Journal of Multivariate Analysis 124 (2014) 480–487

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Note(s)

A note on the asymptotic behavior of the Bernstein estimator of the copula density

Paul Janssen [a], Jan Swanepoel [b], Noël Veraverbeke [a,b,*]

[a] Hasselt University, Center for Statistics, Agoralaan, Gebouw D, 3590 Diepenbeek, Belgium

[b] North-West University, Potchefstroom Campus, Potchefstroom, South Africa

ARTICLE INFO

Article history:
Received 21 March 2013
Available online 28 October 2013

AMS 2010 subject classifications:
primary 62G05
62G07
secondary 62G20

Keywords:
Asymptotic normality
Bernstein estimator
Copula density

ABSTRACT

Copulas and their corresponding densities are functions of a multivariate joint distribution and the one-dimensional marginals. Bernstein estimators have been used as smooth nonparametric estimators for copulas and copula densities. The purpose of this note is to study the asymptotic distributional behavior of the Bernstein estimator of a copula density. Compared to the existing results, our general theorem does not assume known marginals. This makes our theorem applicable for real data.

1. Introduction

Copulas and copula densities are building blocks to study the dependence between components of a random vector. For many problems it is not evident how to select a parametric family of copulas to describe the data at hand and, hence, nonparametric estimation of the copula and the copula density is an option. The Bernstein estimator of the copula and the copula density is one such nonparametric estimator that received attention in recent papers. The authors in [7] study the asymptotic distributional behavior of the Bernstein estimator of the copula. Assuming the marginals to be known, the results on the asymptotic distributional behavior of the Bernstein estimator of the copula density are given in [10,2,3]. In this note we show that the central limit theorem for the Bernstein estimator of the copula density is valid without imposing this assumption.

To facilitate the discussion we first collect some preliminary definitions. For simplicity we consider bivariate random vectors. Given a random vector (X, Y) with the joint distribution function H and marginal distribution functions F and G, there exists a bivariate distribution function C on $[0, 1]^2$ [12] such that

$$H(x, y) = C(F(x), G(y)).$$

C is the copula corresponding to H; see [9] for a detailed discussion on copulas. We assume throughout that F and G are continuous, which implies that C is unique and that

$$C(u, v) = H\{F^{-1}(u), G^{-1}(v)\}$$

with F^{-1} and G^{-1} being the usual quantile functions.

* Corresponding author at: Hasselt University, Center for Statistics, Agoralaan, Gebouw D, 3590 Diepenbeek, Belgium.
E-mail addresses: paul.janssen@uhasselt.be (P. Janssen), jan.swanepoel@nwu.ac.za (J. Swanepoel), noel.veraverbeke@uhasselt.be (N. Veraverbeke).

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.10.009

A copula $C(u, v)$, being continuous on $[0, 1]^2$ [9, p. 11], can be approximated by a Bernstein copula [10] in the following way:

$$B_m(u, v) = \sum_{k=0}^{m}\sum_{\ell=0}^{m} C\left(\frac{k}{m}, \frac{\ell}{m}\right) P_{m,k}(u)P_{m,\ell}(v)$$

with for $k = 0, 1, \ldots, m$ and $0 \le u \le 1$

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

the binomial probabilities. We indeed have

$$\lim_{m\to\infty} B_m(u, v) = C(u, v).$$

The Bernstein estimator of the copula $C(u, v)$ is obtained by replacing $C\left(\frac{k}{m}, \frac{\ell}{m}\right)$ in the corresponding Bernstein copula by $C_n\left(\frac{k}{m}, \frac{\ell}{m}\right)$ with

$$C_n(u, v) = H_n\{F_n^{-1}(u), G_n^{-1}(v)\}$$

where for a random sample $(X_1, Y_1), \ldots (X_n, Y_n)$ from H,

$$H_n(x, y) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x, Y_i \le y)$$

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x) \qquad G_n(y) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \le y).$$

This gives

$$C_{m,n}(u, v) = \sum_{k=0}^{m}\sum_{\ell=0}^{m} C_n\left(\frac{k}{m}, \frac{\ell}{m}\right) P_{m,k}(u)P_{m,l}(v).$$

We call m the order of the estimator; the order will typically depend on n and we have $m \to \infty$ as $n \to \infty$.

