -mixing dependent data. Section 4 concludes.

2 Bernstein copula estimator

Let $X = \{(X_{i1}, \dots, X_{id})', i = 1, \dots, n\}$ be a sample of n observations of α -mixing vectors in \mathbb{R}^d , with distribution function F and density function f . A sequence is α -mixing of order h if the mixing coefficient $\alpha(h)$ goes to zero as the order h goes to infinite, where

$$\alpha(h) = \sup_{A \in \mathcal{F}_1^t(X), B \in \mathcal{F}_{t+h}^\infty(X)} |P(A \cap B) - P(A)P(B)|,$$

$\mathcal{F}_1^t(X)$ and $\mathcal{F}_{t+h}^\infty(X)$ are the σ -field of events generated by $\{X_l, l \leq t\}$ and $\{X_l, l \geq t+h\}$, respectively. The concept of α -mixing is omnipresent in time series analysis and is less restrictive than β and ρ -mixing. The following condition requires an α -mixing coefficient with exponential decay. We assume that the process X is α -mixing such that

$$\alpha(h) \leq \rho^h, \quad h \geq 1, \tag{1}$$

for some constant $0 < \rho < 1$.

According to Sklar (1959), the distribution function of X can be expressed via a copula:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \tag{2}$$

where F_j is the marginal distribution function of random variable $X^j = \{X_{1j}, \dots, X_{nj}\}$ and C is a copula function which captures the dependence in X . Deheuvels (1979) uses a nonparametric approach, based on the empirical distribution function, to estimate the distribution copula. Using Bernstein polynomials, to smooth the empirical distribution, Sancetta and Satchell (2004) propose the empirical Bernstein copula which is defined as follows: for $s = (s_1, \dots, s_d) \in [0, 1]^d$

$$\hat{C}(s_1, \dots, s_d) = \sum_{v_1=0}^{k-1} \dots \sum_{v_d=0}^{k-1} F_n \left(\frac{v_1}{k}, \dots, \frac{v_d}{k} \right) \prod_{j=1}^d p_{v_j}(s_j), \quad (3)$$

where k is an integer playing the role of the bandwidth parameter, F_n is the empirical distribution function of X , and $p_{v_j}(s_j)$ is the binomial distribution function:

$$p_{v_j}(s_j) = \binom{k-1}{v_j} s_j^{v_j} (1-s_j)^{k-v_j-1}.$$

If we derive (2) with respect to (x_1, \dots, x_d) , we obtain the density function, say $f(x_1, \dots, x_d)$, of X that can be expressed as follows:

$$f(x_1, \dots, x_d) = (f_1(x_1) \times \dots \times f_d(x_d)) \times c(F_1(x_1), \dots, F_d(x_d)),$$

where f_j is the marginal density of the random variable X^j and c is the copula density. Hence, the estimation of the joint density function can be done by estimating the univariate marginal densities and the copula density function. In this paper, we estimate the copula density function using Bernstein polynomials. Indeed, if we derive (3) with respect to (s_1, \dots, s_d) we obtain the Bernstein density copula:

$$\hat{c}(s_1, \dots, s_d) = \frac{1}{n} \sum_{i=1}^n K_k(s, S_i) \quad (4)$$

where $S_i = (F_1(X_{i1}), \dots, F_d(X_{id}))$,

$$K_k(s, S_i) = k^d \sum_{v_1=0}^{k-1} \dots \sum_{v_d=0}^{k-1} A_{S_i, v} \prod_{j=1}^d p_{v_j}(s_j),$$

and

$$A_{S_i, v} = \mathbf{1}_{\{S_i \in B_v\}}, \quad \text{with} \quad B_v = \left[\frac{v_1}{k}, \frac{v_1+1}{k} \right] \times \dots \times \left[\frac{v_d}{k}, \frac{v_d+1}{k} \right].$$

In what follows, we denote the multiple sums $\sum_{v_1=0}^{k-1} \dots \sum_{v_d=0}^{k-1}$ by \sum_v . The Bernstein estimator for the density copula function is simple to implement, non-negative, and integrates to one. Sancetta and Satchell (2004) give the upper bounds of the bias and variance of the Bernstein copula density estimator for *i.i.d* observations. In this paper, we provide the asymptotic properties of the Bernstein copula density for α -mixing dependent data. For such data, we give the exact asymptotic bias and variance, we prove the uniform almost sure (*a.s.*) convergence of the Bernstein density copula, and we establish its asymptotic normality.

