Regression estimation based on Bernstein density copulas*

Taoufik Bouezmarni [†‡] Benedikt Funke[§] Félix Camirand Lemyre [‡]

November 4, 2014

ABSTRACT

The regression function can be expressed in term of copula density and marginal distributions. In this paper, we propose a new method of estimating a regression function using the Bernstein estimator for the copula density and the empirical distributions for the marginal distributions. The method is fully non-parametric and easy to implement. We provide some asymptotic results related to this copula-based regression modeling by showing the almost sure convergence and the asymptotic normality of the estimator by providing the asymptotic variance. Also, we study the finite sample performance of the estimator.

Key words: Regression estimation, Nonparametric estimation; Bernstein copula density estimator; asymptotic properties; asymptotic normality.

*Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and financial support from the German Science Foundation (DFG), SFB 823 are gratefully acknowledged

[†]Corresponding author. E-mail addressee: Taoufik.Bouezmarni@USherbrooke.ca.

[‡]Département de Mathématiques, Université de Sherbrooke, Sherbrooke, Québec, Canada J1K 2R1.

[§]Technische Universität Dortmund. Fakultät für Mathematik. Lehrstuhl LSIV. Vogelpothsweg 87, 44227 Dortmund.

1 Introduction

The regression analysis is one of the most used tools in practice. The linear and non linear regression models are the popular regression models. The ordinary least squares and the nonlinear least squares are usually used to estimated their parameters. However, the two models are too restrictive and misidentification of the model leads to a wrong conclusions. The nonparametric approach provides an excellent alternative since it does not suppose any specific model for data. Nadaraya-Watson and the local linear kernel are the nonparametric methods that are widely used in order to estimate the regression functions.

Given $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ be a random vector of dimension $d \geq 1$ and Y be a random variable, the regression model is the conditional mean of Y given $\boldsymbol{X}$ which can deduced from the conditional density function of Y given $\boldsymbol{X}$. Thanks to (Sklar 1959), the conditional density can be expressed in term of copula density and marginal distributions. Based on this idea, recently, regression function is expressed in term of copula density and marginal distributions in (Noh, Anouar, and Bouezmarni 2013). This relation has been already applied in (Sungur 2005), (Leong and Valdez 2005) and (Crane and Van Der Hoek 2008) to compute the mean regression function corresponding to several well known copula families (Gaussian, t, Farlie-Gumbel-Morgenstern (FGM), Iterated FGM, Archimedean, etc.) with single ($d = 1$) and multiple covariate(s). To estimate the regression function, (Noh, Anouar, and Bouezmarni 2013) have suggested a semiparametric way by estimating the marginal distributions nonparametrically and by assuming a parametric model for the copula density function. This method avoid the curse of dimensionality and simple to implement. However, the misspecification of the parametric copula can lead to a considerable bias. Indeed, (Dette, Van Hecke, and Volgushev 2014) have considered a simple example where the specification of the parametric copula does not fit the underlying regression.

To avoid the problem of misspecfication, we propose a new estimator by estimating the copula density and the marginal distributions using nonparametric method. We estimate the marginal distributions by the rescaled empirical estimator. Several nonparametric approaches have been proposed to estimate the copula density functions. (Gijbels and Mielniczuk 1990) estimate the bivariate copula density using smoothing kernel methods. They also suggest the reflection method in order to solve the well known 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. Bernstein estimator for copula density was proposed by (Sancetta and Satchell 2004) for independent and identically distributed (i.i.d.) data and investigated by (Bouezmarni, Rombouts, and Taamouti 2010) for dependent data. This method is easy to implement and free from boundary bias problem. Recently, (Bouezmarni and Rombouts 2014) have shown the performance of the Bernstein copula estimator for density that is not necessarily bounded at the corners. Recently, (Janssen, Swanepoel, and Ververbeke 2014) investigate Bernstein estimator when the marginal distributions are estimated. Bernstein estimator for copula density is used by (Bouezmarni, El ghouch, and Taamouti 2014) for testing the conditional independence. Also, in (Bouezmarni and Rombouts 2014), based on Bernstein estimator copula density, they have proposed Granger causality measures.

This paper is organized as follows. The estimator of regression function based on Bernstein copula density estimator is introduced in Section 2. Section 3 provides the asymptotic properties of this estimator. We show the almost sure convergence and the asymptotic normality with the asymptotic variance of the new estimator. In Section 4, we provide simulation results that show the performance of the proposed estimator. Section 5 concludes and the proofs of the asymptotic results are presented in the Appendix.

2 Regression estimator based on Bernstein copula density estimator

Let $\boldsymbol{X} = (X_1, \ldots, X_d)^\top$ be a random vector of dimension $d \geq 1$, a set of covariates, and Y be a random variable, the response variable, with cumulative distribution function (c.d.f.) F_0 and density function f_0. We denote by F_j the c.d.f. of X_j and we denote by f_j its corresponding density. For some $\boldsymbol{x} = (x_1, \ldots, x_d)^\top$ we will use $\boldsymbol{F}(\boldsymbol{x})$ as a shortcut for $(F_1(x_1), \ldots, F_d(x_d))$. From (Sklar 1959), the c.d.f. of $(Y, \boldsymbol{X}^\top)^\top$ evaluated at $(y, \boldsymbol{x}^\top)$ can be expressed as $C(F_0(y), \boldsymbol{F}(\boldsymbol{x}))$, where C is the copula distribution of $(Y, \boldsymbol{X}^\top)^\top$, that is, the distribution function from $[0,1]^{d+1}$ to $[0,1]$ defined by $C(u_0, u_1, \ldots, u_d) = P(F_0(Y) \leq u_0, F_1(X_1) \leq u_1, \ldots, F_d(X_d) \leq u_d)$. Indeed, the distribution function can be controlled by the marginal distributions, which provide the information on each component, and the copula that captures the dependence between components. This fact allows a great flexibility for modeling approach where in the first step one models the marginal distributions and in the second step one characterizes the dependence using a copula model. For more details, see (Nelsen 2006) and (Joe 1997). The conditional density of Y given $\boldsymbol{X}^\top$ is given by $f_0(y)\dfrac{c(F_0(y), \boldsymbol{F}(\boldsymbol{x}))}{c^{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}))}$, where $c(u_0, \boldsymbol{u}) \equiv c(u_0, u_1, \ldots, u_d) = \dfrac{\partial^{d+1} C(u_0, u_1, \ldots, u_d)}{\partial u_0 \partial u_1 \ldots \partial u_d}$ is the copula density corresponding to C and $c^{\boldsymbol{X}}(\boldsymbol{u}) \equiv c^{\boldsymbol{X}}(u_1, \ldots, u_d) = \dfrac{\partial^d C(1, u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d}$ is the copula density of $\boldsymbol{X}$. Obviously, the conditional mean, $m(\boldsymbol{x})$, of Y given $\boldsymbol{X} = \boldsymbol{x}$ can be written as

$$m(\boldsymbol{x}) = \mathbb{E}(Y w(F_0(Y), \boldsymbol{F}(\boldsymbol{x}))) = \frac{e(\boldsymbol{F}(\boldsymbol{x}))}{c^{\boldsymbol{X}}(\boldsymbol{F}(\boldsymbol{x}))}, \tag{1}$$

where $w(u_0, \boldsymbol{u}) = c(u_0, \boldsymbol{u})/c^{\boldsymbol{X}}(\boldsymbol{u})$ and

$$e(\boldsymbol{u}) = \mathbb{E}(Y c(F_0(Y), \boldsymbol{u})) = \int_0^1 F_0^{-1}(u_0) c(u_0, \boldsymbol{u}) du_0. \tag{2}$$

The regression function is a weighted mean with weights induced by the unknown copula density function c defined above.

Note that in the single covariate case and if the covariates are mutually independent, we have, $c^{\boldsymbol{X}}(\boldsymbol{u}) \equiv c_{X_1}(u_1) = 1$ for all $u_1 \in [0,1]$. In such a case the weight function w coincides with the copula density c and (1) reduces to $m(x_1) = e(F_1(x_1)) = \mathbb{E}(Y c(F_0(Y), F_1(x_1)))$.

For given estimator $\hat{w}$, $\hat{F}_0$ and $\hat{F}_j$ for w, F_0 and F_j, respectively, the regression m can be estimated by

$$\hat{m}(\boldsymbol{x}) = \int_{-\infty}^{\infty} y \hat{w}(\hat{F}_0(y), \hat{\boldsymbol{F}}(\boldsymbol{x})) d\hat{F}_0(y), \tag{3}$$

where $\hat{\boldsymbol{F}}(\boldsymbol{x}) = (\hat{F}_1(x_1), \ldots, \hat{F}_d(x_d))^\top$.

Depending on the method to estimate the components in (3), $\hat{m}(\boldsymbol{x})$ can be a nonparametric or a semiparametric or a fully parametric estimator. (Noh, Anouar, and Bouezmarni

2013) have proposed a semiparametric way by estimating the marginal distributions non-parametrically and they suppose a parametric model for the copula density function. This suffers seriously from misspecification of the parametric copula.

Here, we use a nonparametric estimators for c, F_0 and F_j, $j = 1 \ldots, d$, this leads to a fully nonparametric estimator. Nonparametric methods for estimating c is based on Bernstein copula density estimator (see (Sancetta and Satchell 2004)) and given by :

$$c_{k,n}(u_0, \boldsymbol{u}) = \sum_{v_0=0}^{k} \sum_{v_1=0}^{k} \ldots \sum_{v_d=0}^{k} C_n\left(\frac{v_0}{k}, \frac{v_1}{k}, \ldots, \frac{v_d}{k}\right) \prod_{j=0}^{d} p'_{v_j,k}(u_j), \tag{4}$$

where $k \equiv k_n$ is an integer that depends on the sample size n, C_n is the empirical copula function of the vector $(Y, \boldsymbol{X}) = (Y, X_1, \ldots, X_d)$ given by,

$$C_n(u_0, \boldsymbol{u}) = F_n\left(F_{n0}^{-1}(u_0), F_{1n}^{-1}(u_1), \ldots, F_{nd}^{-1}(u_d)\right)$$

with F_n (resp. F_{n0}, F_{1n},...., F_{nd}) is the empirical distribution function of $(Y, \boldsymbol{X})$ (resp. of Y, X_1,..., X_d) and $p'_{v_j,k}(u_j)$ is the derivative, with respect to u_j, of $p_{v_j,k}(u_j)$, for $j = 0, 1, \ldots, d$, which is the binomial distribution function:

$$P_{v_j,k}(s_j) = \begin{pmatrix} k \\ v_j \end{pmatrix} u_j^{v_j}\ (1 - s_j)^{k-v_j}.$$

A nonparametric estimator of the regression function m using the estimator $\hat{c}$ when $d = 1$ or the covariates are mutually independent is given by :

$$\hat{m}_n(x) \quad = \quad \frac{1}{n}\sum_{i=1}^{n} Y_i c_{k,n}(F_{0n}(Y_i), F_{1n}(x))$$