The copula density corresponding to C is denoted as c and is given by (if it exists)

$$c(u, v) = \partial^2 C(u, v)/\partial u \partial v.$$

The corresponding Bernstein estimator of the copula density is

$$c_{m,n}(u, v) = \sum_{k=0}^{m}\sum_{\ell=0}^{m} C_n\left(\frac{k}{m}, \frac{\ell}{m}\right) P'_{m,k}(u)P'_{m,l}(v) \tag{1}$$

with $P'_{m,k}(u)$ being the derivative with respect to u.

In terms of these definitions we can detail the contribution we make in this note. Our asymptotic normality result for $c_{m,n}(u, v)$ extends the results in [10,3]. They assume that the marginals F and G are known. Both papers indeed replace the pair $(F_n^{-1}(u), G_n^{-1}(v))$ by $(F^{-1}(u), G^{-1}(v))$. However, given a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from H, a fully nonparametric Bernstein estimator for the copula density should use F_n and G_n rather than F and G. We show that the asymptotic normality result remains valid if we take this double stochastic nature of the problem into account, i.e., the stochastics coming from H_n and the stochastics coming from F_n and G_n. This covers the real data situation. Two references to rank-based inference procedures for copulas are [5,6].

The paper is organized as follows. In Section 2 we state and prove the asymptotic normality result. In Section 3 we use asymptotic bias and variance expressions to derive an optimal order m. Appendix contains two interesting properties of binomial probabilities that are used in the proof of the theorem.

2. Asymptotic normality of the Bernstein estimator of the copula density

Our main result reads as follows.

Theorem. *Assume*

(1) *The order $m > 0$ depends on n such that $m = o\{n^{1/2}(\log n)^{-1}(\log\log n)^{-1/2}\}$.*
(2) *The second order partial derivatives $C^{(1,1)}$, $C^{(2,2)}$ and $C^{(1,2)} = c$ of C exist and are continuous on $[0, 1]^2$.*

Then, for $0 < u, v < 1$, *as* $n \to \infty$,

$$\left(\frac{n}{m}\right)^{1/2} \{c_{m,n}(u, v) - b_m(u, v)\} \xrightarrow{d} N\left(0; c(u, v)\frac{1}{4\pi}\frac{1}{\sqrt{u(1-u)v(1-v)}}\right)$$

where

$$\begin{aligned} b_m(u, v) &= m^2 \sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}\left\{C\left(\frac{k+1}{m}, \frac{\ell+1}{m}\right) - C\left(\frac{k}{m}, \frac{\ell+1}{m}\right)\right. \\ &\quad \left. - C\left(\frac{k+1}{m}, \frac{\ell}{m}\right) + C\left(\frac{k}{m}, \frac{\ell}{m}\right)\right\} P_{m-1,k}(u)P_{m-1,\ell}(v) \\ &= m^2 \sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}\left\{\int_{k/m}^{(k+1)/m}\int_{\ell/m}^{(\ell+1)/m} c(s, t)ds\, dt\right\} P_{m-1,k}(u)P_{m-1,\ell}(v). \end{aligned} \tag{2}$$

Remark 1. As shown in the proof below, the bias term $b_m(u, v)$ is equal to $c(u, v)+O(m^{-1})$. Therefore the centering $b_m(u, v)$ may be replaced by the copula density $c(u, v)$ if $n/m^3 \to 0$. This combined with condition (1) shows that a good choice for m is given by $m = n^\alpha$ with $\frac{1}{3} < \alpha < \frac{1}{2}$.

Remark 2. The proof of the theorem will make use of a stochastic representation of the empirical copula process given in [13]. Conditions (1) and (2) are needed in that context. However the theorem also holds if condition (2) is replaced by the set of weaker conditions given in [11]. The strength of this representation is that it handles at once the double stochastic nature of the problem.

Proof. Recall a result of [13]: if the second order partial derivatives of C exist and are continuous on $[0, 1]^2$, then

$$\begin{aligned} \sup_{0\le u,v\le 1}&\left|C_n(u, v) - C(u, v) - \frac{1}{n}\sum_{i=1}^{n}\{I(U_i \le u, V_i \le v) - C(u, v)\}\right. \\ &\left. + C^{(1)}(u, v)\frac{1}{n}\sum_{i=1}^{n}\{I(U_i \le u) - u\} + C^{(2)}(u, v)\frac{1}{n}\sum_{i=1}^{n}\{I(V_i \le v) - v\}\right| \\ &= O\{n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}\} \quad \text{a.s., as } n \to \infty. \end{aligned}$$