3 Main results

This section studies the asymptotic properties of the Bernstein density copula estimator. We first show that the asymptotic bias of the Bernstein density copula estimator has a uniform rate of convergence. Hence, asymptotically there is no boundary bias problem. Second, we provide the exact asymptotic variance of the estimator in the interior region. Finally, we establish the uniform strong consistency and the asymptotic normality of the Bernstein density copula estimator. We start by studying the bias of the Bernstein density copula estimator. The following proposition gives the exact asymptotic bias. To stress that the bandwidth depends on n , we replace k by k_n .

Proposition 1 (Asymptotic Bias). *Suppose that the density copula function c is twice differentiable. Let \hat{c} be the Bernstein density copula estimator of c as defined in (4). Then, for $s = (s_1, \dots, s_d) \in (0, 1)^d$, if $k_n \rightarrow \infty$, we have*

$$E(\hat{c}(s)) = c(s) + k_n^{-1} \gamma^*(s) + o(k_n^{-1})$$

where

$$\gamma^*(s) = \frac{1}{2} \sum_{j=1}^d \left\{ \frac{dc(s)}{ds_j} (1 - 2s_j) + \frac{d^2c(s)}{ds_j^2} s_j (1 - s_j) \right\}.$$

Proof. Using a second order Taylor expansion and various sums we have

$$\begin{aligned} E(\hat{c}(s)) - c(s) &= k_n^d \sum_v \left\{ \int_{B_v} (c(u) - c(s)) du \right\} \prod_{j=1}^d p_{v_j}(s_j) \\ &= k_n^d \sum_v \left\{ \sum_{l=1}^d \frac{dc(s)}{du_l} \int_{B_v} (u_l - s_l) du \right\} \prod_{j=1}^d p_{v_j}(s_j) \\ &+ \frac{k_n^d}{2} \sum_v \left\{ \sum_{l \neq m}^d \frac{d^2c(s)}{du_l du_m} \int_{B_v} (u_l - s_l)(u_m - s_m) du \right\} \prod_{j=1}^d p_{v_j}(s_j) \\ &+ \frac{k_n^d}{2} \sum_v \left\{ \sum_{l=1}^d \frac{d^2c(s)}{du_l^2} \int_{B_v} (u_l - s_l)^2 du \right\} \prod_{j=1}^d p_{v_j}(s_j) + o(k_n^{-1}) \\ &= \frac{1}{2} k_n^{-1} \left\{ \sum_{j=1}^d \frac{dc}{du_j}(s) (1 - 2s_j) + \sum_{j=1}^d \frac{d^2c}{du_j^2}(s) s_j (1 - s_j) \right\} + o(k_n^{-1}). \end{aligned}$$

The last equality is obtained by using the mean and variance of the Binomial distribution and the fact that

$$\frac{k_n^d}{2} \sum_v \left\{ \sum_{l \neq m}^d \frac{d^2c(s)}{du_l du_m} \int_{B_v} (u_l - s_l)(u_m - s_m) du \right\} \prod_{j=1}^d p_{v_j}(s_j) = O(k_n^{-d}).$$

□

Now, we compute the variance of the Bernstein density copula estimator. This is given by the following proposition.

Proposition 2 (Asymptotic Variance). *Let \hat{c} be the Bernstein density copula estimator of c as defined in (4). Then, for $s \in (0, 1)^d$, under condition (1) and if $n^{-1}k_n^{-d/2} \rightarrow 0$, we have*

$$\text{Var}(\hat{c}(s)) = n^{-1}k_n^{-d/2}V(s) + o(n^{-1}k_n^{-k_n/2})$$

where $V(s) = (4\pi)^{-d/2} \frac{c(s)}{\prod_{j=1}^d (s_j(1-s_j))^{1/2}}$.

Note that the formula of the variance at $s = 0$ and $s = 1$ is given by Sancetta and Satchell (2004). We see that the variance increases with dimension d and it increases near the boundary because of the term $(s_j(1-s_j))^{1/2}$ in the denominator of $V(s)$.

