Bernstein estimator for conditional copulas

Peer-reviewed author version

VERAVERBEKE, Noel (2024) Bernstein estimator for conditional copulas. In:
Statistical papers (1988) = Statistische Hefte (1988),.

DOI: 10.1007/s00362-024-01573-x
Handle: http://hdl.handle.net/1942/43299

BERNSTEIN ESTIMATOR FOR CONDITIONAL COPULAS

Noël VERAVERBEKE
University of Hasselt, Belgium
North-West University, South Africa
noel.veraverbeke@uhasselt.be

February 22, 2024

Abstract

The use of Bernstein polynomials in smooth nonparametric estimation of copulas has been well established in recent years. Their good properties in terms of bias and variance are well known. In this note we generalize some of the asymptotic theory to conditional copulas, that is where the dependence structure between the variables changes with a value of a random covariate. We obtain asymptotic representations and asymptotic normality for a conditional copula.

Keywords

Asymptotic properties, Bernstein estimation, Conditional Copula, Covariate.

1 Introduction

Consider a bivariate random vector (Y_1, Y_2) and a random covariate X. The joint conditional distribution function of (Y_1, Y_2) is denoted by $H_x(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2 \mid X = x)$ and the marginal conditional distribution functions by $F_{1x}(y_1) = P(Y_1 \leq y_1 \mid X = x)$ and $F_{2x}(y_2) = P(Y_2 \leq y_2 \mid X = x)$.
According to Sklar's theorem (see e.g. Nelsen (2006), Patton (2006), we have that

$$H_x(y_1, y_2) = C_x(F_{1x}(y_1), F_{2x}(y_2))$$

where C_x is a conditional copula function. To guarantee uniqueness of C_x, we assume that F_{1x} and F_{2x} are continuous. The conditional copula function can be expressed as $C_x(u_1, u_2) = H_x(F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2))$ $(0 \leq u_1, u_2 \leq 1)$, where $F_{1x}^{-1}(u) = \inf\{y : F_{1x}(y) \geq u\}$ is the quantile function of F_{1x} and F_{2x}^{-1} is the quantile function of F_{2x}.
For the estimation of C_x we assume that we have a sample $(Y_{11}, Y_{21}, X_1), \ldots, (Y_{1n}, Y_{2n}, X_n)$ from (Y_1, Y_2, X). Based on this we have the following empirical estimator for $H_x(y_1, y_2)$:

$$H_{xh}(y_1, y_2) = \sum_{i=1}^{n} w_{ni}(x; h_n) I(Y_{1i} \leq y_1, Y_{2i} \leq y_2)$$

where $\{w_{ni}(x; h_n)\}$ is a sequence of weights that smooth in the X-direction. Here we will take Nadaraya-Watson weights given by

$$w_{ni}(x; h_n) = K\left(\frac{X_i - x}{h_n}\right) / \sum_{j=1}^{n} K\left(\frac{X_j - x}{h_n}\right)$$

for $i = 1, \ldots, n$. The function K is a probability density function (kernel) and$\{h_n\}$ is a sequence of positive constants, tending to 0 as n tends to infinity (bandwidth). We assume that K has finite support $[-L, L]$ for some $L > 0$, $\mu_1(K) = \int uK(u)du = 0$, K is of bounded variation and Lipschitz of order 1.

The marginals of H_{xh} are denoted by $F_{1xh}(y_1) = H_{xh}(y_1, +\infty) = \sum_{i=1}^{n} w_{ni}(x; h_n) I(Y_{1i} \leq y_1)$ and $F_{2xh}(y_2) = \sum_{i=1}^{n} w_{ni}(x; h_n) I(Y_{2i} \leq y_2)$. We then define the empirical copula estimator by plugging in empirical versions for H_x, F_{1x} and F_{2x} in the expression for $C_x(u_1, u_2)$:

$$C_{xh}(u_1, u_2) = H_{xh}(F_{1xh}^{-1}(u_1), F_{2xh}^{-1}(u_2)) \quad (0 \leq u_1, u_2 \leq 1).$$

Conditional copulas are needed to model a dependence structure between Y_1 and Y_2 which changes with the value x of *a* random covariate X. An important application of conditional copulas is that they enable to define and study estimators for conditional versions of the classical association measures. The reason is that for example Kendall's tau and Spearman's rho can be expressed as functionals of

the underlying conditional copula. This topic has been the subject of Gijbels et al (2011) and Veraverbeke et al (2011) with theory, simulations and examples.
The purpose of this note is to study the Bernstein conditional copula estimator, which is a smoothened version of the estimator C_{xh}. It is defined as

$$C_{xm}(u_1,u_2)=\sum_{k=0}^{m}\sum_{\ell=0}^{m}C_{xh}\left(\frac{k}{m},\frac{\ell}{m}\right)P_{m,k}(u_1)P_{m,\ell}(u_2)$$