Here $(U_1, V_1), \ldots, (U_n, V_n)$ is a random sample from C, $C^{(1)}(u, v) = \frac{\partial}{\partial u}C(u, v)$, $C^{(2)}(u, v) = \frac{\partial}{\partial v}C(u, v)$. Plugging in in (1) gives

$$\begin{aligned} c_{m,n}(u, v) &= \sum_{k=0}^{m}\sum_{\ell=0}^{m} C\left(\frac{k}{m}, \frac{\ell}{m}\right) P'_{m,k}(u)P'_{m,\ell}(v) \\ &\quad + \sum_{k=0}^{m}\sum_{\ell=0}^{m}\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i \le \frac{k}{m}, V_i \le \frac{\ell}{m}\right) - C\left(\frac{k}{m}, \frac{\ell}{m}\right)\right\} P'_{m,k}(u)P'_{m,\ell}(v) \\ &\quad - \sum_{k=0}^{m}\sum_{\ell=0}^{m} C^{(1)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i \le \frac{k}{m}\right) - \frac{k}{m}\right\} P'_{m,k}(u)P'_{m,\ell}(v) \\ &\quad - \sum_{k=0}^{m}\sum_{\ell=0}^{m} C^{(2)}\left(\frac{k}{m}, \frac{\ell}{m}\right)\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(V_i \le \frac{\ell}{m}\right) - \frac{\ell}{m}\right\} P'_{m,k}(u)P'_{m,\ell}(v) \\ &\quad + O\{mn^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}\} \quad \text{a.s.} \end{aligned}$$

uniformly in $(u, v) \in (0, 1)^2$.

The factor m in the order term appears because for each $0 < u < 1$,

$$\sum_{k=0}^{m}|P'_{m,k}(u)| \sim \sqrt{\frac{2}{\pi}}\frac{m^{1/2}}{\sqrt{u(1-u)}} = O(m^{1/2})$$

as $m \to \infty$. This relation is derived in Lemma 1 of Appendix.

Denote the above decomposition as

$$c_{m,n}(u, v) = b_m(u, v) + \text{(I)} - \text{(II)} - \text{(III)} + \text{(IV)}.$$

The term $b_m(u, v)$ is a deterministic bias term. By the continuity of c we have

$$\int_{k/m}^{(k+1)/m}\int_{\ell/m}^{(\ell+1)/m} c(s,t)ds\,dt = \frac{1}{m^2}c\left(\frac{k}{m-1}, \frac{\ell}{m-1}\right) + O\left(\frac{1}{m^3}\right) \tag{3}$$

uniformly in $k, \ell \le m-1$.

This gives

$$b_m(u, v) = c(u, v) + O\left(\frac{1}{m}\right).$$

The term (I) can be rewritten as

$$\begin{aligned}(\mathrm{I}) \;=\; & m^2\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(\frac{k}{m} < U_i \le \frac{k+1}{m}, \frac{\ell}{m} < V_i \le \frac{\ell+1}{m}\right)\right.\\ & \left. - P\left(\frac{k}{m} < U_i \le \frac{k+1}{m}, \frac{\ell}{m} < V_i \le \frac{\ell+1}{m}\right)\right\} P_{m-1,k}(u)P_{m-1,\ell}(v)\\ := & \sum_{i=1}^{n} Z_{\mathrm{in}}.\end{aligned}$$

We check the Liapunov condition for the array $\{Z_{\mathrm{in}}\}$ of independent random variables. Clearly $E(Z_{1n}) = 0$. Further

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} Z_{\mathrm{in}}\right) = nE(Z_{1n}^2) = \frac{m^4}{n}E & \left\{\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}\sum_{k'=0}^{m-1}\sum_{\ell'=0}^{m-1}\left\{I\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)\right.\right.\\ & - P\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)\Big\} \cdot \left\{I\left(\frac{k'}{m} < U_1 \le \frac{k'+1}{m}, \frac{\ell'}{m} < V_1 \le \frac{\ell'+1}{m}\right)\right.\\ & \left.\left. - P\left(\frac{k'}{m} < U_1 \le \frac{k'+1}{m}, \frac{\ell'}{m} < V_1 \le \frac{\ell'+1}{m}\right)\right\} \cdot P_{m-1,k}(u)P_{m-1,\ell}(v)P_{m-1,k'}(u)P_{m-1,\ell'}(v).\right\}\end{aligned}$$