Proof. We have

$$\text{Var}(\hat{c}(s)) = \frac{1}{n} \text{Var}(K_{k_n}(s, S_i)) + \frac{2}{n} \sum_{i=1}^{n-1} (1 - n^{-1}i) \text{Cov}((K_{k_n}(s, S_1), K_{k_n}(s, S_i))).$$

First, under condition (1) and using Billingsley's inequality, Lemma (3.1) in Bosq (1996) and that $\|K_{k_n}(s, S_i)\|_\infty = O(k_n^{d/2})$, we obtain

$$\frac{2}{n} \sum_{i=1}^{n-1} (1 - n^{-1}i) \text{Cov}((K_{k_n}(s, S_1), K_{k_n}(s, S_i))) = o(n^{-1}k_n^{d/2}).$$

Second,

$$\frac{1}{n} \text{Var}(K_{k_n}(s, S_i)) = \frac{k_n^{2d}}{n} \sum_v \text{Var}(A_{S_i, v}) \prod_{j=1}^d p_{v_j}^2(s_j)$$

From Sancetta and Satchell (2004)

$$\text{Var}(A_{S_i, v}) = \frac{c(\frac{v_1}{k_n}, \dots, \frac{v_d}{k_n})}{k_n^d} + o(k_n^{-d}).$$

Hence

$$\begin{aligned} \frac{1}{n} \text{Var}(K_{k_n}(s, S_i)) &= \frac{k_n^{2d}}{n} \sum_v \left(\frac{c(\frac{v_1}{k_n}, \dots, \frac{v_d}{k_n})}{k_n^d} + o(k_n^{-d}) \right) \prod_{j=1}^d p_{v_j}^2(s_j) \\ &= \frac{k_n^{2d}}{n} \left[\sum_{v \in I} (\cdot) + \sum_{v \in I^c} (\cdot) \right] \\ &\equiv V_1 + V_2 \end{aligned}$$

where

$$I = \left\{ v = (v_1, \dots, v_d); \left| \frac{v_j}{k_n} - s_j \right| < k_n^{-\delta}, j = 1, \dots, d \quad 1/3 < \delta < 1/2 \right\}.$$

Let's start with the second term V_2 . By considering the notations $m = \sup_s(c(s))$, and when $v \in I^c$ this means that there exists d_0 , for $1 \leq d_0 \leq d$, elements of v_j such that $\left| \frac{v_j}{k_n} - s_j \right| > k_n^{-\delta}$, we have

$$\begin{aligned} V_2 &= \sum_{v \in I^c} \left(\frac{c\left(\frac{v_1}{k_n}, \dots, \frac{v_d}{k_n}\right)}{k_n^d} + o(k_n^{-d}) \right) \prod_{j=1}^d p_{v_j}^2(s_j) \\ &\leq \frac{mk_n^d}{n} \left(\prod_{j=1}^{d_0} \sum_{\left| \frac{v_j}{k_n} - s_j \right| > k_n^{-\delta}} p_{v_j}^2(s_j) \right) \left(\prod_{j=d_0+1}^d \sum_{\left| \frac{v_j}{k_n} - s_j \right| < k_n^{-\delta}} p_{v_j}^2(s_j) \right) \\ &\leq \frac{m}{n} k_n^{d/2} \left(k_n^{-7/2 d_0} \right) = o(n^{-1} k_n^{d/2}). \end{aligned}$$

For the last inequality, on the one hand, we use

$$\begin{aligned} \sum_{\left| v_j/k_n - s_j \right| > k_n^{-\delta}} p_{v_j}^2(s_j) &\leq \left(\sum_{\left| v_j/k_n - s_j \right| > k_n^{-\delta}} p_{v_j}(s_j) \right)^2 \\ &= O(k_n^{-4}), \quad \text{from Lorentz (1953)}. \end{aligned}$$

On the other hand, from Laplace's formula we have

$$\frac{k_n^{\frac{1}{2}} p_{v_j}^2(s_j)}{P_{v_j}(s_j)} \rightarrow 1, \quad \text{as } k_n \rightarrow \infty$$

where

$$P_{v_j}(s_j) = \frac{k_n^{1/2}}{2\pi s_j(1-s_j)} \int_{\frac{j}{k_n}}^{\frac{j+1}{k_n}} \exp \left[-\frac{k_n}{s_j(1-s_j)} (t - s_j)^2 \right] dt.$$