where, for $k=0,1,\ldots,m$:

$$P_{m,k}(u)=\binom{m}{k}u^k(1-u)^{m-k}.$$

The natural number m is called the order and for asymptotics it will be assumed that $m\to\infty$ as $n\to\infty$.
The idea of using Bernstein polynomials in copula estimation started with Sancetta and Satchell (2004). In several papers since then, it has been shown that Bernstein estimators have good asymptotic bias and variance properties compared to the classical kernel smoothers. In a series of papers by Janssen, Swanepoel and Veraverbeke (2012, 2014, 2016), the asymptotic theory and finite sample results have been obtained for copulas, copula densities and copula derivatives.

2 Overview

In this paper we first prove an asymptotic representation for $C_{xm}(u_1,u_2)$ as a weighted sum plus a bias term and a remainder term (Section 3). Our second result is the asymptotic normality of $C_{xm}(u_1,u_2)$ (Section 4). Section 5 contains the proof of a technical lemma. An important ingredient in the proof of these results is of course the asymptotic representation for the empirical copula estimator C_{xh}. The latter result has been obtained in Veraverbeke et al (2011) with a remainder term $o_P((nh_n)^{-1/2})$. Although this suffices for the asymptotic normality of $C_{xm}(u_1,u_2)$, we prefer to use the more recent version of Veraverbeke (2023) with a stronger remainder term $O((nh_n)^{-3/4}(\log n)^{3/4})$ a.s.
All our results will require conditions on the bandwidth h_n and on the order m and also some of the regularity conditions that we have listed in below.
We remark that derivatives will be denoted as for example:
$\dot{F}_{1x}(t)=\frac{\partial}{\partial x}F_{1x}(t)$, $C_x^{(1)}(u_1,u_2)=\frac{\partial}{\partial u_1}C_x(u_1,u_2)$, $C_x^{(1,2)}(u_1,u_2)=\frac{\partial^2}{\partial u_1\partial u_2}C_x(u_1,u_2)$,
$\dot{C}_x(u_1,u_2)=\frac{\partial}{\partial x}C_x(u_1,u_2)$, etc.

Conditions

(C) C_z has bounded third order partial derivatives for all $(u_1,u_2)\in[0,1]^2$ and z in a neighborhood of x.

(F1) The density f_X of X is uniformly continuous and strictly positive at x.

(F2) f_X'' is finite at x.

(Y) $\dot{F}_{1x}(t)$, $\ddot{F}_{1x}(t)$, $\dot{F}_{2x}(t)$, $\ddot{F}_{2x}(t)$ are continuous in (x,t).

(Z) $F_{1z}(F_{1x}^{-1}(u))$ and $F_{2z}(F_{2x}^{-1}(u))$ are Lipschitz continuous in u for z in a neighborhood of x.

(R) $\dot{C}_z(u_1,u_2)$ and $\ddot{C}_z(u_1,u_2)$ exist and are continuous for all (z,u_1,u_2) for z in a neighborhood of x.

Remark 1
The most restrictive condition is condition (C) on the partial derivatives of the conditional copula. It is satisfied for copula families like Frank, Farlie-Gumbel-Morgenstern, Ali-Mikhael-Haq. See Nelsen (2006) for the precise definitions. But there are a number of classial copulas like normal, Student t, Gumbel and Clayton that are ruled out by this condition (C) because their partial derivatives are not continuous at certain boundary points of the unit square.
The theoretial inconvenience is well known and some remedies have been suggested, see Omelka et al (2009) and Segers (2012). It is an interesting open question whether our result can be proven under a set of less restrictive conditions as in Segers (2012).

3 Asymptotic representation for the Bernstein estimator of a conditional copula

Defining

$$B_{xm}(u_1,u_2)=\sum_{k=0}^{m}\sum_{\ell=0}^{m} C_x\left(\frac{k}{m},\frac{\ell}{m}\right)P_{m,k}(u_1)P_{m,\ell}(u_2)$$

we have by the theorem of Weierstrass that $\lim_{m\to\infty} B_{xm}(u_1,u_2)=C_x(u_1,u_2)$, uniformly in $(u_1,u_2)\in[0,1]^2$. The reason is that every copula is continuous on $[0,1]^2$ (see Nelsen (2006)).
We write

$$\begin{aligned} C_{xm}(u_1,u_2)-C_x(u_1,u_2) &= [C_{xm}(u_1,u_2)-B_{xm}(u_1,u_2)] \\ &+ [B_{xm}(u_1,u_2)-C_x(u_1,u_2)]. \end{aligned} \tag{1}$$