Note that the product of the two indicators is zero if $k \ne k'$ or $\ell \ne \ell'$. From (3) we conclude that

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{n} Z_{\mathrm{in}}\right) = & \frac{m^4}{n}E\left\{\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} I\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)P_{m-1,k}^2(u)P_{m-1,\ell}^2(v)\right\}\\ & - 2\frac{m^4}{n}E\left\{\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} I\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)P_{m-1,k}(u)P_{m-1,\ell}(v)\right.\\ & \left.\times \sum_{k'=0}^{m-1}\sum_{\ell'=0}^{m-1} P\left(\frac{k'}{m} < U_1 \le \frac{k'+1}{m}, \frac{\ell'}{m} < V_1 \le \frac{\ell'+1}{m}\right)P_{m-1,k'}(u)P_{m-1,\ell'}(v)\right\}\\ & + \frac{m^4}{n}\left(\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} P\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)P_{m-1,k}(u)P_{m-1,\ell}(v)\right)^2\\ = & \frac{m^4}{n}\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} P\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)P_{m-1,k}^2(u)P_{m-1,\ell}^2(v)\\ & - \frac{m^4}{n}\left(\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} P\left(\frac{k}{m} < U_1 \le \frac{k+1}{m}, \frac{\ell}{m} < V_1 \le \frac{\ell+1}{m}\right)P_{m-1,k}(u)P_{m-1,\ell}(v)\right)^2\\ = & \frac{m^4}{n}\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}\left(\int_{k/m}^{(k+1)/m}\int_{\ell/m}^{(\ell+1)/m} c(s,t)ds\,dt\right)P_{m-1,k}^2(u)P_{m-1,\ell}^2(v) - \frac{1}{n}b_m^2(u, v)\\ = & \frac{m^2}{n}\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1} c\left(\frac{k}{m-1}, \frac{\ell}{m-1}\right)P_{m-1,k}^2(u)P_{m-1,\ell}^2(v)\end{aligned}$$

$$+O\left(\frac{m}{n}\right)\sum_{k=0}^{m-1}\sum_{\ell=0}^{m-1}P_{m-1,k}^2(u)P_{m-1,\ell}^2(v)-\frac{1}{n}\left\{c(u,v)+O\left(\frac{1}{m}\right)\right\}^2$$
$$\sim\ \frac{m}{n}c(u,v)\frac{1}{4\pi\sqrt{u(1-u)v(1-v)}}$$

by Lemma 3.1 in [1] or by Lemma 2 of Appendix with $\ell = 1$. Repeating the same steps as above and using Lemma 2 of Appendix for $\ell = 1, 2$ and 3 we obtain that $\sum_{i=1}^n E(Z_{\text{in}}^4) = O((m/n)^3)$. This gives that the Liapunov condition is satisfied. Indeed, $\sum_{i=1}^n E(Z_{\text{in}}^4)/(\text{Var}(\sum_{i=1}^n Z_{\text{in}}))^2 = O(m/n) = o(1)$, since $m/n \to 0$. Therefore,

$$\left(\frac{n}{m}\right)^{1/2}(I)\xrightarrow{d}N\left(0;c(u,v)\frac{1}{4\pi}\frac{1}{\sqrt{u(1-u)v(1-v)}}\right).$$

The term (II) can be rewritten as

$$(\text{II})=\sum_{k=0}^{m}\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i\le\frac{k}{m}\right)-\frac{k}{m}\right\}P'_{m,k}(u)\cdot\sum_{\ell=0}^{m}C^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)P'_{m,\ell}(v).$$

We have

$$\begin{aligned}\sum_{\ell=0}^{m}C^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)P'_{m,\ell}(v)&=m\sum_{\ell=0}^{m-1}\left\{C^{(1)}\left(\frac{k}{m},\frac{\ell+1}{m}\right)-C^{(1)}\left(\frac{k}{m},\frac{\ell}{m}\right)\right\}P_{m-1,\ell}(v)\\&=m\sum_{\ell=0}^{m-1}\int_{\ell/m}^{(\ell+1)/m}c\left(\frac{k}{m},t\right)dt\,P_{m-1,\ell}(v)\\&=m\sum_{\ell=0}^{m-1}\left\{\frac{1}{m}c\left(\frac{k}{m},\frac{\ell}{m-1}\right)+O(m^{-2})\right\}P_{m-1,\ell}(v)\\&=c\left(\frac{k}{m},v\right)+O(m^{-1})=c\left(\frac{k}{m-1},v\right)+O(m^{-1})\end{aligned}$$