Let $v_{j'}$ and $v_{j''}$ be the smallest and the biggest integers such that $\left| v_j/k_n - s_j \right| \leq k_n^{-\delta}$. Using the Laplace's formula and the change of variables we get

$$\begin{aligned} \sum_{\left| v_j/k_n - s_j \right| \leq k_n^{-\delta}} p_{v_j}^2(s_j) &\approx \frac{1}{2\pi s_j(1-s_j)} \int_{\frac{v_{j'}}{k_n}}^{\frac{v_{j''}}{k_n}} \exp \left[-\frac{k_n}{s_j(1-s_j)} (t - s_j)^2 \right] dt \\ &= \frac{k_n^{-1/2}}{2\sqrt{\pi s_j(1-s_j)}} \frac{1}{\sqrt{2\pi}} \int_{v_{j_1}}^{v_{j_2}} \exp(-y^2/2) dy \end{aligned}$$

where $v_{j_1} = \sqrt{\frac{2k_n}{s_j(1-s_j)}} \left(\frac{v_{j'}}{k_n} - s_j \right)$ and $v_{j_2} = \sqrt{\frac{2k_n}{s_j(1-s_j)}} \left(\frac{v_{j''}}{k_n} - s_j \right)$. Note that, when $k_n \rightarrow \infty$, then $j_1 \rightarrow -\infty$ and $j_2 \rightarrow +\infty$, because $s_j - v_{j'}/k_n \leq k_n^{-\delta}$, $v_{j''}/k_n - s_j \leq k_n^{-\delta}$, and $\delta < 1/2$. Thus,

$$\sum_{\left| v_j/k_n - s_j \right| \leq k_n^{-\delta}} p_{v_j}^2(s_j) = O(k_n^{-1/2}), \quad \text{as } k_n \rightarrow \infty.$$

Now, for V_1

$$\begin{aligned}
V_1 &= \frac{k_n^{2d}}{n} \sum_{v \in I} \left(\frac{c(\frac{v_1}{k_n}, \dots, \frac{v_d}{k_n})}{k_n^d} + o(k_n^{-d}) \right) \prod_{j=1}^d p_{v_j}^2(s_j) \\
&= \frac{k_n^d}{n} c(s) \left(\prod_{j=1}^d \sum_{v \in I} p_{v_j}^2(s_j) \right) + o(n^{-1}k_n^{d/2}), \quad \text{because } s_j \approx \frac{v_j}{k_n} \\
&= n^{-1}k_n^{d/2} (4\pi)^{-d/2} \frac{c(s)}{\prod_{j=1}^d (s_j(1-s_j))^{1/2}} + o(n^{-1}k_n^{d/2}).
\end{aligned}$$

□

For α -mixing dependent data, the uniform almost sure convergence of the Bernstein density copula estimator is stated in the following proposition.

Proposition 3 (Uniform a.s. Convergence). *Suppose that the density copula function c is twice differentiable and that $\{S_i\}$ is an α -mixing sequence with coefficient $\alpha(h) = O(\rho^h)$, for some $0 < \rho < 1$. Let \hat{c} be the Bernstein density copula estimator of c as defined in (4). Then, if $k_n \rightarrow \infty$ such that $n^{-1/2}k_n^{d/4} \ln(n) \rightarrow 0$, we have*

$$\sup_s |\hat{c}(s) - c(s)| = O(k_n^{-1} + n^{-1/2}k_n^{d/4} \ln(n)), \quad a.s.$$

Proof. From the bias term and under the assumption that c is twice differentiable, we have

$$\sup_s |E(\hat{c}(s)) - c(s)| = O(k_n^{-1}).$$

If we denote

$$Y_{n,i} = \frac{1}{n} K_{k_n}(s, S_i),$$

then we can show that

$$R_2(n) = \sup_i E|Y_{n,i}^2|^{1/2} = O(n^{-1}k_n^{k_n/4})$$

and $|Y_{n,i}| \leq n^{-1}k_n^{d/2}$. Hence, under the above conditions on the bandwidth parameter and applying Theorem (3.2) from Liebscher (1996) to $Y_{n,i}$, we get

$$\sup_s |E(\hat{c}(s)) - c(s)| = O(n^{-1/2}k_n^{d/4} \ln(n)).$$

□

The next proposition establishes the asymptotic normality of the Bernstein density copula estimator for α -mixing dependent data. This result can be applied in many contexts. We can use it for example to build copula-based tests of goodness-fit and conditional independence.