The second term in (1) ia a first bias term in our representation. Under condition (C) we have

$$\begin{aligned} & B_{xm}(u_1,u_2)-C_x(u_1,u_2) \\ =\ & \frac{1}{2m}\{u_1(1-u_1)C_x^{(1,1)}(u_1,u_2)+u_2(1-u_2)C_x^{(2,2)}(u_1,u_2)\}+o(m^{-1}). \end{aligned} \tag{2}$$

To deal with the term $C_{xm} - B_{xm}$ in (1), we first rewrite C_x in a more convenient representation in terms of uniforms.
Define, for $i = 1, \ldots, n$

$$U_{1i} = F_{1x}(Y_{1i}), \quad U_{2i} = F_{2x}(Y_{2i}).$$

We have $P(U_{1i} \le u_1 \mid X_i = x) = u_1$, $P(U_{2i} \le u_2 \mid X_i = x) = u_2$ and $P(U_{1i} \le u_1, U_{2i} \le u_2 \mid X_i = x) = C_x(u_1, u_2)$.
Now recall the asymptotic representation for $C_{xh} - C_x$ in Veraverbeke (2023): if $\frac{\log n}{nh_n} \to 0$, $\frac{nh_n^5}{\log n} = O(1)$ and regularity conditions (C), (F1), (Y), (R) hold, then uniformly in $(u_1, u_2) \in [0,1]^2$:

$$\begin{aligned} & C_{xh}(u_1,u_2) - C_x(u_1,u_2) \\ = \; & \sum_{i=1}^{n} w_{ni}(x;h_n)\xi_i(u_1,u_2) + O((nh)^{-3/4}(\log n)^{3/4}) \text{ a.s.} \end{aligned}$$

where

$$\begin{aligned} & \xi_i(u_1,u_2) = I(U_{1i} \le u_1, U_{2i} \le u_2) - C_x(u_1,u_2) \\ & -C_x^{(1)}(u_1,u_2)\{I(U_{1i} \le u_1) - u_1\} - C_x^{(2)}(u_1,u_2)\{I(U_{2i} \le a_2) - u_2\}. \end{aligned}$$

This representation for $C_{xh} - C_x$ leads to a representation for $C_{xm} - B_{xm}$:

$$\begin{aligned} & C_{xm}(u_1,u_2) - B_{xm}(u_1,u_2) \\ = \; & \sum_{i=1}^{n} w_{ni}(x;h_n)Z_{in}(u_1,u_2) + O((nh_n)^{-3/4}(\log n)^{3/4}) \end{aligned} \tag{3}$$

a.s., where

$$\begin{aligned} Z_{in}(u_1,u_2) = \sum_{k=0}^{m}\sum_{\ell=0}^{m} & \left\{ I(U_{1i} \le \frac{k}{m}, U_{2i} \le \frac{\ell}{m}) - C_x\left(\frac{k}{m}, \frac{\ell}{m}\right) \right. \\ & -C_x^{(1)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\left(I(U_{1i} \le \frac{k}{m}) - \frac{k}{m}\right) \\ & \left. -C_x^{(2)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\left(I(U_{2i} \le \frac{\ell}{m}) - \frac{\ell}{m}\right)\right\} P_{m,k}(u_1)P_{m,\ell}(u_2). \end{aligned}$$

Denote

$$\alpha_m(X_i, x, u_1, u_2) = E(Z_{in}(u_1,u_2) \mid X_i = x) \tag{4}$$

$$\beta_m(X_i, x, u_1, u_2) = \mathrm{Var}(Z_{in}(u_1,u_2) \mid X_i = x). \tag{5}$$

These two quantities can be further expanded. This is summarized in the following lemma.

Lemma
For z in a neighborhood of x, we have for $m \to \infty$,

(i) $\alpha_m(z, x, u_1, u_2) = \alpha(z, x, u_1, u_2) + O(m^{-1/2})$
where

$$\begin{aligned} &\alpha(z, x, u_1, u_2) = C_z(u_1, u_2) - C_x(u_1, u_2) \\ &-C_x^{(1)}(u_1, u_2)[F_{1z}(F_{1x}^{-1}(u_1) - u_1] \\ &-C_x^{(2)}(u_1, u_2)[F_{2z}(F_{2x}^{-1}(u_2) - u_2]. \end{aligned} \tag{6}$$

(ii) $\beta_m(z, x, u_1, u_2) = \beta(z, x, u_1, u_2) + O(m^{-1/2})$ where

$$\begin{aligned} &\beta(z, x, u_1, u_2) = C_z(u_1, u_2)(1 - C_z(u_1, u_2)) \\ &+F_{1z}(F_{1x}^{-1}(u_1))(1 - F_{1z}(F_{1x}^{-1}(u_1)))C_z^{(2)}(u_1, u_2) \\ &+F_{2z}(F_{2x}^{-1}(u_2))(1 - F_{2z}(F_{2x}^{-1}(u_2)))C_z^{(2)}(u_1, u_2) \\ &-2(1 - F_{1z}(F_{1x}^{-1}(u_1))C_z(u_1, u_2)C_x^{(1)}(u_1, u_2) \\ &-2(1 - F_{2z}(F_{2x}^{-1}(u_2))C_z(u_1, u_2)C_x^{(2)}(u_1, u_2) \\ &+2C_x^{(1)}(u_1, u_2)C_x^{(2)}(u_1, u_2)[C_z(u_1, u_2) - u_1 u_2]. \end{aligned} \tag{7}$$

4 Asymptotic normality

We can now state and prove the following result on the asymptotic distribution of the Bernstein conditional copula estimator.