This estimator can be rewritten as folllows:

$$\hat{m}_n(x) \quad = \quad \frac{1}{n}\sum_{i=1}^{n} Y_i \frac{1}{n}\sum_{j=1}^{n}\sum_{l=0}^{k}\sum_{h=0}^{k} \widehat{\kappa}_j(l/k, h/k) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)). \tag{5}$$

where $\kappa_i(u, v) = \mathbb{1}\left(Y_i \le F_0^{-1}(u), X_i \le F_1^{-1}(v)\right)$ and $\widehat{\kappa}_i(u, v) = \mathbb{1}\left(Y_i \le F_{0n}^{-1}(u), X_i \le F_{1n}^{-1}(v)\right)$.
For $d > 1$, the regression function will be estimated by :

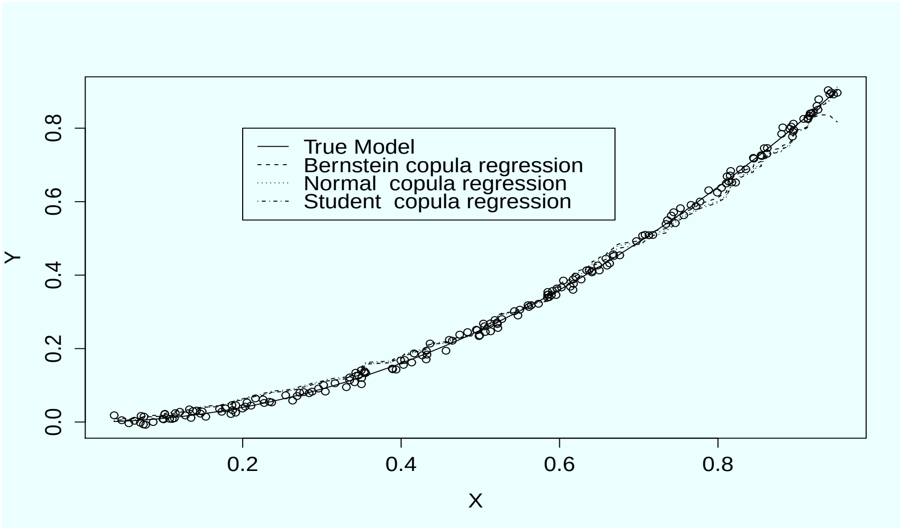
$$\hat{m}_n(\mathbf{x}) = \frac{1}{n c^{\boldsymbol{X}}_{k,n}(F_{1n}(x_1), \ldots, F_{nd}(x_d))} \sum_{i=1}^{n} Y_i c_{k,n}(F_{n0}(Y_i), F_{1n}(x_1), \ldots, F_{nd}(x_d)) \tag{6}$$

The asymptotic normality of $\hat{m}_n(x)$ is hardly tractable since all the observations are used in the regression and in the estimation of the copula density, which causes too much of cross-dependencies. One way to solve this problem is to use one set of data to estimate the copula density and the other to build the regression estimator-part. Thus let $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of i.i.d random variables such that $\mathbb{P}(\xi_i = 0) = \mathbb{P}(\xi_i = 1) = 1/2$. The second version of the regression estimator using Bernstein copula density is defined, when $d = 1$, as

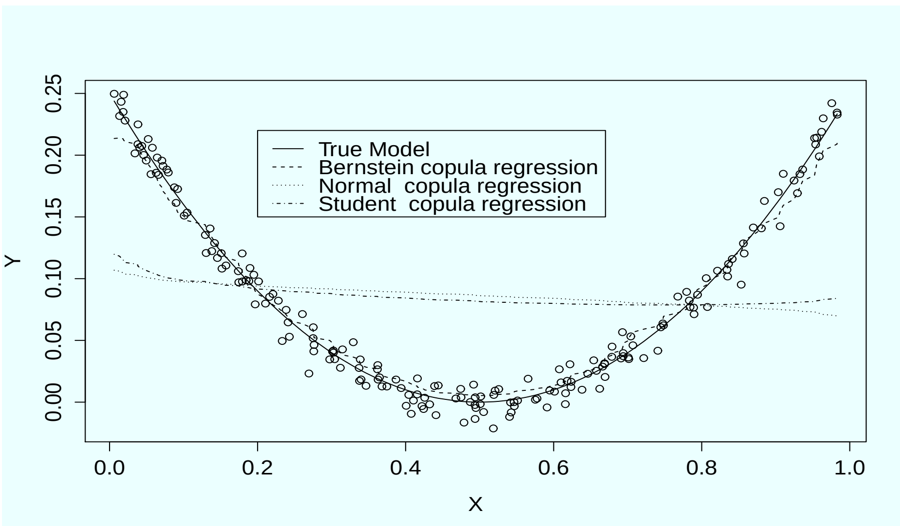
$$\begin{aligned}
m_n^{(\xi)}(x) &= \frac{1}{S_n^{(1)}} \sum_{i=1}^{n} \xi_i Y_i \frac{1}{S_n^{(2)}} \sum_{j=1}^{n} \sum_{l=0}^{k} \sum_{h=0}^{k} (1-\xi_j) \widehat{\kappa}_j(l/k, h/k) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \frac{1}{S_n^{(2)}} \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k} \sum_{h=0}^{k} \widehat{\kappa}_j(l/k, h/k) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i c_{k,n}^{(\xi)}(F_{0n}(Y_i), F_{1n}(x)) \qquad (7)
\end{aligned}$$

where $A_n^{(1)} = \{i : \xi_i = 1, i = 1, ..., n\}$, $A_n^{(2)} = A_n^{(1)c}$, $S_n^{(1)} = \sum_{i=1}^{n} \xi_i$, $S_n^{(2)} = n - S_n^{(1)}$ and $c_{k,n}^{(\xi)}$ is the Bernstein copula density based on data $\{(X_i, Y_i)\}_{i \in A_n^{(2)}}$. Thus conditionally upon the ξ_i's the observations used in the copula density estimator are independent of the observations used in the regression-part of $\hat{m}_n^{(\xi)}$. The strategy is the following. Given almost any sequence of $\{\xi_i\}_{i=1}^{\infty}$ we show the asymptotic normality of $n^{1/2} k^{-1/4} (\hat{m}_n^{(\xi)}(x) - m(x))$ for an arbitrary but fixed value of x.

For illustration, we consider the example given in (Dette, Van Hecke, and Volgushev 2014). This example will be consider in Section 4. In Figure 1 we see that the Bernstein copula estimator and the semiparamatric methods proposed by (Noh, Anouar, and Bouezmarni 2013), based on normal and Student copula, of the first regression model fit well the underlying function. But, for the second model the the semiparametric fails where the Bernstein copula based estimator fits well the regression function. In (Dette, Van Hecke, and Volgushev 2014) the second model is fitted with Frank and mixture of normal copula and both provide a bad estimator.



(a) $m(x) = x^2$



(b) $m(x) = (x - 0.5)^2$

Figure 1: *Bernstein copula and semiparametric copula based regression estimators.*

3 Main results

Here we study the asymptotic properties of the regression estimator using Bernstein estimator for copula densities. The following proposition states the almost sure convergence of the proposed estimator. The following assumptions are required:

Assumption A:

(A.1) The density copula function c has a continuous second-derivative.

(A.2) $k \equiv k_n \sim n^{\alpha}$, $2/(d+4) < \alpha < 1/2$.

(A.3) (i) $\mathbb{E}|Y|^p < \infty, (p \geq 3)$(ii) $\mathbb{E}|Yc_0(F_0(Y), F_1(x))|$ is finite and $\mathbb{E}|Yc_1(F_0(Y), F_1(x))|$ is finite, where $c_0(u,v) = \partial c(u,v)/\partial u$ and $c_1(u,v) = \partial c(u,v)/\partial v$ (iii)$yF_0(y)(1-F_0(y))$ tends to zero when y converges to infinity and (iv) $\int_0^1 (u(1-u))/f_0(F_0^{-1}(u))$ is finite.

Proposition 3.1. *Under Assumptions (A.1), (A.2) and (A.3), we have*

$$\sup_x |\hat{m}_n^{(\xi)}(x) - m(x)| \xrightarrow{a.s.} 0, \textit{ when} \quad n \to \infty. \tag{8}$$

■

The following theorem provides the asymptotic normality of the estimator for $d = 1$ case.

Theorem 3.2. *Let $\hat{m}$ the regression estimator based on Bernstein copula. Under Assumption (A.1), (A.2) and (A.3), we have, for x such that $0 < F_1(x) < 1$ and $d = 1$,*

$$\sqrt{nk^{-1/2}}(\hat{m}_n^{(\xi)}(x) - m(x)) \xrightarrow{d.} N(0, \sigma^2), \textit{ when} \quad n \to \infty. \tag{9}$$

where

$$\sigma^2 = \frac{1}{\sqrt{4\pi F_1(x)(1-F_1(x))}} \left\{ \int_0^1 (F_0^{-1}(u))^2 c(u, F_1(x)) du - \left(\int_0^1 (F_0^{-1}(u)) c(u, F_1(x)) du \right)^2 \right\}.$$

■

The following theorem provides the asymptotic normality of the estimator for $d > 1$ case. For that, we consider the following notaions:

$$\alpha(x) = \sqrt{(4\pi)^d \prod_{j=1}^{d} F_j(x_j)(1 - F_j(x_j))}$$

and

$$\beta(x) = \left\{ \int_0^1 (F_0^{-1}(u))^2 c(u, F_1(x_1), ..., F_d(x_d)) du - \left(\int_0^1 (F_0^{-1}(u)) c(u, F_1(x_1), ..., F_d(x_d)) du \right)^2 \right\}.$$

Theorem 3.3. *Let $\hat{m}$ the regression estimator based on Bernstein copula. Under Assumption (A.1), (A.2) and (A.3), we have, for x such that $0 < F_j(x_j) < 1, j = 1, ..., d$ and $d > 1$,*

$$\sqrt{nk^{-d/2}}(\hat{m}_n^{(\xi)}(x) - m(x)) \xrightarrow{d.} N(0, \sigma^2), \text{ when } \quad n \to \infty. \tag{10}$$

where

$$\sigma^2 \quad = \quad \frac{1}{(c^{\boldsymbol{X}}(x))^2} \frac{\beta(x)}{\alpha(x)}$$

■

Remark 1. *To estimate the asymptotic variance of the regression estimator $\hat{m}_n^{(\xi)}$, it suffices to estimate $c^{\boldsymbol{X}}(x)$ (in the case of multiple covariate), $\alpha(x)$ and $\beta(x)$. We propose to estimate For $c^{\boldsymbol{X}}(x)$ by Bernstein copula density estimator and $\alpha(x)$ by*

$$\alpha(n, x) = \sqrt{(4\pi)^d \prod_{j=1}^{d} F_{nj}(x_j)(1 - F_{nj}(x_j))}.$$

We can see that $\beta(x)$ can be rewritten as follows

$$\beta(x) = \int t^2 c(F_0(t), F_1(x)) f_0(t) dt - \left(\int tc(F_0(t), F_1(x)) f_0(t) dt \right)^2.$$