uniformly in k and ℓ. Therefore

$$\begin{aligned}(\text{II})&=\sum_{k=0}^{m}\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i\le\frac{k}{m}\right)-\frac{k}{m}\right\}P'_{m,k}(u)\left\{c\left(\frac{k}{m-1},v\right)+O(m^{-1})\right\}\\&=m\sum_{k=0}^{m-1}\left[\left(c\left(\frac{k}{m-1},v\right)+O(m^{-1})\right)\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i\le\frac{k+1}{m}\right)-\frac{k-1}{m}\right\}\right.\\&\quad\left.-\left\{c\left(\frac{k}{m-1},v\right)+O(m^{-1})\right\}\frac{1}{n}\sum_{i=1}^{n}\left\{I\left(U_i\le\frac{k}{m}\right)-\frac{k}{m}\right\}\right]P_{m-1,k}(u)\\&=\sum_{i=1}^{n}\widetilde{Z}_{\text{in}}+O(m^{-1})\cdot\sum_{i=1}^{n}\widetilde{\widetilde{Z}}_{\text{in}}\end{aligned}$$

where

$$\widetilde{Z}_{\text{in}}=\frac{m}{n}\sum_{k=0}^{m-1}c\left(\frac{k}{m-1},v\right)\left\{I\left(\frac{k}{m}<U_i\le\frac{k+1}{m}\right)-\frac{1}{m}\right\}P_{m-1,k}(u)$$

and

$$\widetilde{\widetilde{Z}}_{\text{in}}=\frac{m}{n}\sum_{k=0}^{m-1}\left\{I\left(\frac{k}{m}<U_i\le\frac{k+1}{m}\right)-\frac{1}{m}\right\}P_{m-1,k}(u).$$

Calculations as before give that $E(\widetilde{Z}_{\text{in}}) = 0$, $\text{Var}(\sum_{i=1}^n \widetilde{Z}_{\text{in}}) = O(m^{1/2}/n)$, $\sum_{i=1}^n E(\widetilde{Z}_{\text{in}}^4) = O(m^{3/2}/n^3)$ and that the Liapunov ratio is $O(m^{1/2}/n) = o(1)$. Similarly for $\sum_{i=1}^n \widetilde{\widetilde{Z}}_{\text{in}}$.

Therefore

(II) and also (III) are $O_P(m^{1/4}/n^{1/2})$.

With the normalization of (I) we have

$$\left(\frac{n}{m}\right)^{1/2}(\mathrm{II}) = O_P\left(\frac{1}{m^{1/4}}\right) = o_P(1)$$

and similar for (III).

Finally note that $(n/m)^{1/2}(\mathrm{IV}) = o_P(1)$ because of condition (1). This proves the theorem.

Remark 3. The restriction to the bivariate case is not essential. It is clear that the asymptotic normality result can be proved along the same lines for the general d-dimensional case ($d \geq 2$) relying on the stochastic representation of a d-dimensional empirical copula [13,11]. The norming factor becomes $(nm^{-d/2})^{1/2}$and the asymptotic variance is given by

$$c(u_1, \ldots, u_d)(4\pi)^{-d/2}\frac{1}{\sqrt{u_1(1-u_1)\ldots u_d(1-u_d)}}$$

for $0 < u_1, \ldots, u_d < 1$.

Remark 4. The asymptotic normality result can be extended to the weak convergence of the process $(n/m)^{1/2}(c_{n,m}(u, v) - c(u, v))$ on $(0, 1)^2$. The key issue is the asymptotic tightness of the process $(n/m)^{1/2}\sum_{i=1}^{n} Z_{\mathrm{in}}$. This can be obtained by applying Theorem 2.7.1 in [14].