Proposition 4 (Asymptotic Normality). *Suppose that the density copula function c is twice differentiable and that $\{S_i\}$ is an α -mixing sequence with coefficient $\alpha(h) = O(\rho^h)$, for some $0 < \rho < 1$. Let \hat{c} be the Bernstein density copula estimator of c as defined in (4). Then, if $k_n \rightarrow \infty$ such that $k_n = O(n^{2/(4+d)})$, we have*

$$n^{1/2}k_n^{-d/4} \frac{\hat{c}(s) - c(c) - k_n^{-1} \gamma^*(s)}{\sqrt{V(s)}} \rightarrow N(0, 1).$$

Remark that if we choose $k_n = O(n^{2/(4+d)})$, then the bias term disappears.

Proof. Based on Proposition (1), we need to show that

$$n^{1/2}k_n^{-d/4} \left(\frac{\hat{c}(s) - E(\hat{c}(s))}{\sqrt{V(s)}} \right) \rightarrow N(0, 1), \text{ for } s \in (0, 1)^d. \quad (5)$$

If we denote

$$Y_i = \frac{K_{k_n}(s, S_i) - E(K_{k_n}(s, S_i))}{\sqrt{V(s)}}.$$

then

$$\left(n^{1/2}k_n^{-d/4} \frac{\hat{c}(s) - E(\hat{c}(s))}{\sqrt{V(s)}} \right) = n^{-1/2}k_n^{-d/4} \sum_{i=1}^n Y_i \equiv n^{-1/2} I_n.$$

We follow Doob's method to show the asymptotic normality for dependent random vectors, see Doob (1953). We consider the variables $V_i = k_n^{-d/4}(Y_{(i-1)(p+q)+1} + \dots + Y_{ip+(i-1)q})$ and $V_i^* = k_n^{-d/4}(Y_{ip+(i-1)q+1} + \dots + Y_{i(p+q)})$. For $r(p+q) \leq n \leq r(p+q+1)$,

$$I_n = \sum_{i=1}^r V_i + \sum_{i=1}^r V_i^* + k_n^{-d/4} \sum_{i=r(p+q)}^n Y_i. \quad (6)$$

We can show that

$$n^{-1/2} \left(\sum_{i=1}^r V_i^* + k_n^{-d/4} \sum_{i=r(p+q)}^n Y_i \right) \xrightarrow{P} 0.$$

Indeed, if we choose $r \sim n^a$, $p \sim n^{1-a}$, and $q \sim n^c$, where $0 < a < 1$ and $0 < c < 1 - a$, we get

$$n^{-1} \text{Var} \left(\sum_{i=1}^r V_i^* \right) = O(n^{a+c-1}), \quad \text{and} \quad n^{-1} k_n^{-d/2} \text{Var} \left(\sum_{i=r(p+q)}^n Y_i \right) = O(n^{a-1}).$$

The two last terms in the right side of (6) are asymptotically negligible.

Now, we show that V_i are asymptotically mutually independent. Let $U_i = \exp(itV_i)$ which is \mathcal{F}_i^j -measurable, where $i = (i-1)(p+q) + 1$ and $j = ip + (i-1)q$, hence from Volkonski and Razanov

(1959)

$$\left| E \left(\exp \left(it \sum_{i=1}^r V_i \right) \right) - \prod_{i=1}^r E(\exp(itV_i)) \right| \leq 16(r-1)\alpha(q+1).$$

Lastly, we employ the Lyapounov's theorem for the asymptotic normality of $n^{-1/2} \sum_{i=1}^n V_i$. If we choose $a > \frac{d+2}{d+4}$, we obtain,

$$\begin{aligned} \frac{\sum_{i=1}^r E(|V_i|^3)}{(r \operatorname{var}(V_1))^{3/2}} &\leq \|V_i\|_\infty (r \operatorname{var}(V_1))^{-1/2} \\ &\leq p k_n^{-d/4} \|K_{k_n}(s, t)\|_\infty (r \operatorname{var}(V_1))^{-1/2} \\ &= O(n^{\frac{d+2}{d+4}-a}) = o(1) \quad \text{because} \quad \|K_{k_n}(s, X_i)\|_\infty = O(k^{d/2}). \end{aligned}$$