Hence, a consistent estimator of $\beta(x)$ is given by

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} Y_i^2 c_{k,n}(F_{n0}(Y_i), F_{1n}(x)) - \left(\frac{1}{n} \sum_{i=1}^{n} Y_i c_{k,n}(F_{n0}(Y_i), F_{1n}(x)) \right)^2.$$

■

Remark 2. *If jth component of x is in extreme region, that is $F_j(x_j)$ is close to zero or one, $\alpha(x)$ is too small and then the variance is too large. To overcome this problem, we suggest to a local bandwidth following the idea given in (Omelka, Gijbels, and Veraverbeke 2009).*

■

4 Monte Carlo Simulations

In this section we study the finite sample behavior of our proposed nonparametric copula regression estimator. The first data generating process is especially motivated by (Dette, Van Hecke, and Volgushev 2014), who studied the performance of the semiparametric copula regression estimator by (Noh, Anouar, and Bouezmarni 2013) in a relatively simple model under misspecification. Therefore suppose that (X_i, Y_i), $i = 1, ..., n$, are independent and identically distributed random variables with $X_i \sim U[0,1]$ and $Y_i = (X_i - 0,5)^2 + \sigma\varepsilon_i$, where $\sigma > 0$ and ε_i are also independent and identically distributed such that $\varepsilon_i \sim \mathcal{N}(0,1)$. (Dette, Van Hecke, and Volgushev 2014) found very unpleasant results for the semiparametric copula regression estimator when the copula family is not correctly specified. Especially they showed that in models where the regression function is not strictly monotone over the considered interval, the semiparametric estimator still yields monotone estimates for most of the common parametric copula families. This is the motivation for us to test the performance of our new proposed nonparametric estimator in this nonlinear regression model due to its robustness towards misspecification. At first we set $n = 500$, $k = 25$ and $\sigma = 0,01$ and evaluated the performance of our regression estimator.

The next interesting question was to compare the performance of the copula regression estimator to the performance of standard nonparametric regression estimators like local polynomial estimators. We chose the well-known Nadaraya Watson estimator as well as the local linear estimator. We considered again the parabola model where we now chose $n = 100$, $k = 15$ and $\sigma = 0,05$. The bandwidths for the Nadaraya Watson and the local linear estimator were chosen by a Least Squares cross-validation method (LSCV). It is well known that the local linear estimator has preferable properties compared to the Nadaraya Watson estimator, due to its better performance near the boundary and also due to the independence of the asymptotic bias of the first derivative of the unknown regression function m.

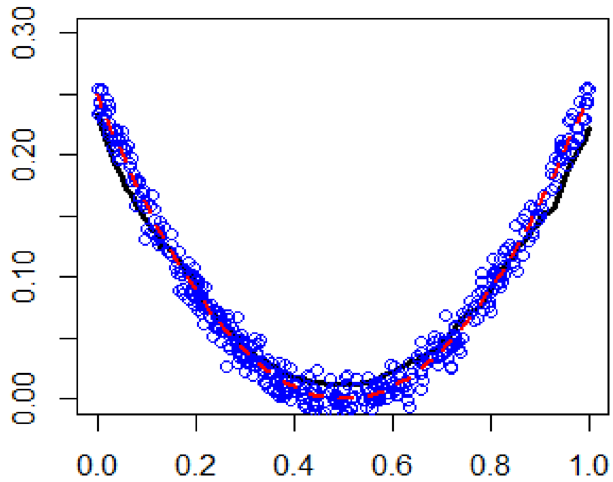
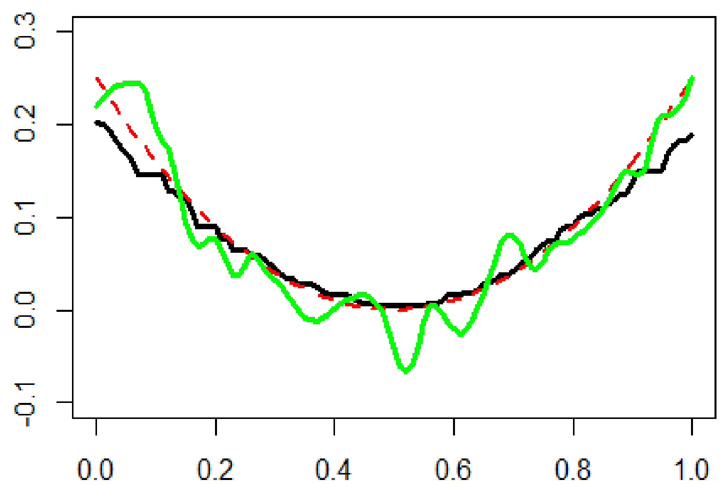
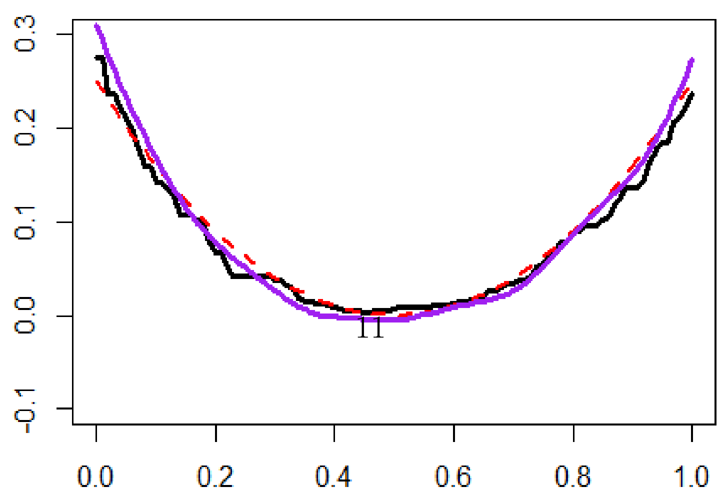


Figure 2: *Observations in blue, Bernstein copula regression estimator in black (MSE≈ $1,43*10^{-4}$) and true regression function in red.*



(a) True regression function in red, Nadaraya Watson estimator in green (MSE≈ $6,63 * 10^{-4}$) and Bernstein copula regression estimator in black (MSE≈ $3,69 * 10^{-4}$)



(b) True regression function in red, local linear estimator in purple

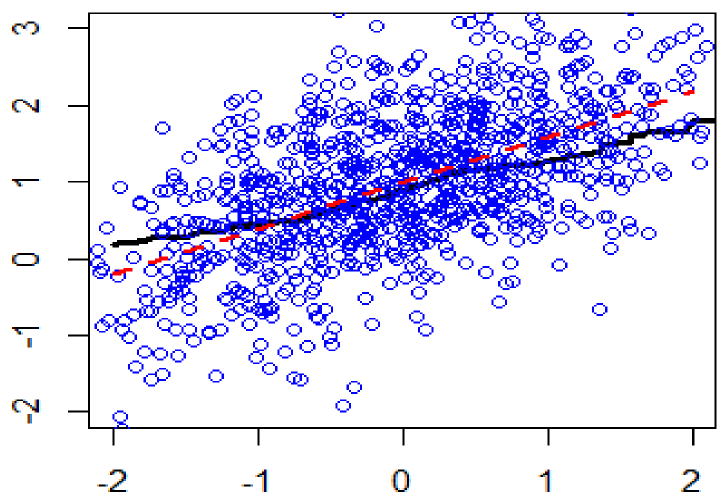


Figure 4: *Observations in blue, Bernstein copula regression estimator in black (MSE*$\approx 1,56*10^{-1}$*) and true regression function in red.*

The next two examples originate from (Noh, Anouar, and Bouezmarni 2013). Therefore let (X_i, Y_i), $i = 1, ..., n$, be independent and identically distributed random variables such that $X_i \sim \mathcal{N}(0,1)$, $Y_i \sim \mathcal{N}(1,1)$ and $(F_Y(Y), F_X(X)) \sim$Gauss-Copula with parameter $\rho = 0,6$. In this case the conditional expectation can be described as $m(x) = E[Y_i|X_i = x] = 1 + 0,6x$. We set $n = 500$, $k = 25$ and checked the performance on the interval $[0,2]$.

The last simulation example contains a misspecified model. We keep the setting up, such that $X_i \sim \mathcal{N}(0,1)$, $Y_i \sim \mathcal{N}(1,1)$ and $(F_Y(Y), F_X(X)) \sim$Gauss-Copula with parameter $\rho = 0,6$. We want to compare the performance of the nonparametric estimator with the performance of a misspecified semiparametric regression estimator. Therefore, instead of the Gaussian copula density, we make use of the density of a Gumbel copula for the estimation of its unknown parameter via the pseudo maximum likelihood estimator. We again set $n = 500$ and estimated the parameter of the Gumbel copula as $\theta \approx 4,3$. The resulting estimates are shown in the following figure. We can see that the nonparametric estimator yields a better result due to its robustness towards misspecification.

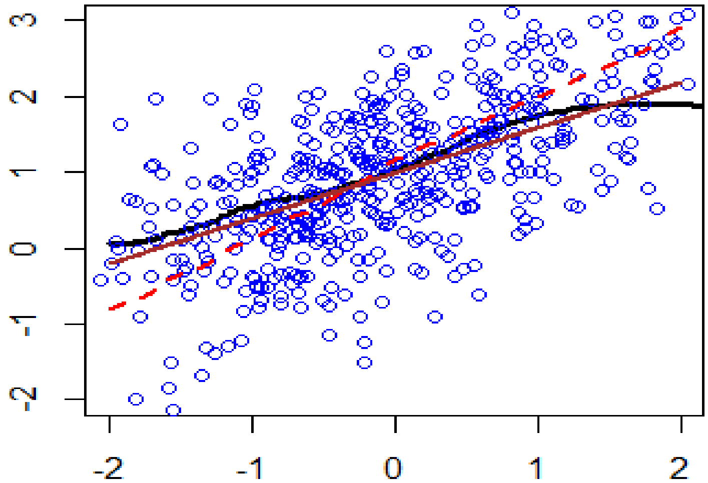


Figure 5: *Observations in blue, Bernstein copula regression estimator in black, semi parametric estimator in red and true regression function in brown.*

5 Conclusion

This paper proposes a new method of estimating a regression function based on the Bernstein estimator for the copula density and the empirical distributions for the marginal distributions. The method is fully non-parametric and easy to implement. We prove the almost sure convergence and the asymptotic normality of the estimator by providing the asymptotic variance. Simulations results shows the performance of the estimator. The comparison of the proposed estimator and the investigation of the optimal bandwidth is left for future research.