3. Asymptotic bias and optimal order

The bias term $b_m(u, v)$ in (2) can be further expanded under additional assumptions on the partial derivatives of c. For instance, if c has second order partial derivatives that are Lipschitz on $(0, 1)^2$, then

$$b_m(u, v) = c(u, v) + \frac{1}{2m}b(u, v) + o\left(\frac{1}{m}\right) \tag{4}$$

where

$$b(u, v) = u(1-u)c_{uu}(u, v) + v(1-v)c_{vv}(u, v) + (1-2u)c_u(u, v) + (1-2v)c_v(u, v) \tag{5}$$

and

$$c_u = \frac{\partial}{\partial u}c, \qquad c_v = \frac{\partial}{\partial v}c, \qquad c_{uu} = \frac{\partial^2}{\partial u^2}c, \qquad c_{vv} = \frac{\partial^2}{\partial v^2}c.$$

This follows by applying various Taylor expansions and using the fact that the first and second central moments of a binomial $(m-1, u)$ variable are 0 and $(m-1)u(1-u)$.

Expression (4) together with the asymptotic variance expression in the theorem leads to a formula for the asymptotic optimal choice of the order m. For the asymptotic mean squared error of $c_{m,n}(u, v)$ we have the expression

$$\frac{m}{n}\sigma^2(u, v) + \frac{1}{4m^2}b^2(u, v)$$

where

$$\sigma^2(u, v) = c(u, v)\frac{1}{4\pi}\frac{1}{\sqrt{u(1-u)v(1-v)}}$$

and $b(u, v)$ is given in (5).

Minimizing with respect to m gives

$$m_0 = m_0(u, v) = \left\{\frac{b^2(u, v)}{2\sigma^2(u, v)}\right\}^{1/3} n^{1/3}$$

and

$$\mathrm{AMSE}(c_{m_0,n}(u, v)) = 3\left\{\frac{b(u, v)\sigma^2(u, v)}{4}\right\}^{2/3} n^{-2/3}.$$

With the choice $m = n^{\alpha}$ with $\frac{1}{3} < \alpha < \frac{1}{2}$, as explained in Remark 1, we have that the optimal value m_0 is close to be included in this range.

Acknowledgments

The work was supported by the IAP Research Network P7/13 of the Belgian State (Belgian Science Policy). The second author thanks the National Research Foundation of South Africa for financial support. The third author acknowledges support from research grant MTM 2008-03129 of the Spanish Ministerio de Ciencia e Innovacion. He is also an extraordinary professor at the North-West University, Potchefstroom, South Africa.

The authors are grateful to the two referees and the associate editor for their valuable comments.

Appendix

In this appendix we prove two lemmas of independent interest that were used in the proof of the theorem.

Lemma 1. *For any $0 < u < 1$, as $m \to \infty$, we have*

$$\sum_{k=0}^{m} \mid P'_{m,k}(u) \mid \sim \sqrt{\frac{2}{\pi}} \frac{m^{1/2}}{\sqrt{u(1-u)}}.$$

Proof. By direct calculation we find

$$P'_{m,k}(u) = \frac{1}{u(1-u)} P_{m,k}(u)(k - mu).$$

Therefore

$$\begin{aligned} \sum_{k=0}^{m} \mid P'_{m,k}(u) \mid &= \frac{2}{u(1-u)} \sum_{k=[mu]+1}^{m} (k - mu) P_{m,k}(u) \\ &= \frac{2}{u(1-u)} ([mu]+1) \binom{m}{[mu]+1} u^{[mu]+1} (1-u)^{m-[mu]} \end{aligned}$$

by an identity in [8].

Now use Stirling's approximation for factorials and some algebra to obtain the result.

Lemma 2. *For any $0 < u < 1$, any $\ell = 1, 2, \ldots$, as $m \to \infty$, we have*

$$\left(\sqrt{2mu(1-u)}\right)^{\ell} \sum_{k=0}^{m} P_{m,k}^{\ell+1}(u) \to \phi_\ell(0, \ldots, 0),$$

where $\phi_\ell(x_1, \ldots, x_\ell)$ is the ℓ-dimensional normal $N_\ell(\mathbf{0}, \mathbf{\Sigma}^0)$ density with variance–covariance matrix $\mathbf{\Sigma}^0 = [\sigma_{ij}]$, with $\sigma_{ij} = 1$ if $i = j$ and $\sigma_{ij} = 1/2$ if $i \neq j$.