□

4 Conclusion

A nonparametric Bernstein polynomial-based estimator of density copula for dependent data is provided. The proposed estimator can be applied in several contexts, and we can use it to build copula-based tests of, for example, goodness-fit and conditional independence, see Bouezmarni, Rombouts, and Taamouti (2008). We provide the exact asymptotic bias and variance of the Bernstein copula density estimator and we establish its uniform strong consistency, and the asymptotic normality. Our results can be extended to the right censored data using smoothed Kaplan-Meier estimator instead of the empirical distribution function. A bandwidth choice in practice remains an open question and existing methods like cross-validation can be investigated.

References

- BOSQ, D. (1996): *Nonparametric Statistics for Stochastic Processes, Estimation and Prediction*. Springer-Verlag, New York.
- BOUEZMARNI, T., AND J. ROLIN (2007): "Bernstein Estimator for Unbounded Density Function," *Journal of Nonparametric Statistics*, 19, 145–161.
- BOUEZMARNI, T., AND J. ROMBOUTS (2008): "Semiparametric Multivariate Density Estimation for Positive Data Using Copulas," *Computational Statistics and Data Analysis*, forthcoming.

- BOUEZMARNI, T., J. ROMBOUTS, AND T. TAAMOUTI (2008): “A Nonparametric Test for Conditional Independence Using Bernstein Density Copulas,” Mimeo.
- BROWN, B., AND S. CHEN (1999): “Beta-Bernstein Smoothing for Regression Curves with Compact Supports,” *Scandinavian Journal of Statistics*, 26, 47–59.
- CHEN, S. X., AND T. HUANG (2007): “Nonparametric Estimation of Copula Functions for Dependent Modeling,” *Canadian Journal of Statistics*, 35, 265–282.
- CHEN, X., AND Y. FAN (2006): “Estimation of Copula-based Semiparametric Time Series Models,” *Journal of Econometrics*, 130, 307–335.
- CHEN, X., Y. FAN, AND V. TSYRENNIKOV (2006): “Efficient Estimation of Semiparametric Multivariate Copula Models,” *Journal of the American Statistical Association*, 101, 1228–1240.
- DEHEUVELS, P. (1979): “La Fonction de Dépendance Empirique et ses Propriétés. Un Test non Paramétrique D’indépendance,” *Bulletin de l’académie Royal de Belgique, Classe des Sciences*, pp. 274–292.
- DOOB, J. (1953): *Stochastic Processes*. John Wiley and Sons, New York.
- GENEST, C., K. GHOUDI, AND L. RIVEST (1995): “Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions,” *Biometrika*, 82, 543–552.
- GENEST, C., AND L. RIVEST (1993): “Statistical Inference Procedure for Bivariate Archimedean Copulas,” *Journal of the American Statistical Association*, 88, 1034–1043.
- GHOSAL, S. (2001): “Convergence Rates for Density Estimation with Bernstein Polynomials,” *Annals of Statistics*, 28, 1264–1280.
- GIJBELS, I., AND J. MIELNICZUK (1990): “Estimating The Density of a Copula Function,” *Communications in Statistics - Theory and Methods*, 19, 445–464.
- GRAWRONSKI, N., AND U. STADTMÜLLER (1981): “Smoothing Histograms by Lattice and Continuous Distributions,” *Metrika*, 28, 155–164.
- JOE, H. (2005): “Asymptotic Efficiency of the Two-Stage Estimation Method for Copula-Based Models,” *Journal of Multivariate Analysis*, 94, 401–419.
- KAKIZAWA, Y. (2004): “Bernstein Polynomial Probability Density Estimation,” *Journal of Non-parametric Statistics*, 16, 709–729.