Appendix

6 Auxiliary lemma

Proof of Proposition 3.1 We start by

$$\begin{aligned}
\hat{m}_n^{(\xi)}(x) &= \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i c_{k,n}^{(\xi)}(F_{n0}(Y_i), F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i c(F_{n0}(Y_i), F_{1n}(x)) \\
&\quad + \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i c_{k,n}^{(\xi)}(F_{n0}(Y_i), F_{1n}(x)) - \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i c(\hat{F}_0(Y_i), \hat{F}_1(x)) \\
&=: I_n + R_n
\end{aligned}$$

From (Bouezmarni, Rombouts, and Taamouti 2010) and Assumption (A.1), (A.2), (A.3)(i), we have

$$\begin{aligned}
|R_n| &\leq \sup_i |c_{k,n}^{(\xi)}(F_{n0}(Y_i), F_{1n}(x)) - c(F_{n0}(Y_i), F_{1n}(x))| \left(\frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} |Y_i|\right) \\
&= o(1), \quad a.s.
\end{aligned}$$

Analogously to the proof of (Noh, Anouar, and Bouezmarni 2013) we define $\tilde{u}_{i,0} := F_0(Y_i) + t(F_{n0}(Y_i) - F_0(Y_i)), t \in [0,1]$ respectively $\tilde{u}_1 := F_1(x) + t(F_{1n}(x) - F_1(x)), t \in [0,1]$. Using a Taylor expansion of the first order we can conclude that

$$\begin{aligned}
I_n &= \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i \left[c(F_0(Y_i), F_1(x)) + (F_{n0}(Y_i) - F_0(Y_i))c_0(\tilde{u}_{i,0}, \tilde{u}_1) + (F_{1n}(x) - F_1(x))c_1(\tilde{u}_{i,0}, \tilde{u}_1)\right] \\
&= \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i \left[c(F_0(Y_i), F_1(x))\right] + \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i (F_{n0}(Y_i) - F_0(Y_i))c_0(F_0(Y_i), F_1(x)) \\
&+ \frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} Y_i \left[(F_{1n}(x) - F_1(x))c_1(F_0(Y_i), F_1(x))\right] + R1 + R_2 \qquad (11)
\end{aligned}$$

We start with the residual terms. Using the Assumption (A.1) and (A.3)(i), we derive:

$$\begin{aligned}
|R_1| &\leq \sup_i |F_{n0}(Y_i) - F_0(Y_i)| \sup_i |c_0(\tilde{u}_{i,0}, \tilde{u}_1) - c_0(F_0(Y_i), F_1(x))| \left(\frac{1}{S_n^{(1)}}\sum_{A_n^{(1)}} |Y_i|\right) \\
&= o(1), \quad a.s.
\end{aligned}$$

and

$$\begin{aligned}|R_2| &\leq \sup_i |F_{1n}(x) - F_1(x)| \sup_i |c_1(\tilde{u}_{i,0}, \tilde{u}_1) - c_1(F_0(Y_i), F_1(x))| \left(\frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} |Y_i| \right) \\ &= o(1), \quad a.s.\end{aligned}$$

We examine the second and the third term and this yields using Assumption (A.3)(ii) and (A.3)(iii)

$$\begin{aligned}\left| \frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} Y_i (F_{n0}(Y_i) - F_0(Y_i)) c_0(F_0(Y_i), F_1(x)) \right| &\leq \sup_i |F_{n0}(Y_i) - F_0(Y_i)| \frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} |Y_i c_0(F_0(Y_i), F_1(x))| \\ &\approx \sup_i |F_{n0}(Y_i) - F_0(Y_i)| \mathbb{E}\left[|Y c_0(F_0(Y), F_1(x))|\right] \\ &= o(1), \quad a.s.\end{aligned}$$

and

$$\begin{aligned}\left| \frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} Y_i (F_{1n}(x) - F_1(x)) c_1(F_0(Y_i), F_1(x)) \right| &\leq \sup_i |F_{1n}(x) - F_1(x)| \frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} |Y_i c_1(F_0(Y_i), F_1(x))| \\ &\approx \sup_i |F_{1n}(x) - F_1(x)| \mathbb{E}\left[|Y c_1(F_0(Y), F_1(x))|\right] \\ &= o(1), \quad a.s.\end{aligned}$$

Now, it is easy to see that the first term is

$$\frac{1}{S_n^{(1)}} \sum_{A_n^{(1)}} Y_i \left[c(F_0(Y_i), F_1(x))\right] = \mathbb{E}[Y c(F_0(Y), F_1(x))] + o(1), \quad a.s. \tag{12}$$

Using these results we can finally conclude the proof of Proposition 3.1.

■

Before showing the asymptotic normality of the nonparametric regression estimator based on Bernstein copula density estimator at point x where $0 < F_1(x) < 1$, we recall an interesting result given in Lemma 1 of (Janssen, Swanepoel, and Ververbeke 2014) which states that for any fixed $0 < v < 1$ we have

$$\sum_{l=0}^{k} \left| P'_{l,k}(v) \right| \sim \mathcal{O}\left(\frac{k^{1/2}}{\sqrt{v(1-v)}} \right). \tag{13}$$

We now investigate the asymptotic representation this sum when $v = v_n$ is a sequence. First notice that If $v_n \to v > 0$, then the approximation (13) still holds. We explore the case where $v_n \searrow 0$. The case $v_n \nearrow 1$ can be dealt similarly.

Suppose $kv_n \to \infty$. Mimicking the proof in (Janssen, Swanepoel, and Ververbeke 2014) for a sequence v_n going down to zero, one gets that

$$\sum_{l=0}^{k} \left|P'_{l,k}(v_n)\right| = \frac{2}{v_n(1-v_n)}([kv_n]+1)\binom{k}{[kv_n]+1} v_n^{[kv_n]+1}(1-v_n)^{k-[kv_n]}.$$

If kv_n goes to infinity one can still use stirling formula to obtain that

$$\sum_{l=0}^{k} \left|P'_{l,k}(v_n)\right| \sim \sqrt{\frac{2}{\pi}} \frac{k^{1/2}}{\sqrt{v_n(1-v_n)}}.$$

However, if $kv_n \to K > 0$. Again mimicking the proof of (Janssen, Swanepoel, and Ververbeke 2014) for a sequence v_n satisfying such a condition one gets that

$$\sum_{l=0}^{k} \left|P'_{l,k}(v_n)\right| \leq \frac{2(K+1)}{v_n(1-v_n)}.$$

Also, if $kv_n \to 0$, since, for large enough k, $kv_n < \frac{1}{2}$ then $[kv_n] = 0$ and therefore

$$\begin{aligned}\sum_{l=0}^{k} \left|P'_{l,k}(v_n)\right| &= \frac{2}{(1-v_n)} k(1-v_n)^k \\ &= \mathcal{O}(k).\end{aligned}$$

One can summaries the previous development in the following lemma.

Lemma 1. *Let v_n be a sequence such that $v_n \in [0,1]$ for all $n \in \mathbb{N}$, and $v_n \to v \in [0,1]$.*

1. *If $v > 0$, then*
$$\sum_{l=0}^{k} |P'_{k,l}(v_n)| \sim \mathcal{O}\left(\frac{k^{1/2}}{\sqrt{v(1-v)}}\right)$$
2. *If $v_n \searrow 0$ or $v_n \nearrow 1$ then*
$$\sum_{l=0}^{k} |P'_{k,k}(v_n)| \sim \mathcal{O}(k).$$

Proof of Theorem 3.2 First notice that

$$c_{k,n}(u,v) \quad = \quad \sum_{l=0}^{k}\sum_{h=0}^{k} C_n(l/k,h/k)P'_{k,l}(u)P'_{k,h}(v) \tag{14}$$

where C_n is the empirical copula estimator. Next, following the idea of (Janssen, Swanepoel, and Ververbeke 2014), since the copula C is differentiable with respect to each of its arguments:

$$\begin{aligned} C_n(u,v) \quad = \quad & C(u,v) + (\frac{1}{n}\sum_{i=1}^{n}\kappa_i(u,v) - C(u,v)) \\ & + C^{(1)}(u,v)(\frac{1}{n}\sum_{i=1}^{n}\kappa_i(u,1) - u) \\ & + C^{(2)}(u,v)(\frac{1}{n}\sum_{i=1}^{n}\kappa_i(1,v) - v) + \frac{1}{n}\sum_{i=1}^{n}\zeta_{i,n}(u,v) \end{aligned} \tag{15}$$

where

$$\zeta_{i,n}(u,v) \quad = \quad \widehat{\kappa}_i(u,v) - \kappa_i(u,v) - C^{(1)}(u,v)(\kappa_i(u,1)-u) - C^{(2)}(u,v)(\kappa_i(1,v)-v). \tag{16}$$

Plugging (15) into (14) leads to

$$c_{k,n}(u,v) \quad = \textstyle\sum_{z=0}^{4} c_{k,n}^{(z)}(u,v) \tag{17}$$

where

$$\begin{aligned} c_{k,n}^{(0)}(u,v) \quad &= \quad \sum_{l=0}^{k}\sum_{h=0}^{k} C(l/k,h/k)P'_{k,l}(u)P'_{k,h}(v) \\ c_{k,n}^{(1)}(u,v) \quad &= \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{l=0}^{k}\sum_{h=0}^{k}(\kappa_i(l/k,h/k) - C(l/k,h/k))P'_{k,l}(u)P'_{k,h}(v) \\ c_{k,n}^{(2)}(u,v) \quad &= \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{l=0}^{k}\sum_{h=0}^{k} C^{(1)}(l/k,h/k)(\kappa_i(l/k,1) - l/k)P'_{k,l}(u)P'_{k,h}(v) \\ c_{k,n}^{(3)}(u,v) \quad &= \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{l=0}^{k}\sum_{h=0}^{k} C^{(2)}(l/k,h/k)(\kappa_i(1,h/k) - h/k)P'_{k,l}(u)P'_{k,h}(v) \\ c_{k,n}^{(4)}(u,v) \quad &= \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{l=0}^{k}\sum_{h=0}^{k} \zeta_{i,n}(l/k,h/k)P'_{k,l}(u)P'_{k,h}(v). \end{aligned} \tag{18}$$

In order to match the setup of the last section one can simply replace the correspondents indices in $c_{m,n}^{(k)}$ for k from 0 to 4 by those of $A_n^{(2)}$ to obtain $c_{k,n}^{(\xi j)}, j = 0,1,2,3,4$. Finally, we replace the decomposition (18) in $\hat{m}_n^{(\xi)}(x)$ and get the following decomposition $\hat{m}_n^{(\xi)}(x) = \sum_{j=0}^{4} m_n^{(\xi j)}(x)$ with

$$\hat{m}_n^{(\xi j)}(x) = \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i c_{k,n}^{(\xi j)}(F_{0n}(Y_i), F_{1n}(x)), \quad \text{for} \quad j = 0, 1, 2, 3, 4.$$

Now, the goal is the derive the asymptotic representation of each of these terms. We prove, in Section 6.1 that $m_n^{(\xi 0)}(x)$, is asymptotically equal to $m(x)$. Later, we show, in Section 6.4 (resp. 6.2), that $\sqrt{nk^{-1/2}} m_n^{(\xi 2)}(x)$ (resp. $\sqrt{nk^{-1/2}} m_n^{(\xi 2)}(x)$) are negligible. Finally, we establish in Section 6.3 and 6.5 the asymptotic normality of $\sqrt{nk^{-1/2}} m_n^{(\xi 1)}(x) + \sqrt{nk^{-1/2}} m_n^{(\xi 3)}(x)$.

