Estimation of Multivariate Discrete Distributions through Bernstein Copulas

Fossaluza, V., Esteves, L. G., Pereira, C. A. B.
IME - USP

Abstract

Measuring dependence between random variables is one of the most celebrated problems in statistics, and, therefore, the determination of the joint distribution of the relevant variables is crucial. Recently, copulas have become an important tool for properly inferring the joint distribution of the variables of interest. Although most of the studies have dealt with the case of continuous variables, a few have focused on the treatment of discrete variables. This paper presents a nonparametric approach to the estimation of joint discrete distributions with bounded support using copulas and Bernstein polynomials.

1 Introduction

The association between random variables is a subject of interest in many scientific fields. The most complete way to characterize the association between random variables is the joint distribution of these random variables. Multivariate density functions, for absolutely continuous variables, and multivariate probability mass functions, for discrete variables, have become the focus of researchers interested in evaluating such association (see, for instance, dos Anjos et al., 2004; Joe, 1997; Nelsen, 2006).

The motivation for the present paper was a study performed by the Obsessive-Compulsive Spectrum Disorder Program of the Institute of Psychiatry, University of São Paulo Medical School. A set of 1001 consecutive adult outpatients diagnosed with primary obsessive-compulsive disorder (OCD) according to DSM-IV criteria (American Psychiatric Association, 1994) were recruited, and some of them were submitted to psychiatric treatment. The OCD severity was evaluated using the Yale-Brown Scale (Y-BOCS; Goodman et al., 1989b,a) at the beginning of the project. At the time the data records were accessed, only 213 patients participated in the re-evaluation using the scale. Y-BOCS is composed of two subscales: obsession (O) and compulsion (C), each assuming values in the set of integers $\{0, 1, \ldots, 20\}$. To measure the OCD severity, we consider the maximum value between the O and C subscale measures: the M-BOCS scale, given by max(O;C) (see the discussions in Pereira et al., 2011; Diniz et al., 2011).

Figure 1 illustrates the results of the initial and final M-BOCS scores for the 213 patients. Our first objective is to estimate the marginal distributions of the initial and final scores. For this purpose, all available information should be used: 1001 patients at the first evaluation and 213 at the end of the study. Using only the complete observations, without missing marginal values, we obtain only 213 pairs of measurements to be used in the estimation. This set of observations could be insufficient to a proper estimation since the sample space contains exactly 441 ($= 21^2$) points almost the double of the sample size: that is, at least half of the sample space will receive zero as estimates. Using the 1001 observations of one of the marginals we should have improved our estimates.