- (2006): “Bernstein Polynomial Estimation of a Spectral Density,” *Journal of Time Series Analysis*, 27, 253–287.
- LIEBSCHER, E. (1996): “Strong Convergence of Sums of α -mixing Random Variables with Applications to Density Estimation,” *Stochastic Processes and their Applications*, 65, 69–80.
- LORENTZ, G. (1953): *Bernstein Polynomials*. University of Toronto Press.
- OAKES, D. (1982): “A Model for Association in Bivariate Survival Data,” *Journal of the Royal Statistical Society, Series B*, 44, 414–422.
- PETRONE, S. (1999a): “Bayesian Density Estimation Using Bernstein Polynomials,” *Canadian Journal of Statistics*, 27, 105–126.
- (1999b): “Random Bernstein Polynomials,” *Scandinavian Journal of Statistics*, 26, 373–393.
- PETRONE, S., AND L. WASSERMAN (2002): “Consistency of Bernstein Polynomial Density Estimators,” *Journal of the Royal Statistical Society, Ser. B*, 64, 79–100.
- RÖDEL, E. (1987): “R-Estimation of Normed Bivariate Density Functions,” *Statistics*, 18, 573–585.
- SANCETTA, A., AND S. SACHELL (2004): “The Bernstein Copula and its Applications to Modeling and Approximations of Multivariate Distributions,” *Econometric Theory*, 20, 535–562.
- SKLAR, A. (1959): “Fonction de Répartition à n Dimensions et leurs Marges,” *Publications de l’Institut de Statistique de l’Université de Paris*, 8, 229–231.
- TENBUSCH, A. (1994): “Two-dimensional Bernstein Polynomial Density Estimation,” *Metrika*, 41, 233–253.
- (1997): “Nonparametric Curve Estimation with Bernstein Estimates,” *Metrika*, 45, 1–30.
- VITALE, R. (1975): “A Bernstein Polynomial Approach to Density Estimation,” *Puri, M.L. (Ed.), Statistical Inference and Related Topics, vol. 2. Academic Press, New York,* pp. 87–99.
- VOLKONSKI, V. A., AND Y. A. RAZANOV (1959): “Some Limit Theorems for Random Functions. I,” *Theory of Probability and Its Applications*, 4, 178–197.

Recent titles

CORE Discussion Papers

- 2008/8. Géraldine STRACK and Yves POCHET. An integrated model for warehouse and inventory planning.
- 2008/9. Yves SMEERS. Gas models and three difficult objectives.
- 2008/10. Pierre DEHEZ and Daniela TELLONE. Data games. Sharing public goods with exclusion.
- 2008/11. Pierre PESTIEAU and Uri POSSEN. Prodigality and myopia. Two rationales for social security.
- 2008/12. Tim COELLI, Mathieu LEFEBVRE and Pierre PESTIEAU. Social protection performance in the European Union: comparison and convergence.
- 2008/13. Loran CHOLLETE, Andréas HEINEN and Alfonso VALDESOGO. Modeling international financial returns with a multivariate regime switching copula.
- 2008/14. Filomena GARCIA and Cecilia VERGARI. Compatibility choice in vertically differentiated technologies.
- 2008/15. Juan D. MORENO-TERNERO. Interdependent preferences in the design of equal-opportunity policies.
- 2008/16. Ana MAULEON, Vincent VANNETELBOSCH and Wouter VERGOTE. Von Neumann-Morgenstern farsightedly stable sets in two-sided matching.
- 2008/17. Tanguy ISAAC. Information revelation in markets with pairwise meetings: complete information revelation in dynamic analysis.
- 2008/18. Juan D. MORENO-TERNERO and John E. ROEMER. Axiomatic resource allocation for heterogeneous agents.
- 2008/19. Carlo CAPUANO and Giuseppe DE FEO. Mixed duopoly, privatization and the shadow cost of public funds.
- 2008/20. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Forced saving, redistribution and nonlinear social security schemes.
- 2008/21. Philippe CHEVALIER and Jean-Christophe VAN DEN SCHRIECK. Approximating multiple class queueing models with loss models.
- 2008/22. Pierre PESTIEAU and Uri M. POSSEN. Interaction of defined benefit pension plans and social security.
- 2008/23. Marco MARINUCCI. Optimal ownership in joint ventures with contributions of asymmetric partners.
- 2008/24. Raouf BOUCEKKINE, Natali HRITONENKO and Yuri YATSENKO. Optimal firm behavior under environmental constraints.
- 2008/25. Ana MAULEON, Vincent VANNETELBOSCH and Cecilia VERGARI. Market integration in network industries.
- 2008/26. Leonidas C. KOUTSOUGERAS and Nicholas ZIROS. Decentralization of the core through Nash equilibrium.
- 2008/27. Jean J. GABSZEWICZ, Didier LAUSSEL and Ornella TAROLA. To acquire, or to compete? An entry dilemma.
- 2008/28. Jean-Sébastien TRANCREZ, Philippe CHEVALIER and Pierre SEMAL. Probability masses fitting in the analysis of manufacturing flow lines.
- 2008/29. Marie-Louise LEROUX. Endogenous differential mortality, non monitored effort and optimal non linear taxation.
- 2008/30. Santanu S. DEY and Laurence A. WOLSEY. Two row mixed integer cuts via lifting.
- 2008/31. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Taxing sin goods and subsidizing health care.
- 2008/32. Jean J. GABSZEWICZ, Didier LAUSSEL and Nathalie SONNAC. The TV news scheduling game when the newscaster's face matters.
- 2008/33. Didier LAUSSEL and Joana RESENDE. Does the absence of competition *in* the market foster competition *for* the market? A dynamic approach to aftermarket.
- 2008/34. Vincent D. BLONDEL and Yurii NESTEROV. Polynomial-time computation of the joint spectral radius for some sets of nonnegative matrices.