6.1 Asymptotic representation of $m_n^{(\xi 0)}(x)$

First

$$\begin{aligned} m_n^{(\xi 0)}(x) &= \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{l=0}^{k} \sum_{h=0}^{k} C(l/k, h/k) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\ &= \frac{k^2}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1 C((l+1)/k, (h+1)/k) P_{k-1,l}(F_{0n}(Y_i) P_{k-1,h}(F_{1n}(x)) \end{aligned}$$

where

$$\Delta_1 C(l/k, h/k) = C(l/k, h/k) - C((l-1)/k, h/k) - C(l/k, (h-1)/k) + C((l-1)/k, (h-1)/k).$$

Since C has a Lebesgue bounded density c, the latter is equal to $\int_{(l-1)/k}^{l/k} \int_{(h-1)/k}^{h/k} c(u, v) du dv$. Thus by the mean value theorem

$$
\begin{aligned}
m_n^{(\xi 0)}(x) &\approx \frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} c(\frac{l}{k-1},\frac{h}{k-1})P_{k-1,l}(F_{0n}(Y_i))P_{k-1,h}(F_{1n}(x))\\
&= \frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} c(\frac{l}{k-1},\frac{h}{k-1})P_{k-1,l}(F_{0}(Y_i))P_{k-1,h}(F_{1}(x))\\
&\quad +\frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} c(\frac{l}{k-1},\frac{h}{k-1})\\
&\quad \times[P_{k-1,l}(F_{0n}(Y_i))-P_{k-1,l}(F_0(Y_i))]P_{k-1,h}(F_{1n}(x))\\
&\quad +\frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} c(\frac{l}{k-1},\frac{h}{k-1})\\
&\quad \times[P_{k-1,h}(F_{1n}(x))-P_{k-1,h}(F_1(x))]P_{k-1,l}(F_0(Y_i))\\
&= (A)+(B)+(C). \qquad (19)
\end{aligned}
$$

(B) and (C) can be treat similarly in the following way. We only show the representation of the term (B). Using the mean value theorem:

$$
(B) = \frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} c(\frac{l}{k-1},\frac{h}{k-1})P'_{k-1,l}(\eta_{n,i})P_{k-1,h}(F_{1n}(x))(F_{0n}(Y_i)-F_0(Y_i))
$$

where $\eta_{n,i}$ lies between $F_{0n}(Y_i)$ and $F_0(Y_i)$. Since c has bounded derivatives one has

$$
(B) = \frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i \sum_{l=0}^{k-2}\sum_{h=0}^{k-1} c^{(1)}(\frac{l}{k-2},\frac{h}{k-1})P_{k-2,l}(\eta_{n,i})P_{k-1,h}(F_{1n}(x))(F_{0n}(Y_i)-F_0(Y_i)).
$$

Finally, since $\frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i = \mathcal{O}_p(1)$ $\mathbb{P}_\xi-$a.s, $c^{(1)}$ is bounded, $\sup_y |F_{0n}(y)-F_0(y)| = \mathcal{O}_p(1/\sqrt{n})$ and the Bernstein weights sums to 1, (B) is $\mathcal{O}_p\left(n^{-1/2}\right)$. Applying the same strategy to (C) yields

$$
m_n^{(\xi 0)}(x) = (A) + \mathcal{O}_p\left(n^{-1/2}\right). \qquad (20)
$$

To conclude with $m_n^{(\xi 0)}(x)$, notice that

$$
\begin{aligned}
(A) &= \frac{1}{S_n^{(1)}}\sum_{i\in A_n^{(1)}} Y_i c(F_0(Y_i),F_1(x)) + \mathcal{O}_p(1/k)\\
&= \mathbb{E}\left(Y c(F_0(Y_i),F_1(x))\right) + \mathcal{O}_p\left(1/\sqrt{n}\right) + \mathcal{O}_p(1/k) \qquad \mathbb{P}_\xi - \text{a.s.} \qquad (21)
\end{aligned}
$$

Since the norming factor is $n^{1/2}k^{-1/4}$ the two terms of the right hand side are negligible provided $k \sim n^{\alpha}$ with $\alpha > 2/5$.

6.2 Asymptotic negligibility of $m_n^{(\xi 4)}(x)$

First recall that

$$m_n^{(\xi 4)}(x) = \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \frac{1}{S_n^{(2)}} \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k} \sum_{h=0}^{k} \zeta_{j,n}(l/k, h/k) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x))$$

Denote $\mathbb{E}_A$ the expectation conditionally upon $\{Y_i : i \in A\}$. One has

$$\begin{aligned}
\mathbb{E}_{A_n^{(1)}}\left(m_n^{(\xi 4)}(x)\right) &= \frac{1}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k} \sum_{h=0}^{k} \mathbb{E}\left(\zeta_{j,n}(l/k, h/k)\right) P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k} \sum_{h=0}^{k} \mathbb{E}\left(\widehat{\kappa}_j(l/k, h/k) - \kappa_j(l/k, h/k)\right) \\
&\quad \times P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{l=0}^{k} \sum_{h=0}^{k} \Big(\boldsymbol{F}(F_{0n}^{-1}(l/k), F_{1n}^{-1}(h/k)) - \boldsymbol{F}(F_0^{-1}(l/k), F_1^{-1}(h/k)) \Big) \\
&\quad \times P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&= \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{l=0}^{k} \sum_{h=0}^{k} \Big(\boldsymbol{F}^{(1)}(\eta_{n,0}(l/k), F_{1n}^{-1}(h/k))[F_{0n}^{-1}(h/k) - F_0^{-1}(h/k)] \Big) \\
&\quad \times P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x)) \\
&\quad + \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{l=0}^{k} \sum_{h=0}^{k} \Big(\boldsymbol{F}^{(2)}(F_{0n}^{-1}(l/k), \eta_{n,1}(h/k))[F_{1n}^{-1}(h/k) - F_1^{-1}(h/k)] \Big) \\
&\quad \times P'_{k,l}(F_{0n}(Y_i)) P'_{k,h}(F_{1n}(x))
\end{aligned}$$

where $\eta_{n,z}(t/k)$ lies between $F_{zn}^{-1}(t/k)$ and $F_z^{-1}(t/k)$ for $z = 0, 1$. Since $\boldsymbol{F}^{(1)}$ and $\boldsymbol{F}^{(2)}$ are uniformly continuous, and since the derivatives of $\boldsymbol{F}^{(1)}(F_0^{-1}(u), F_1^{-1}(v))$ and $\boldsymbol{F}^{(2)}(F_0^{-1}(u), F_1^{-1}(v))$ with respect to each of their arguments are continuous and uniformly bounded,

$$\frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i = \mathcal{O}_p(1) \qquad \mathbb{P}_{\xi} - \text{a.s},$$

and since $[F_{jn}^{-1} - F_j^{-1}] = \mathcal{O}_p\left(1/\sqrt{n}\right)$, $j = 0$ or 1, one gets that

$$\mathbb{E}\Big(m_n^{(\xi 4)}(x)\Big) \quad = \quad \mathcal{O}_p\left(1/\sqrt{n}\right) \qquad \mathbb{P}_\xi - \text{a.s.} \tag{22}$$

Next we calculate the variance of $m_n^{(\xi 4)}(x)$. That is,

$$\begin{aligned} \mathbb{V}\left(m_n^{(\xi 4)}(x)\right) \quad = \quad & \Big(\frac{1}{S_n^{(1)}S_n^{(2)}}\Big)^2 \sum_{i,i' \in A_n^{(1)}} \sum_{j,j' \in A_n^{(2)}} \sum_{l,l'=0}^{k} \sum_{h,h'=0}^{k} P'_{k,h}(F_{1n}(x))P'_{k,h'}(F_{1n}(x)) \\ & \times \mathrm{Cov}\Big(Y_i\zeta_{j,n}(l/k,h/k)P'_{k,l}(F_{0n}(Y_i)), Y_{i'}\zeta_{j',n}(l'/k,h'/k)P'_{k,l'}(F_{0n}(Y_{i'}))\Big) \end{aligned}$$

Notice that if simultaneously $i \neq i'$ and $j \neq j'$ then the covariance is 0. We deal the three other cases separately.

Case 1. $i = i'$, $j \neq j'$.

$$\begin{aligned} & \sum_{l,l'=0}^{k} \sum_{h,h'=0}^{k} \mathbb{E}_{A_n^{(1)}}\Big(Y_i^2 P'_{k,l}(F_{0n}(Y_i))\zeta_{j,n}(l/k,h/k)\zeta_{j',n}(l'/k,h'/k)P'_{k,l'}(F_{0n}(Y_i))\Big) \\ & \times P'_{k,h}(F_{1n}(x))P'_{k,h'}(F_{1n}(x)) \\ & = Y_i^2\Big(\sum_{l=0}^{k}\sum_{h=0}^{k} \mathbb{E}\left(\zeta_{j,n}(l/k,h/k)\right) P'_{k,l}(F_{0n}(Y_i))P'_{k,h}(F_{1n}(x))\Big)^2 \\ & = Y_i^2 \times \mathcal{O}_p\left(1/n\right) \end{aligned} \tag{23}$$

which implies that

$$\begin{aligned} & \Big(\frac{1}{S_n^{(1)}S_n^{(2)}}\Big)^2 \sum_{i \in A_n^{(1)}} \sum_{j,j' \in A_n^{(2)}: j \neq j'} \sum_{l,l'=0}^{k} \sum_{h,h'=0}^{k} P'_{k,h}(F_{1n}(x))P'_{k,h'}(F_{1n}(x)) \\ & \times \mathrm{Cov}\Big(Y_i\zeta_{j,n}(l/k,h/k)P'_{k,l}(F_{0n}(Y_i)), Y_i\zeta_{j',n}(l'/k,h'/k)P'_{k,l'}(F_{0n}(Y_i))\Big) \\ & \sim \mathcal{O}\left(1/n^2\right) \qquad \mathbb{P}_\xi - \text{a.s} \end{aligned} \tag{24}$$

since $S_n^{(z)}/n \overset{\mathbb{P}_\xi - \text{a.s}}{\to} 1/2$ by the strong law of large numbers.

Case 2. $i \neq i'$, $j = j'$.