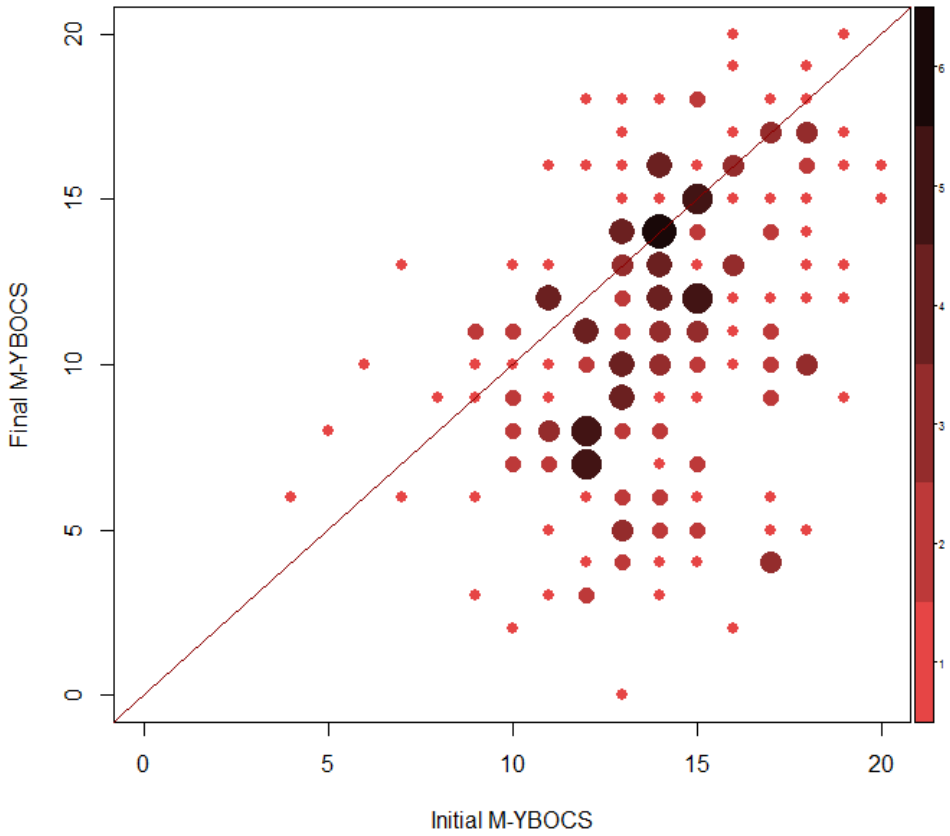

Figure 1: Initial X final OCD severity

The aim of the present paper is to introduce a method to estimate multivariate discrete probability mass functions in the presence of (marginal) missing data. For this purpose, we developed an estimation method that uses both empirical distribution functions and Bernstein polynomials. The procedure consists on the estimation of a smooth joint distribution function, followed by a method that transforms it into a discrete function, the estimated joint probability mass function. The new method is compared with alternative methods, both visually and by evaluating standard distances.

Section 2 describes the existing methods found in literature that will be considered for comparisons. Section 3 describes our estimator for the joint probability functions. Section 4 presents a discussion of the method and, using simulated samples and the OCD real example, it is compared to alternative ones. Finally, in Section 5, we present our final comments with considerations for future work.

2 Existing solutions

First, we introduce the mathematical setup. Let F be the unknown distribution function of a random vector $\boldsymbol{X}$ taking values in a subset of $\mathbb{R}^p$. A sample of size n of $\boldsymbol{X}$

is represented by $\boldsymbol{X_1}, \ldots, \boldsymbol{X}_n$, $\boldsymbol{X_i} = (X_{i1}, \ldots, X_{ip})$, $i = 1, \ldots, n$. That is, the $\boldsymbol{X_i}$'s are conditionally exchangeable for any given F. The observations of $\boldsymbol{X_i}$ are denoted by $\boldsymbol{x_i}$.

Assuming that the distribution F comes from a known family of distributions, we represent the statistical by $(\mathcal{X}, \mathcal{F}, \mathcal{P})$, where $\mathcal{X}$ is the sample space, $\mathcal{F}$ is a sigma-algebra of its subsets and $\mathcal{P} = \{P(\cdot|\theta) : \theta \in \Theta\}$ is a family of distributions indexed by the parameter θ belonging to the parameter space Θ. The estimation of F is then reduced to that of the parameter θ and the dependence structure is limited to that supported by the underlying statistical model. For many years, the multivariate normal distribution has been used for most multivariate analyses (see, for example, Johnson and Wichern, 2002). Recently, in many random phenomena, the distributions of which are skewed and have heavier tails than the normal, alternative distributions, such as multivariate skew-elliptical distributions, have been adopted (Branco and Dey, 2001; Genton and Loperfido, 2005). In recent approaches, copulas have become a popular tool in order to model multivariate dependence structures and for obtaining new multivariate distributions with given marginals.