Proof. Let $\mathbf{A}_i = (A_{i1}, A_{i2}, \ldots, A_{i\ell})$, for $i = 1, \ldots, m$, and assume that for each $j = 1, \ldots, \ell$, $A_{1j}, A_{2j}, \ldots, A_{mj}$ are i.i.d. with $E(A_{1j}) = 0$ and $E(A_{1j}^2) = 1$. Suppose that A_{ij} is lattice with mass points $b, b \pm h, b \pm 2h, \ldots$. Define for $j = 1, \ldots, \ell$

$$S_{mj} = \sum_{i=1}^{m} A_{ij}, \qquad x_j = \frac{mb + kh}{\sqrt{m}},$$

for $k = 0, \pm 1, \pm 2, \ldots$. Mimicking the proof of Theorem 3 of Section XV.5 of [4] we arrive at the following multivariate extension of his one-dimensional local central limit theorem for lattice random variables:

$$\frac{m^{\ell/2}}{h^\ell} P\left(\frac{S_{mj}}{\sqrt{m}} = x_j, j = 1, \ldots, \ell\right) - \phi_\ell(x_1, \ldots, x_\ell) \to 0 \tag{6}$$

uniformly in $(x_1, \ldots, x_\ell)$, where ϕ_ℓ is the ℓ-dimensional $N_\ell(\mathbf{0}, \mathbf{\Sigma})$ density.

Now, suppose that $\{B_{ij},\ i = 1, \ldots, m;\ j = 1, \ldots, \ell\}$ are independent Bernoulli variables with parameter u and that the random vectors $\mathbf{B}_i = (B_{i1}, \ldots, B_{i\ell})$ are independent of the random vector $\mathbf{W} = (W_1, \ldots, W_m)$, where $W_1, \ldots, W_m$ are also independent Bernoulli variables with parameter u.

Furthermore, set $A_{ij} = (B_{ij} - W_i)/\sqrt{2u(1-u)}$, then $E(A_{ij}) = 0$, $E(A_{ij}^2) = 1$ and $\mathrm{Cov}(A_{ij}, A_{ik}) = 1/2$ for $j \neq k$; $i = 1, \ldots, m$. Also the span of A_{ij} is $h = 1/\sqrt{2u(1-u)}$. Conditioning on the event $\{\sum_{i=1}^{m} W_i = k\}$ it follows that

$$P\left(S_{mj} = 0, j = 1, \ldots, \ell\right) = \sum_{k=0}^{m} P_{m,k}^{\ell+1}(u). \tag{7}$$

Hence the lemma follows from (6) and (7) with $\mathbf{\Sigma} = \mathbf{\Sigma}^0$.

References

[1] G.J. Babu, A.J. Canty, Y.P. Chaubey, Application of Bernstein polynomials for smooth estimation of a distribution and density function, J. Statist. Plann. Inference 105 (2002) 377–392.
[2] G.J. Babu, Y.P. Chaubey, Smooth estimation of a distribution function and density function on a hypercube using Bernstein polynomials for dependent random vectors, Statist. Probab. Lett. 76 (2006) 959–969.
[3] T. Bouezmarni, J.V.K. Rombouts, A. Taamouti, Asymptotic properties of the Bernstein density copula estimator for α-mixing data, J. Multivariate Anal. 101 (2010) 1–20.
[4] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, John Wiley, New York, 1966.
[5] C. Genest, A. Favre, Everything you always wanted to know about copula modeling but were afraid to ask, J. Hydrol. Eng. 12 (2007) 347–368.
[6] C. Genest, J. Neslehova, Copulas and copula models, in: A.H. El-Shaarawi, W.W. Piergorsch (Eds.), second ed., in: Encyclopedia of Environmetrics, vol. 2, Wiley, Chichester, 2012, pp. 541–553.
[7] P. Janssen, J. Swanepoel, N. Veraverbeke, Large sample behavior of the Bernstein copula estimator, J. Statist. Plann. Inference 142 (2012) 1189–1197.
[8] N.L. Johnson, A note on the mean deviation of the binomial distribution, Biometrika 44 (1957) 532–533.
[9] R. Nelsen, An Introduction to Copulas, second ed., Springer, New York, 2006.
[10] A. Sancetta, S. Satchell, The Bernstein copula and its applications to modeling and approximations of multivariate distributions, Econometric Theory 20 (2004) 535–562.
[11] J. Segers, Asymptotics of empirical copula processes under non-restrictive smoothness assumptions, Bernoulli 18 (2012) 764–782.
[12] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229–231.
[13] W. Stute, The oscillation behavior of empirical processes: the multivariate case, Ann. Probab. 12 (1984) 361–379.
[14] A. van der Vaart, J.A. Wellner, Weak Convergence of Empirical Processes, Springer, New York, 2000.