We first need to compute

$$
\begin{aligned}
&\sum_{l,l'=0}^{k}\sum_{h,h'=0}^{k}\mathbb{E}_{A_n^{(1)}}\Big(Y_i'Y_jP'_{k,l}(F_{0n}(Y_i))\zeta_{j,n}(l/k,h/k)^2P'_{k,l'}(F_{0n}(Y_i'))\Big)\\
&\times P'_{k,h}(F_{1n}(x))P'_{k,h'}(F_{1n}(x))\\
&\leq |Y_iY_{i'}|\Big(\sum_{l=0}^{k}|P'_{k,l}(F_{0n}(Y_i)|\Big)\times\Big(\sum_{h=0}^{k}|P'_{k,h}(F_{0n}(Y_i')|\Big) \qquad (25)\\
&\times\Big(\sum_{h=0}^{k}|P'_{k,h}(F_{1n}(x)|\Big)^2\mathcal{O}_p\left(n^{-3/2}\right) \qquad (26)
\end{aligned}
$$

from the fact that $\sup_{u,v}|\zeta_{j,n}(u,v)| \sim \mathcal{O}_p\left(n^{-3/4}\right)$, see (Janssen, Swanepoel, and Ververbeke 2014). From the continuous mapping theorem together with lemma 13-1 combined with the fact that $F_1(x)(1-F_1(x))>0$:

$$
\sum_{h=0}^{k}|P'_{k,h}(F_{1n}(x))| \sim \mathcal{O}\left(k^{1/2}\right). \qquad (27)
$$

Next, since one has no control over the values of $(F_{0n}(Y_i))$ one has to use lemma 13-2 to conclude that

$$
\sum_{l=0}^{k}|P'_{k,l}(F_{0n}(Y_i))| \sim \mathcal{O}_p\left(k\right). \qquad (28)
$$

Thus

$$
\begin{aligned}
&\Big(\frac{1}{S_n^{(1)}S_n^{(2)}}\Big)^2\sum_{i,i'\in A_n^{(1)}:i\neq i'}\sum_{j\in A_n^{(2)}}\sum_{l,l'=0}^{k}\sum_{h,h'=0}^{k}P'_{k,h}(F_{1n}(x))P'_{k,h'}(F_{1n}(x))\\
&\times\mathrm{Cov}\Big(Y_i\zeta_{j,n}(l/k,h/k)P'_{k,l}(F_{0n}(Y_i)),Y_{i'}\zeta_{j,n}(l'/k,h'/k)P'_{k,l'}(F_{0n}(Y_{i'}))\Big)\\
&\sim \mathcal{O}_p\left(k^3/n^{5/2}\right) \qquad \mathbb{P}_\xi-\text{a.s.} \qquad (29)
\end{aligned}
$$

Similarly, if $i=i'$ and $j=j'$ one gets the bound $\mathcal{O}_p\left(k^3/n^{7/2}\right)$ $\mathbb{P}_\xi-$a.s. Combining the three previous results yields

$$
\mathbb{V}\left(m_n^{(\xi 4)}(x)\right) = \mathcal{O}_p\left(k^3/n^{5/2}\right)+\mathcal{O}_p\left(1/n^2\right) \qquad \mathbb{P}_\xi-\text{a.s.} \qquad (30)
$$

Equation (30) together with the factor $n^{1/2}k^{-1/4}$ yields the negligibility of $m_n^{(\xi 4)}(x)$ provided $k^{5/2}n^{-3/2}\to 0$, that is, $k\sim n^\alpha$, $\alpha<3/5$.

6.3 Asymptotic representation of $m_n^{(\xi 1)}(x)$

Recall that

$$
\begin{aligned}
m_n^{(\xi 1)}(x) &= \frac{k^2}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1(\kappa_j(\frac{l+1}{k}, \frac{h+1}{k}) - C(\frac{l+1}{k}, \frac{h+1}{k})) \\
&\quad \times P_{k-1,l}(F_{0n}(Y_i)) P_{k-1,h}(F_{1n}(x)).
\end{aligned}
$$

The first step is to investigate the effect of replacing the estimated marginales F_{0n} and F_{1n} by the true ones in the bernstein weights. To do this denote

$$
\begin{aligned}
\widetilde{m}_n^{(\xi 1)}(x) &= \frac{k^2}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1(\kappa_j(\frac{l+1}{k}, \frac{h+1}{k}) - C(\frac{l+1}{k}, \frac{h+1}{k})) \\
&\quad \times P_{k-1,l}(F_0(Y_i)) P_{k-1,h}(F_1(x)). \qquad (31)
\end{aligned}
$$

Now

$$
\begin{aligned}
m_n^{(\xi 1)}(x) - \widetilde{m}_n^{(\xi 1)}(x) &= \frac{k^2}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1(\kappa_j(\frac{l+1}{k}, \frac{h+1}{k}) - C(\frac{l+1}{k}, \frac{h+1}{k})) \\
&\quad \times [P_{k-1,l}(F_{0n}(Y_i)) - P_{k-1,l}(F_0(Y_i))] P_{k-1,h}(F_1(x)) \\
&\quad + \frac{k^2}{S_n^{(1)} S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1(\kappa_j(\frac{l+1}{k}, \frac{h+1}{k}) - C(\frac{l+1}{k}, \frac{h+1}{k})) \\
&\quad \times [P_{k-1,h}(F_{1n}(x)) - P_{k-1,h}(F_1(x))] P_{k-1,l}(F_{0n}(Y_i)) \\
&= A_n(x) + B_n(x). \qquad (32)
\end{aligned}
$$

It is clear that both $\mathbb{E}\left(A_n(x)\right)$ and $\mathbb{E}\left(B_n(x)\right)$ are 0. Next,

$$
\begin{aligned}
\mathbb{E}_{A_n^{(1)}}\left(A_n(x)^2\right) &= \left(\frac{k^2}{S_n^{(1)}S_n^{(2)}}\right)^2 \sum_{i,i'\in A_n^{(1)}} Y_iY_{i'} \sum_{j\in A_n^{(2)}} \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k},\frac{h+1}{k}\right)P_{k-1,h}(F_1(x))^2 \\
&\quad \times[P_{k-1,l}(F_{0n}(Y_i)) - P_{k-1,l}(F_0(Y_i))] \times [P_{k-1,l}(F_{0n}(Y_{i'})) - P_{k-1,l}(F_0(Y_{i'}))] \\
&\quad +\left(\frac{k^2}{S_n^{(1)}S_n^{(2)}}\right)^2 \Big\{ \sum_{i\in A_n^{(1)}} Y_i \sum_{j\in A_n^{(2)}} \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k},\frac{h+1}{k}\right)P_{k-1,h}(F_1(x)) \\
&\quad \times[P_{k-1,l}(F_{0n}(Y_i)) - P_{k-1,l}(F_0(Y_i))]\Big\} \\
&\quad \times\Big\{ \sum_{i'\in A_n^{(1)}} Y_{i'} \sum_{j\in A_n^{(2)}} \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k},\frac{h+1}{k}\right)P_{k-1,h}(F_1(x)) \\
&\quad \times[P_{k-1,l}(F_{0n}(Y_{i'})) - P_{k-1,l}(F_0(Y_{i'}))]\Big\} \\
&= A_{1n}(x) + A_{2n}(x). \qquad (33)
\end{aligned}
$$

Again using the mean value theorem

$$
\begin{aligned}
A_{1n}(x) &= \left(\frac{k^2}{S_n^{(1)}}\right)^2 \frac{1}{S_n^{(2)}} \sum_{i,i'\in A_n^{(1)}} Y_iY_{i'} \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k},\frac{h+1}{k}\right)P_{k-1,h}(F_1(x))^2 \\
&\quad \times P'_{k-1,l}(\eta_{n,i})P'_{k-1,l}(\eta_{n,i'})\left[F_{0n}(Y_i) - F_0(Y_i)\right] \times \left[F_{0n}(Y_{i'}) - F_0(Y_{i'})\right]
\end{aligned}
$$

where $\eta_{n,z}$ lies between $F_0(Y_z)$ and $F_{0n}(Y_z)$. Because of lemma 2 of (Janssen, Swanepoel, and Ververbeke 2014), and since C has bounded second order derivative:

$$
\begin{aligned}
A_{1n}(x) &= k^{5/2}\left(\frac{1}{S_n^{(1)}}\right)^2 \frac{1}{S_n^{(2)}} \sum_{i,i'\in A_n^{(1)}} Y_iY_{i'} \sum_{l=0}^{k-1} \Delta_1 c\left(\frac{l+1}{k}, F_1(x)\right) \\
&\quad \times P'_{k-1,l}(\eta_{n,i})P'_{k-1,l}(\eta_{n,i'})\left[F_{0n}(Y_i) - F_0(Y_i)\right] \times \left[F_{0n}(Y_{i'}) - F_0(Y_{i'})\right] + \mathcal{O}\left(k^{-1}\right). \qquad (34)
\end{aligned}
$$

Next we need to break $A_{1n}(x)$ into 3 cases, according to the values of $F_0(Y_i)$ and $F_0(Y_{i'})$.

First, according to lemma 13:

$$
\begin{aligned}
A_{1n}^{(1)}(x) &= k^{5/2}\big(\frac{1}{S_n^{(1)}}\big)^2\frac{1}{S_n^{(2)}}\sum_{i\neq i'\in A_n^{(1)}} Y_iY_{i'}\mathbb{1}\left(|F_0(Y_i)-F_0(Y_i')|<k^{-1/2+\epsilon/3}, F_0(Y_i)<k^{-1/2+2\epsilon/3}\right)\\
&\quad\times\sum_{l=0}^{k-1}\Delta_1 c(F_1(x),\frac{h+1}{k})\times P'_{k-1,l}(\eta_{n,i})P'_{k-1,l}(\eta_{n,i'})\\
&\quad\times\big[F_{0n}(Y_i)-F_0(Y_i)\big]\times\big[F_{0n}(Y_{i'})-F_0(Y_{i'})\big]\\
&\quad+k^{5/2}\big(\frac{1}{S_n^{(1)}}\big)^2\frac{1}{S_n^{(2)}}\sum_{i\in A_n^{(1)}} Y_i^2\sum_{l=0}^{k-1}\Delta_1 c(F_1(x),\frac{h+1}{k})\times P'_{k-1,l}(\eta_{n,i})^2\times\big[F_{0n}(Y_i)-F_0(Y_i)\big]^2\\
&\leq 2n^{-1}k^{7/2}\big(\frac{1}{S_n^{(1)}}\big)^2\frac{1}{S_n^{(2)}}\left\|c\right\|_\infty\\
&\quad\times\sum_{i\neq i'\in A_n^{(1)}}|Y_iY_{i'}|\mathbb{1}\left(|F_0(Y_i)-F_0(Y_i')|<k^{-1/2+\epsilon/3}, F_0(Y_i)<k^{-1/2}+2\epsilon/3\right)\\
&\quad+k^{3/2}\big(\frac{1}{S_n^{(1)}}\big)^2\Big\{\frac{1}{S_n^{(2)}}\sum_{i\in A_n^{(1)}}Y_i^2\Big\}\times 2k^2n^{-1}\left\|c\right\|_\infty.
\end{aligned}
$$