In short, a copula is a multivariate distribution whose marginals are uniform allover $[0, 1]$. There are many parametric families of copulas, allowing the modeling of many different dependence structures (dos Anjos et al., 2004; Joe, 1997; Nelsen, 2006). Let F be a p-dimensional distribution function with the margins $F_1, \ldots, F_p$. Sklar (1959) first showed that there exists a p-dimensional copula C such that for all $\boldsymbol{x} = (x_1, \ldots, x_p)$ in the domain of F,

$$F(x_1, \ldots, x_p) = C(F_1(x_1), \ldots, F_p(x_p)).$$

If the variables $X_1, \ldots, X_p$ are absolutely continuous, then the copula C is unique; otherwise, C is uniquely determined on $Ran(F_1) \times \ldots \times Ran(F_p)$, where $Ran(F_i)$ is the image of the function F_i, $i = 1, \ldots, p$ (Sklar, 1959). Thus, the copula could be used to model the margins and the dependence structure separately. The non unique representation of a copula for discrete distributions is a theoretical issue that needs to be considered in the light of analytical proof, but this does not limit its empirical applications (Trivedi and Zimmer, 2007). However, the above theorem (Sklar, 1959) does not tell us how to find the copula C. This problem is widely discussed in the literature and there are several proposed solutions (see, for example, Durrleman et al., 2000b). The most widely used approach is to adjust several families of (parametric) copulas and choose one of them using some selection criteria or a test of goodness of fit (Rakonczai and Zempléni, 2007; Berg, 2009; Genest et al., 2009).

Nonparametric techniques may also be applied to estimate a multivariate distribution. A popular solution under this approach is the use of the empirical distribution function $F^{(n)} : \mathbb{R}^p \to [0, 1]$, defined as

$$F^{(n)}(t_1, \ldots, t_p) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1} \left\{ x_{1i} \leq t_1, \ldots, x_{pi} \leq t_p \right\}, \text{ for } (t_1, \ldots, t_p) \in \mathbb{R}^p,$$

where $\mathbb{1}\{A\}$ is the indicator of the set A. This is equivalent to using the relative frequencies to estimate the joint probability mass function. The relative frequencies coincide with the maximum likelihood estimate under the assumption that the data come from a multinomial distribution. A problem with these approaches is that the probability of the non-observed cells will be estimated as zero.

Another possibility is to use some function to smooth the empirical distribution. We can consider the Bernstein polynomials (DeVore and Lorentz, 1993; Lorentz, 1986) for this propose owing to their simplicity and good mathematical properties (Babu et al., 2002; Babu and Chaubey, 2006). Let $h : [0,1]^p \to \mathbb{R}$ be a continuous function. The m degree (multivariate) Bernstein polynomial for the function h, $B_h^m : [0,1]^p \to \mathbb{R}$, is defined as

$$B_h^m(x_1, \ldots, x_p) = \sum_{j_1=0}^{m} \ldots \sum_{j_p=0}^{m} h\left(\frac{j_1}{m}, \ldots, \frac{j_p}{m}\right) \prod_{i=1}^{p} \binom{m}{j_i} x_i^{j_i} (1 - x_i)^{m-j_i}.$$

The multivariate Bernstein polynomials for the function h converges uniformly to the function h as $m \to \infty$ (Heitzinger et al., 2003, 2004) and its derivatives are very simple to obtain. The function h must be defined in $[0,1]^p$ and, therefore, data that do not take values in $[0,1]^p$ must be transformed for practical purposes (Babu et al., 2002). Bernstein polynomials have been used to approximate a copula, say C, by simply replacing the function h with the copula C. The resultant Bernstein polynomial, B_C^m, which is also a copula that converges strongly to C, is called the Bernstein copula (Li et al., 1997, 1998; Kulpa, 1999; Sancetta and Satchell, 2001; Taylor, 2009). When the copula is unknown, the empirical copula can be used instead (Durrleman et al., 2000a,b; Sancetta, 2004; Sancetta and Satchell, 2004; Bouezmarni et al., 2010). The empirical copula is defined as

$$C_n(u_1, \ldots, u_p) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\left\{F_1(x_{1i}) \leq u_1, \ldots, F_p(x_{pi}) \leq u_p\right\}.$$

Notice that even when F_i, $i \in 1, \ldots, n$, is unknown, we can use the empirical marginal distribution $F_i^{(n)}$ as a consistent estimator of F_i, according to the Glivenko-Cantelli theorem (e.g., Vaart and Wellner, 1996).