Theorem
Assume

$$\begin{aligned} &h_n \to 0, m \to \infty \\ &n h_n m^{-2} \to c^2 \geq 0 \\ &n h_n^5 \to \widetilde{c}^2 \geq 0. \end{aligned}$$

Also assume the regularity conditions (C), (F1), (F2), (Y), (Z), (R).
Then, as $n \to \infty$,

$$\begin{aligned} &\sqrt{n h_n}(C_{xm}(u_1, u_2) - C_x(u_1, u_2)) \\ &\xrightarrow{d} N(c b_x(u_1, u_2) + \widetilde{c}\widetilde{b}_x(u_1, u_2); \frac{\sigma_x^2(u_1, u_2)}{f_X(x)}\|K\|_2^2) \end{aligned}$$

where

$$b_x(u_1,u_2) = \frac{1}{2}\left\{u_1(1-u_1)C_x^{(1,1)}(u_1,u_2) + u_2(1-u_2)C_x^{(2,2)}(u_1,u_2)\right\}$$

$$\widetilde{b}_x(u_1,u_2) = \mu_2(K)\left\{\frac{1}{2}\ddot{\alpha}(x,x,u_1,u_2) + \dot{\alpha}(x,x,u_1,u_2)\frac{f'_X(x)}{f_X(x)}\right\}$$
$$\dot{\alpha}(x,x,u_1,u_2) = \dot{C}_x(u_1,u_2) - C_x^{(1)}(u_1,u_2)[F_{1X_i}(F_{1x}^{-1}(u_1)-u_1] - C_x^{(2)}(u_1,u_2)[F_{2X_i}(F_{2x}^{-1}(u_2)-u_2]$$
$$\ddot{\alpha}(x,x,u_1,u_2) = \ddot{C}_x(u_1,u_2) - C_x^{(1)}(u_1,u_2)[F_{1X_i}(F_{1x}^{-1}(u_1)-u_1] - C_x^{(2)}(u_1,u_2)[F_{2X_i}(F_{2x}^{-1}(u_2)-u_2]$$

$$\sigma_x^2(u_1,u_2) = Var\{I(U_1\le u_1, U_2\le u_2) - C_x(u_1,u_2)$$

$$-C_2^{(1)}(u_1,u_2)(I(U_1\le u_1)-u_1) - C_x^{(2)}(u_1,u_2)(I(U_2\le u_2)-u_2)\}$$

$\mu_2(K) = \int u^2K(u)du, \|K\|_2^2 = \int K^2(u)du.$

Proof

We look at the limiting distribution of the main term in (3), which we rewrite as follows:

$$\sqrt{nh_n}\sum_{i=1}^n w_{ni}(x;h_n)Z_{in}(u_1,u_2) = \sqrt{nh_n}\frac{\widehat{T}_n(x)}{\widehat{f}_n(x)}$$
$$= \sqrt{nh_n}\left[\frac{\widehat{T}_n(x)}{\widehat{f}_n(x)} - \frac{E(\widehat{T}_n(x))}{E(\widehat{f}_n(x))}\right] + \sqrt{nh_n}\left[\frac{E(\widehat{T}_n(x)}{E(\widehat{f}_n(x))}\right] \qquad (8)$$

where
$$\widehat{T}_n(x) = \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)Z_{in}(u_1,u_2)$$

and
$$\widehat{f}_n(x) = \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{X_i-x}{h_n}\right)$$

is the Parzen kernel density estimator.
We have, with α_m as in (4), and using the Lemma:

$$\begin{aligned}
E(\widehat{T}_n(x)) &= \frac{1}{h_n}E\left(K\left(\frac{X-x}{h_n}\right)\alpha_m(X,x,u_1,u_2)\right) \\
&= \int K(u)\alpha(x+h_nu,x,u_1,u_2)f_X(x+h_nu)du + O(m^{-1/2}) \\
&= \mu_2(K)h_n^2\left\{\frac{1}{2}\ddot{\alpha}(x,x,u_1,u_2)f_X(x)\right. \\
&\quad \left.+\dot{\alpha}(x,x,u_1,u_2)f'_X(x)\right\} + o(h_n^2) + O(m^{-1/2}).
\end{aligned}$$