Taking the expectation with respect to the $Y_i, i\in A_n^{(1)}$ and since $\mathbb{E}\left(|Y_i|^p\right)<\infty$, using holder inquality with p and $q=p/(p-1)$:

$$
\begin{aligned}
\mathbb{E}\left(A_{1n}^{(1)}(x)\right) &\leq 2\frac{k^{7/2}n^{-1}}{S_n^{(2)}}\left\|c\right\|_\infty\mathbb{E}\left(|Y_i|\right)^{1/p}\\
&\quad\times\mathbb{P}\left(|F_0(Y_i)-F_0(Y_{i'})|<k^{-1/2+\epsilon/3}, F_0(Y_i)<k^{-1/2+2\epsilon/3}\right)^{(p-1)/p}\\
&\quad+\mathcal{O}\left(k^{7/2}n^{-3}\right)\qquad\mathbb{P}_\xi-\text{a.s}\\
&\leq 2\frac{k^{7/2}n^{-1}}{S_n^{(2)}}\left\|c\right\|_\infty\mathbb{E}\left(|Y_i|\right)^{1/p}\times k^{\frac{p-1}{p}(-1+\epsilon)} \qquad (35)
\end{aligned}
$$

Next using lemma 13 again :

$$
\begin{aligned}
A_{1n}^{(2)}(x) &= k^{5/2}\big(\frac{1}{S_n^{(1)}}\big)^2 \frac{1}{S_n^{(2)}} \sum_{i,i' \in A_n^{(1)}} Y_i Y_{i'} \mathbb{1}\left(|F_0(Y_i) - F_0(Y_i')| < k^{-1/2+\epsilon/3}, F_0(Y_i) > k^{-1/2+2\epsilon/3}\right) \\
&\quad \times \sum_{l=0}^{k-1} \Delta_1 c(F_1(x), \frac{h+1}{k}) \times P'_{k-1,l}(\eta_{n,i}) P'_{k-1,l}(\eta_{n,i'}) \\
&\quad \times \big[F_{0n}(Y_i) - F_0(Y_i)\big] \times \big[F_{0n}(Y_{i'}) - F_0(Y_{i'})\big] \\
&\quad + k^{5/2}\big(\frac{1}{S_n^{(1)}}\big)^2 \frac{1}{S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i^2 \sum_{l=0}^{k-1} \Delta_1 c(F_1(x), \frac{h+1}{k}) \times P'_{k-1,l}(\eta_{n,i})^2 \times \big[F_{0n}(Y_i) - F_0(Y_i)\big]^2 \\
&\leq 2k^{3/2-2\epsilon/3} n^{-1} k^{3/2} \big(\frac{1}{S_n^{(1)}}\big)^2 \frac{1}{S_n^{(2)}} \|c\|_\infty \\
&\quad \times \sum_{i \neq i' \in A_n^{(1)}} |Y_i Y_{i'}| \mathbb{1}\left(|F_0(Y_i) - F_0(Y_i')| < k^{-1/2+\epsilon/3}, F_0(Y_i) < k^{-1/2} + 2\epsilon/3\right) \\
&\quad + k^{3/2}\big(\frac{1}{S_n^{(1)}}\big)^2 \Big\{\frac{1}{S_n^{(2)}} \sum_{i \in A_n^{(1)}} Y_i^2\Big\} \times 2k^2 n^{-1} \|c\|_\infty . \qquad (36)
\end{aligned}
$$

Thus

$$
\begin{aligned}
\mathbb{E}\left(A_{1n}^{(2)}(x)\right) &\leq 2\frac{k^{3-2\epsilon/3}}{S_n^{(2)}} \|c\|_\infty \mathbb{E}\left(Y^p\right)^{1/p} \\
&\quad \times \mathbb{P}\left(|F_0(Y_i) - F_0(Y_{i'})| < k^{-1/2+\epsilon/3}, F_0(Y_i) > k^{-1/2+2\epsilon/3}\right)^{(p-1)/p} \\
&\quad + \mathcal{O}\left(k^{7/2} n^{-3}\right) \qquad \mathbb{P}_\xi - \text{a.s} \\
&\leq 2\frac{k^{3-2\epsilon/3}}{S_n^{(2)}} \|c\|_\infty \mathbb{E}\left(Y^p\right)^{1/p} k^{\frac{p-1}{p}(\epsilon/3-1/2)} + \mathcal{O}\left(k^{7/2} n^{-3}\right) \quad \mathbb{P}_\xi - \text{a.s} \qquad (37)
\end{aligned}
$$

Finally, lemma 1.5.3 of (Lorentz 1953) implies that as soon as $|F_0(Y_i) - F_0(Y_{i'})| > k^{-1/2+\epsilon/3}$ then

$$
\sum_{l=0}^{k} |P'_{k-1,l}(\eta_{n,i}) P'_{k-1,l}(\eta_{n,i'})| \sim \mathcal{O}\left(e^{-k^{2\epsilon/3}}\right). \qquad (38)
$$

Since

$$
\mathbb{E}\left(A_{1n}(x)\right) = \mathbb{E}\left(A_{1n}^{(1)}(x)\right) + \mathbb{E}\left(A_{1n}^{(2)}(x)\right) + \mathcal{O}\left(e^{-k^{2\epsilon/3}}\right)
$$

it follows from equations (35),(37) and (38) that

$$
\begin{aligned}
\mathbb{E}\left(A_{1n}(x)\right) &\leq 2\|c\|_\infty \mathbb{E}\left(Y^p\right)^{1/p} \frac{k^3 n^{-1}}{S_n^{(2)}} \{k^{1/2+\frac{p-1}{p}(-1+\epsilon)} + k^{\frac{p-1}{p}(\epsilon/3-1/2)}\} + \mathcal{O}\left(e^{-m^{2\epsilon/3}}\right) \\
&\quad + \mathcal{O}\left(k^{7/2} n^{-3}\right) \qquad \mathbb{P}_\xi - \text{a.s} \qquad (39)
\end{aligned}
$$

Since the norming factor is $n^{1/2}m^{-1/4}$, the latter is negligible as soon as $m = n^\alpha$ with

$$\alpha \quad < \quad \min\Big\{(3 - \frac{p-1}{p})^{-1}, 2(5 - \frac{p-1}{p})^{-1}\Big\} \tag{40}$$

with $p \geq 3$ and by the choice of

$$\epsilon < \min\Big\{\frac{p}{p-1}(\alpha^{-1} - 3) + 1, \frac{2}{3}\big[\frac{p}{p-1}(2\alpha^{-1} - 5) + 1\big]\Big\}. \tag{41}$$

The condition $p \geq 3$ guaranties that the bound on α is greater than 2/5. Next, using the mean value theorem, since C as bounded density and the Bernstein weights sum to one combined with lemma 13 :

$$\begin{aligned} |A_{2n}| \quad &\leq \quad (\frac{1}{S_n^{(1)}})^2 \frac{1}{S_n^{(2)}} \Big\{ \sum_{i \in A_n^{(1)}} |Y_i| \, \|c\|_\infty \sum_{l=0}^{k-1} |P'_{k-1,l}(\eta_{n,i})| \Big\} \\ &\qquad \times \Big\{ \sum_{i' \in A_n^{(1)}} |Y_{i'}| \, \|c\|_\infty \sum_{l=0}^{k-1} |P'_{k-1,l}(\eta_{n,i'})| \Big\} \\ &\qquad \times \big[\sup_y |F_{0n}(y) - F_0(y)|\big]^2 \\ &\leq \quad \frac{k^2}{S_n^{(2)}} \Big\{ \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} |Y_i| \Big\}^2 \times \mathcal{O}\left(n^{-1}\right) \sim \mathcal{O}_p\left(k^2 n^{-2}\right) \qquad \mathbb{P}_\xi - \text{a.s.} \end{aligned} \tag{42}$$

Applying the same strategy to $B_n(x)$ concludes the negligibility of $m_n^{(\xi 1)}(x) - \widetilde{m}_n^{(\xi 1)}(x)$. $\widetilde{m}_n^{(\xi 1)}(x)$ does not involve any estimated marginals in the Bernstein weights. The next step is to show that the random variable $D_n^{(\xi 1)}(x) = \widetilde{m}_n^{(\xi 1)}(x) - \overline{m}_n^{(\xi 1)}(x)$ is asymptotically negligible, where $\overline{m}_n^{(\xi 1)}(x) = \mathbb{E}_{A_n^{(1)}}\big(\widetilde{m}_n^{(\xi 1)}(x)\big)$

First notice that one can rewrite

$$D_n^{(\xi 1)}(x) \quad = \quad \frac{k^2}{S_n^{(2)}} \sum_{j \in A_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1(\kappa_j(\frac{l+1}{k}, \frac{h+1}{k}) - C(\frac{l+1}{k}, \frac{h+1}{k})) M_{k-1,l}(n) P_{k-1,h}(F_1(x))$$

where

$$M_{k-1,l}(n) \quad = \quad \frac{1}{S_n^{(1)}} \sum_{i \in A_n^{(1)}} Y_i P_{k-1,l}(F_0(Y_i)) - \mathbb{E}\left(Y P_{k-1,l}(F_0(Y))\right). \tag{43}$$

Now

$$\begin{aligned} \mathbb{E}_{A_n^{(1)}}\left(D_n^{(\xi 1)}(x)^2\right) \quad &= \quad \frac{k^4}{S_n^{(2)}} \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1 C(\frac{l+1}{k}, \frac{h+1}{k}) M_{k-1,l}(n)^2 P_{k-1,h}(F_1(x))^2 \\ &\qquad + \frac{k^4}{S_n^{(2)}} \Big\{ \sum_{l=0}^{k-1} \sum_{h=0}^{k-1} \Delta_1 C(\frac{l+1}{k}, \frac{h+1}{k}) M_{k-1,l}(n) P_{k-1,h}(F_1(x)) \Big\}^2. \end{aligned} \tag{44}$$

We need to investigate how big can be the $M_{k-1,l}(n)$'s. The expectation of $M_{k-1,l}(n)$ is 0 for all l and

$$\begin{aligned}\mathbb{E}\left(M_{k-1,l}(n)^2\right) &= \frac{1}{n}\left(\mathbb{E}\left(Y^2 P_{k-1,l}(F_0(Y))^2\right) - \mathbb{E}\left(Y P_{k-1,l}(F_0(Y))\right)^2\right. &(45)\\ &\leq \frac{1}{n}\left(\mathbb{E}\left(Y^2\right) + \mathbb{E}\left(|Y|\right)^2\right). &(46)\end{aligned}$$

Since C admits a bounded density c, and since lemma 2 of (Janssen, Swanepoel, and Ververbeke 2014) implies

$$\sum_{h=0}^{k-1} P_{k-1,h}(F_1(x))^2 \sim \mathcal{O}\left(k^{-1/2}\right) \tag{47}$$

one deduces that

$$\mathbb{E}_{A_n^{(1)}}\left(D_n^{(\xi 1)}(x)^2\right) \sim \mathcal{O}_p\left(\frac{k^{5/2}}{n^2}\right) + \mathcal{O}_p\left(\frac{k^2}{n^2}\right) \qquad \mathbb{P}_\xi - \text{a.s.} \tag{48}$$

Again the norming factor being $n^{1/2}k^{-1/4}$ the latter is negligible if $k \sim n^\alpha$, $\alpha < 1/2$. It only remains to show that the random variable $\overline{m}_n^{(\xi 1)}(x)$ is normally distributed. To do this we calculate the Lyapunov ratio.