Recent titles

CORE Discussion Papers - continued

- 2008/35. David DE LA CROIX and Clara DELAVALLADE. Democracy, rule of law, corruption incentives and growth.
- 2008/36. Jean J. GABSZEWICZ and Joana RESENDE. Uncertain quality, product variety and price competition.
- 2008/37. Gregor ZOETTL. On investment decisions in liberalized electricity markets: the impact of price caps at the spot market.
- 2008/38. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Habit formation and labor supply.
- 2008/39. Marie-Louise LEROUX and Grégory PONTIERE. Optimal tax policy and expected longevity: a mean and variance approach.
- 2008/40. Kristian BEHRENS and Pierre M. PICARD. Transportation, freight rates, and economic geography.
- 2008/41. Gregor ZOETTL. Investment decisions in liberalized electricity markets: A framework of peak load pricing with strategic firms.
- 2008/42. Raouf BOUCEKKINE, Rodolphe DESBORDES and Hélène LATZER. How do epidemics induce behavioral changes?
- 2008/43. David DE LA CROIX and Marie VANDER DONCKT. Would empowering women initiate the demographic transition in least-developed countries?
- 2008/44. Geoffrey CARUSO, Dominique PEETERS, Jean CAVAILHES and Mark ROUNSEVELL. Space-time patterns of urban sprawl, a 1D cellular automata and microeconomic approach.
- 2008/45. Taoufik BOUEZMARNI, Jeroen V.K. ROMBOUTS and Abderrahim TAAMOUTI. Asymptotic properties of the Bernstein density copula for dependent data.

Books

- Y. POCHET and L. WOLSEY (eds.) (2006), *Production planning by mixed integer programming*. New York, Springer-Verlag.
- P. PESTIEAU (ed.) (2006), *The welfare state in the European Union: economic and social perspectives*. Oxford, Oxford University Press.
- H. TULKENS (ed.) (2006), *Public goods, environmental externalities and fiscal competition*. New York, Springer-Verlag.
- V. GINSBURGH and D. THROSBY (eds.) (2006), *Handbook of the economics of art and culture*. Amsterdam, Elsevier.
- J. GABSZEWICZ (ed.) (2006), *La différenciation des produits*. Paris, La découverte.
- L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
- P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.

CORE Lecture Series

- C. GOURIÉROUX and A. MONFORT (1995), *Simulation Based Econometric Methods*.
- A. RUBINSTEIN (1996), *Lectures on Modeling Bounded Rationality*.
- J. RENEGAR (1999), *A Mathematical View of Interior-Point Methods in Convex Optimization*.
- B.D. BERNHEIM and M.D. WHINSTON (1999), *Anticompetitive Exclusion and Foreclosure Through Vertical Agreements*.
- D. BIENSTOCK (2001), *Potential function methods for approximately solving linear programming problems: theory and practice*.
- R. AMIR (2002), *Supermodularity and complementarity in economics*.
- R. WEISMANTEL (2006), *Lectures on mixed nonlinear programming*.