Here $\dot{\alpha}(z,x,u_1,u_2)$ and $\ddot{\alpha}(z,x,u_1,u_2)$ are the first and second derivative of $\alpha(z,x,u_1,u_2)$ with respect to z. With β_m as in (5), and using the Lemma:

$$\begin{aligned} Var(\widehat{T}_n(x)) &= \frac{1}{nh_n^2} E\left(K^2\left(\frac{X-x}{h_n}\right)\beta_m(X,x,u_1,u_2)\right) + O\left(\frac{h_n^4}{n}+\frac{1}{nm}\right) \\ &= \frac{1}{nh_n}\int K^2(u)\beta(x+h_n u,x,u_1,u_2) f_X(x+h_n u)du \\ &\quad + O\left(\frac{h_n^4}{n}+\frac{1}{nm}+\frac{m^{-1/2}}{nh_n}\right) \\ &= \frac{1}{nh_n}\|K\|_2^2 \beta(x,x,u_1,u_2) f_X(x) + o\left(\frac{1}{nh_n}\right). \end{aligned}$$

From (7):

$$\begin{aligned} \beta(x,x,u_1,u_2) &= C_x(u_1,u_2)(1-C_x(u_1,u_2)) + u_1(1-u_1)C_x^{(1)}(u_1,u_2) \\ &\quad + u_2(1-u_2)C_x^{(2)}(u_1,u_2) - 2u_1(1-u_1)C_x(u_1,u_2)C_x^{(1)}(u_1,u_2) \\ &\quad - 2u_2(1-u_2)C_x(u_1,u_2)C_x^{(2)}(u_1,u_2) \\ &\quad + 2C_x^{(1)}(u_1,u_2)C_x^{(2)}(u_1,u_2)(C_x(u_1,u_2)-u_1u_2) \\ &= Var\{I(U_1\le u_1, U_2\le u_2) - C_x(u_1,u_2) \\ &\quad - C_x^{(1)}(u_1,u_2)(I(U_1\le u_1)-u_1) - C_x^{(2)}(u_1,u_2)(I(U_2\le u_2)-u_2)\} \\ &= \sigma_x^2(u_1,u_2). \end{aligned}$$

From (9) it follows that the last term in (8)

$$\sim \sqrt{nh_n^5}\,\mu_2(K)\left\{\frac{1}{2}\ddot{\alpha}(x,x,u_1,u_2) + \dot{\alpha}(x,x,u_1,u_2)\frac{f_X'(x)}{f_X(x)}\right\}$$

which leads to the bias term $\widetilde{cb}_x(u_1,u_2)$. For the first term in (8) we have by linearization that it has the same asymptotic distribution as

$$\frac{1}{f_X(x)}\sqrt{nh_n}(\widehat{T}_n(x) - E(\widehat{T}_n(x))) - \frac{E(\widehat{T}_n(x))}{f_X^2(x)}\sqrt{nh_n}(\widehat{f}_n(x) - E(\widehat{f}_n(x)).$$

Since $E(\widehat{T}_n(x)) \to 0$ and $\sqrt{nh_n}(\widehat{f}_n(x) - E(\widehat{f}_n(x)) = O_P(1)$, the limiting distribution is governed by

$$\frac{1}{f(x)}\sqrt{nh_n}(\widehat{T}_n(x) - E(\widehat{T}_n(x))).$$

$\widehat{T}_n(x)$ is a sum $\sum_{i=1}^{n} W_{ni}$ of a double array of random variables $W_{ni} = \frac{1}{nh_n} K\left(\frac{X_i - x}{h_n}\right) Z_{in}(u_1, u_2)$. By checking the Liapunov condition

$$\sum_{i=1}^{n} E[(W_{ni} - E(W_{ni})^4]/(Var \sum_{i=1}^{n} W_{ni})^2 = O\left(\frac{1}{nh_n}\right) \to 0,$$

we obtain the asymptotic normality of $\widehat{T}_n(x)$. Combining this with (1), (2) and (3) proves the theorem.