Clearly by Fubini's theorem $\mathbb{E}\left(\overline{m}_n^{(\xi 1)}(x)\right) = 0$. Now

$$\begin{aligned}\mathbb{E}\left(\overline{m}_n^{(\xi 1)}(x)^2\right) &= \frac{k^4}{S_n^{(2)}} \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k}, \frac{h+1}{k}\right)\left(\int_0^1 F_0^{-1}(u) P_{k-1,l}(u)du\right)^2 P_{k-1,h}(F_1(x))^2\\ &\quad - \frac{k^4}{S_n^{(2)}}\left\{\sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k}, \frac{h+1}{k}\right)\left(\int_0^1 F_0^{-1}(u) P_{k-1,l}(u)du\right) P_{k-1,h}(F_1(x))\right\}^2\\ &= V_n(x) + W_n(x).\end{aligned}$$

If $\int_0^1 |F_0^{-1}(u)|du < \infty$ then by the dominated convergence theorem:

$$\begin{aligned}W_n(x) &= \frac{k^4}{S_n^{(2)}}\left\{\int_0^1 F_0^{-1}(u) \sum_{l=0}^{k-1}\sum_{h=0}^{k-1} \Delta_1 C\left(\frac{l+1}{k}, \frac{h+1}{k}\right) P_{l,k-1}(u) P_{h,k-1}(F_1(x))du\right\}^2\\ &\leq \frac{1}{S_n^{(2)}}||c||_\infty^2 \left(\int_0^1 |F_0^{-1}|(u)du\right)^2 + O\left(\frac{k^{-2}}{n}\right) \qquad \mathbb{P}_\xi - \text{a.s.} &(49)\end{aligned}$$

On the other hand one has

$$V_n(x) = V_{1n}(x) + V_{2n}(x),$$

where

$$V_{1n}(x) = \frac{k^4}{S_n^{(2)}} \int_0^1 \int_0^1 \sum_{l,h=0}^{k-1} \Delta_1 C(\frac{l+1}{k}, \frac{h+1}{k})(F_0^{-1}(u))^2 P_{l,k-1}(u) du P_{l,k-1}(v) dv P_{h,k-1}(F_1(x))^2,$$

and

$$V_{2n}(x) = \frac{k^4}{S_n^{(2)}} \int_0^1 \int_0^1 \sum_{l,h=0}^{k-1} \Delta_1 C(\frac{l+1}{k}, \frac{h+1}{k}) F_0^{-1}(u)(F_0^{-1})'(\tilde{u})(v-u) P_{l,k-1}(u) du P_{l,k-1}(v) dv P_{h,k-1}(F_1(x))^2.$$

Now because, $k\int_0^1 P_{l,k-1}(v)dv = 1$ and $\Delta_1 C(\frac{l+1}{k}, \frac{h+1}{k}) = \frac{1}{k^2} c\left(\frac{l}{k}, \frac{h}{k}\right) + \frac{1}{k^2}$ (see (Sancetta and Satchell 2004)), we have

$$V_{1n}(x) \quad = \quad \frac{k^{1/2}}{2\sqrt{\pi F_1(x)(1-F_1(x))}} \int_0^1 \left(F_0^{-1}(u)\right)^2 c(u, F_1(x)) du + o(k^{1/2})$$

and because $(F_0^{-1})'$ and c are finite, we have

$$\begin{aligned}
|V_{2n}(x)| &\leq \frac{k^2}{S_n^{(2)}} \|c\|_\infty \int_0^1 \int_0^1 |F_0^{-1}(u)||v-u| \sum_{l=0}^{k-1} P_{l,k-1}(u) du P_{l,k-1}(v) dv \sum_{h=0}^{k-1} P_{h,k-1}(F_1(x))^2 \\
&= \frac{k^2}{S_n^{(2)}} \|c\|_\infty \int_0^1 \int_0^1 |F_0^{-1}(u)||v-u| (\sum_{k\in A} + \sum_{k\in A^c}) P_{l,k-1}(u) du P_{l,k-1}(v) dv \sum_{h=0}^{k-1} P_{h,k-1}(F_1(x))^2
\end{aligned}$$

where

$$A = \{l, |l/k-u| \leq k^{-\delta} \, \text{and} \, |l/k - v| \leq k^{-\delta}, \text{where} \, 1/3 < \delta < 1/2\}$$

Hence, using again theorem 1.5.3 of (Lorentz 1953) one has

$$\begin{aligned}
|V_{2n}(x)| &\leq \frac{k}{S_n^{(2)}} \|c\|_\infty k^{-\delta} \left\{ \sum_{k\in A} \int_0^1 |F_0^{-1}(u)| P_{l,k-1}(u) du \right\} \left\{ \int_0^1 k P_{l,k-1}(v) dv \right\} \sum_{h=0}^{k-1} P_{h,k-1}(F_1(x))^2 \\
&\quad + \mathcal{O}\left(e^{-(1/2-\delta)^2}\right) \\
&= O\left(\frac{k^{1/2-\delta}}{n}\right) \qquad \mathbb{P}_\xi - \text{a.s.}
\end{aligned}$$

Therefore,

$$\mathbb{E}\left(\overline{m}_n^{(\xi_1)}(x)^2\right) = \frac{k^{1/2} n^{-1}}{2\sqrt{\pi F_1(x)(1-F_1(x))}} \int_0^1 (F_0^{-1}(u))^2 c(u, F_1(x)) du + o\left(\frac{k^{1/2}}{n}\right) + O\left(\frac{k^{1/2-\delta}}{n}\right) \quad \mathbb{P}_\xi - \text{a.s.}$$

Similarly using again lemma 2 of (Janssen, Swanepoel, and Ververbeke 2014) one gets

$$\mathbb{E}\left(\left(\frac{k^2}{S_n^{(2)}}\int_0^1 F_0^{-1}(u)\sum_{l=0}^{k-1}\sum_{h=0}^{l-1}\Delta_1(\kappa_j(\frac{l+1}{k},\frac{h+1}{k})-C(\frac{l+1}{k},\frac{h+1}{k}))P_{k-1,l}(u)P_{k-1,h}(F_1(x))\right)^4\right)$$
$$\sim \mathcal{O}\left(\frac{k^{3/2}}{n^4}\right) \qquad \mathbb{P}_\xi - \text{a.s} \tag{50}$$

Thus the Lyapunov condition is satisfied since the Lyapunov ratio is of order $\mathcal{O}\left(\frac{k^{1/2}}{n}\right)$.

6.4 Asymptotic negligibility of $m_n^{(\xi 2)}(x)$

Using similar arguments, we can show that $\sqrt{nk^{-1/2}}m_n^{\xi_2}(x)$ is negligible.

6.5 Asymptotic representation of $m_n^{(\xi 3)}(x)$

Next, similar derivations, and similar treatments of the term (II) as in the proof of (Janssen, Swanepoel, and Ververbeke 2014) leads to

$$\sqrt{nk^{-1/2}}m_n^{(\xi_3)}(x) = \sqrt{nk^{-1/2}}\overline{m}_n^{(\xi_1,3)}(x) + o(1) \qquad \mathbb{P}_\xi - \text{ a.s}$$

where

$$\begin{aligned}\overline{m}_n^{(\xi_1,3)}(x) &= \frac{k}{S_n^{(2)}}\sum_{i\in A_n^{(2)}}\int_0^1\sum_{h=0}^{k-1}\left\{\mathbb{1}\left(\frac{h}{k}\le V_i\le\frac{h+1}{k}\right)-\frac{1}{k}\right\}\\ &\times \sum_{l=0}^{k} C^{(2)}(\frac{l}{k},\frac{h}{k-1})F_0^{-1}(u)P'_{k,l}(u)P_{k-1,l}(F_1(x))du.\end{aligned}$$

As for $\overline{m}_n^{(\xi_1)}(x)$, one shows that the Lyapunov condition is satisfied for $\sqrt{nk^{-1/2}}\overline{m}_n^{(\xi_1,3)}(x)$. One then concludes that

$$\sqrt{nk^{-1/2}}(\hat{m}_n^{(\xi)}(x)-m(x)) = \sqrt{nk^{-1/2}}\left\{\overline{m}_n^{(\xi_1)}(x)+\overline{m}_n^{(\xi_1,3)}(x)\right\}.$$

This consist representation consist in a sum of independent terms which can be shown, with similar development to be normally distributed with mean 0 and variance σ^2.

References

Bouezmarni, T., J. El ghouch, and A. Taamouti (2014). Test for conditional independence via copula: Application to causality test. *Journal of Business and Economic Statistics 30*, 275–287.

Bouezmarni, T. and J. Rombouts (2014). Bernstein estimator for unbounded copula densities. *Statistics and Risk Modelling 30*, 343–360.

Bouezmarni, T., J. Rombouts, and T. Taamouti (2010). Asymptotic properties of the Bernstein density copula estimator for α-mixing data. *Journal of Multivariate Analysis 101*, 1–10.

Chen, S. X. and T. Huang (2007). Nonparametric estimation of copula functions for dependent modeling. *Canadian Journal of Statistics 35*, 265–282.

Crane, G. J. and J. Van Der Hoek (2008). Conditional Expectation Formulae for Copulas. *Australian and New Zealand Journal of Statistics 50*, 53–67.

Dette, H., R. Van Hecke, and S. Volgushev (2014). Some comments on copula-based regression. *Journal of American Society of Statistics 109*, 1319–1324.

Gijbels, I. and J. Mielniczuk (1990). Estimating the density of a copula function. *Communications in Statistics - Theory and Methods 19*, 445–464.

Janssen, P., J. Swanepoel, and N. Ververbeke (2014). A note on the asymptotic behavior of the Bernstein estimator of the copula density. *Journal of Multivariate Analysis 122*, 480–487.

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

Leong, Y. K. and E. A. Valdez (2005). Claims Prediction with Dependence using Copula Models. Technical report.

Lorentz, G. (1953). *Bernstein Polynomials*. Mathematical Expositions. University of Toronto Press.

Nelsen, R. (2006). *An Introduction to Copulas*. Springer, New York.

Noh, H., E. G. Anouar, and T. Bouezmarni (2013). Copula-based regression estimation and inference. *Journal ofAmerican Society of Statistics 108*(502), 676–688.

Omelka, M., I. Gijbels, and N. Veraverbeke (2009). Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing. *The Annals of Statistics 37*(5B), 3023–3058.

Sancetta, A. and S. Satchell (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory 20*, 535–562.

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

Sungur, E. A. (2005). Some Observations on Copula Regression Functions. *Communications in Statistics-Theory and Methods 34*, 1967–1978.