Clearly our objective is to estimate a (discrete) probability mass function of a random vector. In fact we have obtained a continuous function as the first estimate and hence it should be discretized in order to have a proper estimate. Suppose, with no loss of generality, that all variables assume integer values in the set $\Omega = \{0, 1, \ldots, k\}$ with probability 1. Let F be the (estimated) joint distribution function and $B = [\boldsymbol{a}, \boldsymbol{b}] = [a_1, b_1] \times \ldots \times [a_p, b_p]$ be a p-dimensional rectangle with all its vertices in Ω. The F-volume of B (Nelsen, 2006) is given by

$$V_F(B) = \sum_{\boldsymbol{c}} sgn(\boldsymbol{c}) F(\boldsymbol{c}), \tag{1}$$

where the sum is taken over all the vertices $\boldsymbol{c} = (c_1, \ldots, c_p)$ of B, and $sgn(\boldsymbol{c})$ is given by

$$sgn(\boldsymbol{c}) = \begin{cases} 1, & \text{if } c_j = a_j \text{ for an even number of } j\text{'s}, \\ -1, & \text{if } c_j = a_j \text{ for an odd number of } j\text{'s}. \end{cases}$$

If we take $B = [\boldsymbol{b} - \boldsymbol{1}, \boldsymbol{b}] = [b_1 - 1, b_1] \times [b_2 - 1, b_2] \times \ldots \times [b_p - 1, b_p]$, with $b_i \in \Omega$, $\forall i = 1, \ldots, p$, the probability of the event $\{\boldsymbol{X} = \boldsymbol{b}\} = \{X_1 = b_1, \ldots, X_p = b_p\}$ can be calculated as

$$P(\boldsymbol{X} = \boldsymbol{b}) = V_F(B).$$

3 Proposed solution

Our proposal to estimate the joint distribution of a discrete random vector consists of the use of Bernstein polynomials to estimate both the marginals and the copula. The advantage of this method is the possibility of using all observations, even in the case of missing values in some variable. Furthermore, it is a nonparametric approach and there are no restrictions on the dependence structure.

First, for each random variable X_i, we estimate the marginal distributions using the empirical marginal distribution with n_i observations, $F_i^{(n_i)}(x) = \frac{1}{n_i}\sum_{j=1}^{n_i} \mathbb{1}(x_{ij} \leq x)$, $i = 1, \ldots, p$, and then using the m_i degree of the Bernstein polynomial to smooth this function

$$B_i^{m_i}(x) = \sum_{j=1}^{m_i} F_i^{(n_i)}\left(\frac{j}{m_i}\right)\binom{m_i}{j} x^j (1-x)^{(m_i-j)}.$$

Next, we estimate the copula using an alternative version of the empirical copula based on the n complete observations and the estimates $B_i^{m_i}$, $i = 1, \ldots, p$, $C_n(u_1, \ldots, u_p) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}\left\{B_1^{m_1}(x_{1j}) \leq u_1, \ldots, B_p^{m_p}(x_{pj}) \leq u_p\right\}$, and smooth this function to obtain the corresponding Bernstein copula,

$$B_{C_n}^m(u_1, \ldots, u_p) = \sum_{j_1=0}^{m} \cdots \sum_{j_p=0}^{m} C_n\left(\frac{j_1}{m}, \ldots, \frac{j_p}{m}\right) \prod_{i=1}^{p} \binom{m}{j_i} x_i^{j_i}(1-x_i)^{m-j_i}.$$

The estimate of the joint distribution function will be a discretization (equation 1) of the following function

$$\hat{F}_{m,n}(x_1, \ldots, x_p) = B_{C_n}^m\left(B_1^m(x_1), \ldots, B_p^m(x_p)\right). \tag{2}$$

The algorithm used to obtain the proposed solution is quite simple and is described below:

1. for all n_i observations of each variable X_i, estimate the marginal empirical distribution function $F_i^{(n_i)}$;
2. smooth each function $F_i^{(n_i)}$ with a Bernstein polinomial $B_i^{(m_i)}$ with degree m_i;
3. for all complete observations of the random vector $\boldsymbol{X}$, estimate de empirical copula C_n;
4. estimate the Bernstein copula by smoothing the empirical copula C_n using the m degree multivariate Bernstein polinomial $B_{C_n}^m$;
5. obtain a continuous estimate of the multivariate distribution function $\hat{F}_{m,n}$ given by equation 2;
6. the estimate of the discrete multivariate probability mass function is achieved by a discretization of $\hat{F}_{m,n}$ using equation 1.