Remark 2
In Leblanc (2012) there are explicit expressions for the quantities $S_m(u)$ and $R_{1,m}(u)$ that appear in the proof of our Lemma (Section 5). This enables to calculate the order term $O(m^{-1/2})$ in (7) in an explicit way (in the same way as in Janssen et al (2012)).
Consequently, the asymptotic variance of $C_{xm}(u_1, u_2)$ is given by

$$\frac{1}{nh}\sigma_x^2(u_1, u_2) - \frac{m^{-1/2}}{nh} V_x(u_1, u_2) + o\left(\frac{m^{-1/2}}{nh}\right) \tag{9}$$

where

$$\begin{aligned} V_x(u_1, u_2) &= C_x^{(1)}(u_1, u_2)(1 - C_x^{(1)}(u_1, u_2)) \left(\frac{u_1(1-u_1)}{\pi}\right)^{1/2} \\ &+ C_x^{(2)}(u_1, u_2)(1 - C_x^{(2)}(u_1, u_2)) \left(\frac{u_2(1-u_2)}{\pi}\right)^{1/2}. \end{aligned}$$

The term with the minus sign in (9) clearly shows that there is a gain in the asymptotic variance of the estimator $C_{xm}(u_1, u_2)$ compared to that of $C_{xh}(u_1, u_2)$.

5 Proof of the lemma

(i) We have:

$$\begin{aligned} &\alpha_m(X_i, x, u_1, u_2) \\ &= \sum_{k=0}^{m} \sum_{\ell=0}^{m} \left\{ C_{X_i}\left(\frac{k}{m}, \frac{\ell}{m}\right) - C_x\left(\frac{k}{m}, \frac{\ell}{m}\right) \right. \\ &-C_x^{(1)}\left(\frac{k}{m}, \frac{\ell}{m}\right) \left[F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right) - \frac{k}{m}\right] \\ &\left. - C_x^{(2)}\left(\frac{k}{m}, \frac{\ell}{m}\right) \left[F_{2X_i}\left(F_{2x}^{-1}\left(\frac{\ell}{m}\right)\right) - \frac{\ell}{m}\right] \right\} P_{m,k}(u_1) P_{m,\ell}(u_2). \end{aligned}$$

Therefore,

$$\alpha_m(z,x,u_1,u_2)-\alpha(z,z,u_1,u_2)$$

$$=\sum_{k=0}^{m}\sum_{\ell=0}^{m}\{[C_z(u_1,u_2)-C_x(u_1,u_2)]$$

$$-\left[C_z\left(\frac{k}{m},\frac{\ell}{m}\right)-C_x(u_1,u_2)\right]$$

$$-\left[C_z\left(\frac{k}{m},\frac{\ell}{m}\right)-C_x(u_1,u_2)\right]$$

$$-\left[C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)F_{1z}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right)-C_x^{(1)}(u_1,u_2)F_{1z}(F_{1x}^{-1}(u_1))\right]$$

$$+\left[C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)\frac{k}{m}-C_x^{(1)}(u_1,u_2)u_1\right]$$

$$-\left[C_x^{(2)}\left(\frac{k}{m},\frac{\ell}{m}\right)F_{2z}\left(F_{2x}^{-1}\left(\frac{\ell}{m}\right)\right)-C_x^{(2)}(u_1,u_2)F_{2z}(F_{2x}^{-1}(u_2))\right]$$

$$+\left.\left[C_x^{(2)}\left(\frac{k}{m},\frac{\ell}{m}\right)\frac{\ell}{m}-C_x^{(2)}(u_1,u_2)u_2\right]\right\}P_{m,k}(u_1)P_{m,\ell}(u_2).$$

With conditions (C) and (Z), we can use Lipschitz continuity to obtain that

$$\begin{aligned}&\alpha_m(z,x,u_1,u_2)-\alpha(z,z,u_1,u_2)\\&=O\left(\sum_{k=0}^{m}\mid\frac{k}{m}-u_1\mid P_{m,k}(u_1)+\sum_{\ell=0}^{m}\mid\frac{\ell}{m}-u_2\mid P_{m,\ell}(u_2)\right)\\&=O(m^{-1/2})\text{ see Bojanic and Cheng, 1989).}\end{aligned}$$

(ii) We have:

$$
\begin{aligned}
&\beta_m(X_i, x, u_1, u_2) \\
&= E[(Z_{in}(u_1,u_2) - E(Z_{in}(u_1,u_2) \mid X_i))^2 \mid X_i] \\
&= \sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{k'=0}^{m}\sum_{\ell'=0}^{m} E\left\{\left[I\left(U_{1i} \leq \frac{k}{m}, U_{2i} \leq \frac{\ell}{m}\right) - C_{Xi}\left(\frac{k}{m}, \frac{\ell}{m}\right)\right.\right. \\
&-C_x^{(1)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\left(I\left(U_{1i} \leq \frac{k}{m}\right) - F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right)\right) \\
&\left. - C_x^{(2)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\left(I\left(U_{2i} \leq \tfrac{\ell}{m}\right) - F_{2X_i}\left(F_{2x}^{-1}\left(\frac{\ell}{m}\right)\right)\right)\right] \\
&\qquad \left. \times\, [\text{same with } k', \ell'] \mid X_i \right\} P_{m,k}(u_1)P_{m,\ell}(u_2)P_{m,k'}(u_1)P_{m,\ell'}(u_2).
\end{aligned}
$$

Working out the expectation gives 9 terms and each of these terms is a quadruple sum. This long calculation is similar to the one in the proof of Lemma 3 (iii) in the Appendix of Janssen et al (2012). We show how the calculation works for one of these terms. We denote this term by T.

$$
\begin{aligned}
T = \quad & \sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{k'=0}^{m}\sum_{\ell'=0}^{m} E\left\{ C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right) C_x^{(1)}\left(\frac{k'}{m},\frac{\ell'}{m}\right)\right. \\
& \left[I\left(U_{1i} \le \frac{k}{m}\right) - F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right)\right]\left[I\left(U_{1i} \le \frac{k'}{m}\right) - F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k'}{m}\right)\right)\right] \\
& P_{m,k}(u_1)P_{m,\ell}(u_2)P_{m,k'}(u_1)P_{m,\ell'}(u_2) \\
= \quad & \sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{k'=0}^{m}\sum_{\ell'=0}^{m} C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right) C_x^{(1)}\left(\frac{k'}{m},\frac{\ell'}{m}\right) \\
& \left[F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k\wedge k'}{m}\right) - F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right) F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k'}{m}\right)\right)\right)\right] \\
& P_{m,k}(u_1)P_{m,\ell}(u_2)P_{m,k'}(u_1)P_{m,\ell'}(u_2) \\
= \quad & \sum_{k=0}^{m}\sum_{\ell=0}^{m}\left(C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)\right)^2 F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right) P_{m,k}^2(u_1)P_{m,\ell}^2(u_2) \\
& -\left(\sum_{k=0}^{m}\sum_{\ell=0}^{m} C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right) F_{1X_i}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right) P_{m,k}(u_1)P_{m,\ell}(u_2)\right)^2 \\
& +\sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{\substack{k'=0\\ k'\neq k}}^{m}\sum_{\substack{\ell'=0\\ \ell'\neq \ell}}^{m} C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right) C_x^{(1)}\left(\frac{k'}{m},\frac{\ell'}{m}\right) F_{1X_i}\left(F_{1X_i}^{-1}\left(\left(\frac{k\wedge k'}{m}\right)\right)\right) \\
& P_{m,k}(u_1)P_{m,\ell}(u_2)P_{m,k'}(u_1)P_{m,\ell'}(u_2) \\
& +\sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{\substack{k'=0\\ k'\neq k}}^{m} C_x^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right) C_x^{(1)}\left(\frac{k'}{m},\frac{\ell'}{m}\right) F_{1X_i}\left(F_{1X_i}^{-1}\left(\left(\frac{k\wedge k'}{m}\right)\right)\right) \\
& P_{m,k}(u_1)P_{m,\ell}^2(u_2)P_{m,k'}(u_1) \\
& +\sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{\substack{\ell'=0\\ \ell'\neq \ell}}^{m} C_x^{(1)}\left(\frac{k}{m},\frac{\ell'}{m}\right) C_x^{(1)}\left(\frac{k}{m},\frac{\ell'}{m}\right) F_{1X_i}\left(F_{1X_i}^{-1}\left(F_{1x}^{-1}\left(\frac{k}{m}\right)\right)\right) \\
& P_{m,k}^2(u_1)P_{m,\ell}(u_2)P_{m,\ell'}(u_2) \\
& := (T_1)+(T_2)+(T_3)+(T_4)+(T_5).
\end{aligned}
$$

For (T_2) we have: $(T_2) = -(C_x^{(1)}(u_1,u_2)F_{1X_i}(F_{1x}^{-1}(u_1))^2 + O(m^{-1})$.
For (T_1) we make Taylor expansion around (u_1,u_2) which gives:

$$\begin{aligned}(T_1) &= (C_x^1(u_1,u_2))^2 F_{1X_i}(F_{1x}^{-1}(u_1))S_m(u_1)S_m(u_2)\\ &\quad +O(S_m(u_1)I_m(u_2) + S_m(u_2)I_m(u_1))\end{aligned}$$

where $S_m(u) = \sum_{k=0}^{m} P_{m,k}^2(u)$ and $I_m(u) = \sum_{k=0}^{m} \mid \frac{k}{m} - u \mid P_{m,k}^2(u)$.
An expansion of (T_3) gives $(T_3) = (T_3') + (T_3'')$, with

$$\begin{aligned}(T_3') &= (C_x^{(1)}(u_1,u_2))F_{X_i}(F_{1x}^{-1}(u_1))(1 - S_m(u_1))(1 - S_m(u_2)).\\ (T_3'') &= (C_x^{(1)}(u_1,u_2))^2 \sum_{k=0}^{m}\sum_{\ell=0}^{m}\sum_{k'\neq k}^{m}\sum_{\ell'\neq \ell}^{m}\left(\frac{k\wedge k'}{m} - u_1\right)P_{m,k}(u_1)P_{m,\ell}(u_2)P_{m,k'}(u_1)P_{m,\ell'}(u_2)\\ &= (C_x^{(1)}(u_1,u_2))^2(1 - S_m(u_2))2R_{1,m}(u_1)\end{aligned}$$

where $R_{1,m}(u) = m^{-1} \sum_{\substack{k=0\\ k<k'}}^{m} (k - mu)P_{m,k}(u)P_{m,\ell}(u)$.