4 Applications

In this section, we present some applications of the proposed method and compare its performance with a few existing solutions presented in Section 2. To illustrate the robustness of the method, we simulate a data set from two bivariate discrete distributions generated by copulas (examples 4.1 and 4.2) and apply the method to the observed data in the OCD example.

For each simulated distribution, we present the estimated probabilities in three cases:

1. 600 pairs of observations with no censured data

2. censured data only in one marginal, with 1000 observations in one marginal and 200 in another and

3. censured data in both variables, with 600 observations for each variable, 300 of which are complete pairs.

For each case, we present the estimates using the following estimation methods:

a. the empirical distribution using only the complete pairs

b. the multivariate skew-t approximation using only the complete pairs

c. the discretization of the normal copula with normal marginal approximation to the distribution function using all observations for marginal distribution estimation and the complete pairs for copula estimation

d. the discretization of the Bernstein polynomial approximation to the distribution function using only the complete pairs and

e. our proposed solution, using the Bernstein polynomial to approximate the margins using all observations and the copula using the complete pairs.

In all examples, we illustrate graphically the estimates for the probability mass distributions and evaluate some distances between the estimated and the theoretical distribution. Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ be the theoretical probabilities and $\hat{\boldsymbol{\theta}} = (\theta_1, \ldots, \theta_k)$ the estimated probabilities. We use the following distances to compare the estimates:

i. Aitchison's distance:

$$\Delta(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sqrt{\sum_{i=1}^{k} \left[\ln\left(\frac{\hat{\theta}_i}{\theta_i}\right) - \bar{L} \right]^2}, \text{ where } \bar{L} = \frac{1}{k} \sum_{i=1}^{k} \ln\left(\frac{\hat{\theta}_i}{\theta_i}\right)$$

ii. Euclidean distance:

$$\delta(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \sqrt{\sum_{i=1}^{k} \left[\hat{\theta}_i - \theta_i \right]^2}$$

iii. Total variation distance:

$$\tau(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^{k} \left| \hat{\theta}_i - \theta_i \right|$$

iv. Kullback-Leibler symmetrized divergence:

$$\mathcal{D}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \frac{1}{2} \left[\sum_{i=1}^{k} \theta_i \ln \left(\frac{\theta_i}{\hat{\theta}_i} \right) + \sum_{i=1}^{k} \hat{\theta}_i \ln \left(\frac{\hat{\theta}_i}{\theta_i} \right) \right]$$

Aitchison (2003, 2008) and Pawlowsky et al. (2007) give many reasons to use the Aitchison's distance for compositional vectors; that is, when the sum of the vector's components is constant (in our case, the sum of probabilities is equals to one). Moreover, the orderings induced by these distances agree in most cases. For these reasons, we will restrict ourselves to the Aitchison's distance to compare estimates.

At the end of this section, we present the estimates for the distribution of the real data described in the introduction. In this case, we do not know the theoretical distribution; we present only the estimates and calculate distances from the empirical distribution.

4.1 Simulated symmetrical distribution

In this section, we simulate data from a symmetrical distribution with marginals $X_1 \sim$ beta-binomial$(N_x = 20, \alpha = 5, \beta = 5)$ and $Y_1 \sim$ binomial$(N_y = 20, \pi = 0.5)$ and normal copula with parameter $\rho = 0.7$.