For term (T_4):

$$\begin{aligned}(T_4) &= (C_x^{(1)}(u_1,u_2))^2 F_{1X_i}(F_{1x}^{-1}(u_1))S_m(u_2)(1 - S_m(u_1))\\ &\quad +2(C_x^{(1)}(u_1,u_2))^2 S_m(u_2)R_{1,m}(u_1).\end{aligned}$$

For term (T_5):

$$\begin{aligned}(T_5) &= (C_x^{(1)}(u_1,u_2))^2 F_{1X_i}(F_{1x}^{-1}(u_1))S_m(u_2)(1 - S_m(u_2))\\ &\quad +O(I_m(u_1)).\end{aligned}$$

From Leblanc (2012) we have that $S_m(u)$ and $R_{1,m}(u)$ are $O(m^{1/2})$ and $I_m(u) = O(m^{-3/4})$.
Therefore,

$$(T) = (C_x^{(1)}(u_1,u_2))^2 F_{1X_i}(F_{1x}^{-1}(u_1))(1 - F_{X_i}(F_{1x}^{-1}(u_1))) + O(m^{-1/2}).$$

A similar treatment of the 8 other terms leads to

$$\begin{aligned}&\beta_m(X_i,x,u_1,u_2) = C_{X_i}(u_1,u_2)(1 - C_{X_i}(u_1,u_2))\\ &+F_{1X_i}(F_{1x}^{-1}(u_1))(1 - F_{1X_i}(F_{1x}^{-1}(u_1)))C_{X_i}^2(u_1,u_2)\\ &+F_{2X_i}(F_{2x}^{-1}(u_2))(1 - F_{2X_i}(F_{2x}^{-1}(u_1)))C_{X_i}^2(u_1,u_2)\\ &-2(1 - F_{1X_i}(F_{1x}^{-1}(u_1)))C_{X_i}(u_1,u_2)C_x^{(1)}(u_1,u_2)\\ &-2(1 - F_{2X_i}(F_{2x}^{-1}(u_2)))C_{X_i}(u_1,u_2)C_x^{(2)}(u_1,u_2)\\ &+2(C_x^{(1)}(u_1,u_2)C_x^{(2)}(u_1,u_2)[C_{X_i}(u_1,u_2) - u_1u_2] + O(m^{-1/2}).\end{aligned}$$

Hence,

$$\beta_m(z,x,u_1,u_2) = \beta(z,x,u_1,u_2) + O(m^{-1/2}).$$

References

[1] Bojanic, R. and Cheng, F. (1989). Rate of convergence of Bernstein polynomials for functions with derivatives of bounded variation. J. Math. Anal. Appl., **141**, 136-151.

[2] Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulas, association measures and their applications. Computational Statistics and Data Analysis, **55**, 1919-1932.

[3] Janssen, P., Swanepoel, J. and Veraverbeke, N. (2012). Large sample behavior of the Bernstein copula estimator. J. Statist. Planning and Inference, **142**, 1189-1197.

[4] Janssen, P., Swanepoel, J. and Veraverbeke, N. (2014). A note on the asymptotic behavior of the Bernstein estimator of the copula density. J. Multivariate Analysis, **124**, 480-487.

[5] Janssen, P., Swanepoel, J. and Veraverbeke, N. (2016). Bernstein estimation for a copula derivative with application to conditional distribution and regression functionals. TEST, **25**, 351-374.

[6] Leblanc, A. (2012). On estimating distribution functions using Bernstein polynomials. Annals of the Institute of Statistical Mathematics, **64**, 919-943.

[7] Nelsen, R.B. (2006). An introduction to copulas. Second edition. Springer-Verlag, New York.

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

[9] Patton, A.J. (2006). Modelling asymmetric exchange rate dependence. International Economic Review, **47**, 527-556.

[10] Sancetta, A. and Satchell, S. (2004). The Bernstein copula and its application to modeling and approximations of multivariate distributions. Econometric Theory, **20**, 535-562.

[11] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. Bernoulli, **18**, 764-782.

[12] Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. Scandinavian J. of Statistics, **38**, 766-780.

[13] Veraverbeke, N. (2023). Almost sure asymptotic representation for the conditional copula estimator. Communications in Statistics. Theory and Methods. DOI 10.1080/03610926.2023.2300318.