Example 4.1.1. 600 complete pairs of observations

Example 4.1.1	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	4.98521	0.02116	0.09988	0.04154
Skew T	1.44499	0.00629	0.02915	0.00345
Normal Copula	**1.28402**	**0.00476**	**0.02418**	**0.00236**
Bernstein Polynomial	3.45943	0.01388	0.07159	0.01870
Bernstein Copula	3.23712	0.01217	0.06360	0.01578

Table 1: Distances between the estimates and theoretical probabilities for the example 4.1.1

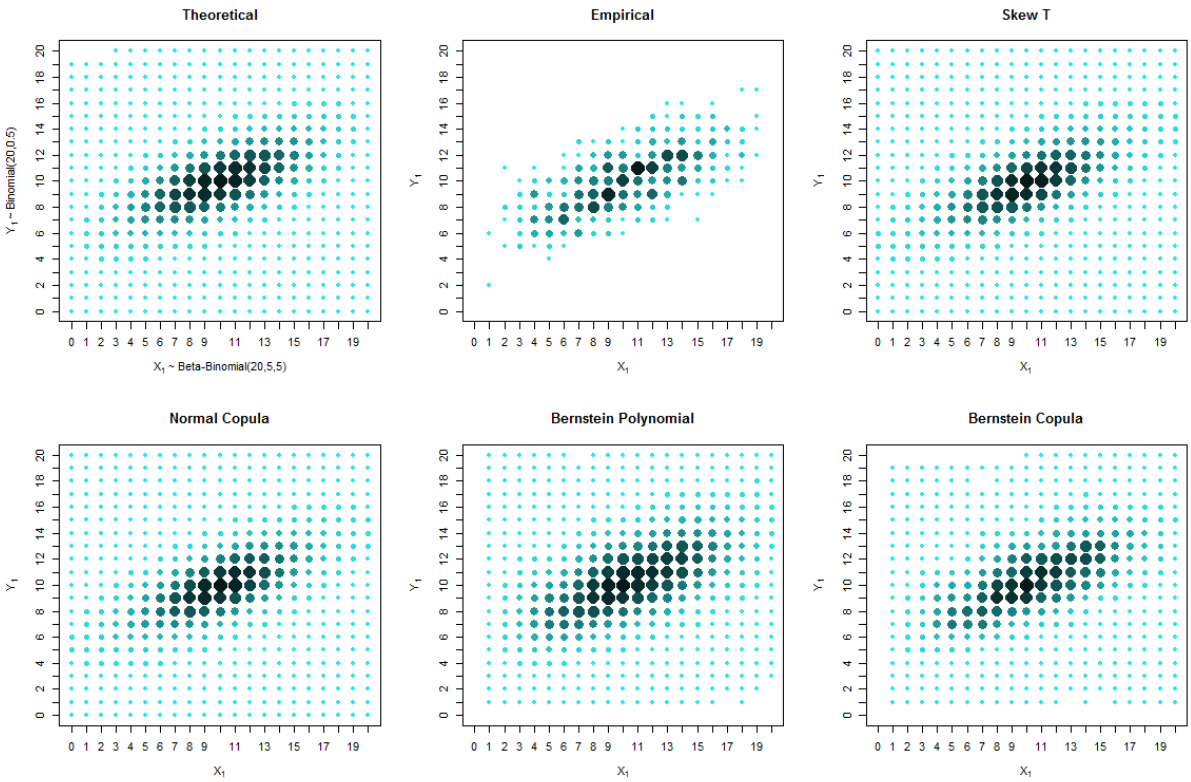

Figure 2: Estimates and theoretical probabilities for example 4.1.1.

Example 4.1.2. Censured data only in one marginal, with 1000 observations in one marginal and 200 in another

Example 4.1.2	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	8.97909	0.03454	0.17083	0.12490
Skew T	1.28441	**0.00493**	**0.02291**	**0.00239**
Normal Copula	**1.28040**	0.00554	0.02738	0.00284
Bernstein Polynomial	4.84901	0.02049	0.11110	0.03982
Bernstein Copula	3.28689	0.01171	0.06340	0.01530

Table 2: Distances between the estimates and theoretical probabilities for the example 4.1.2

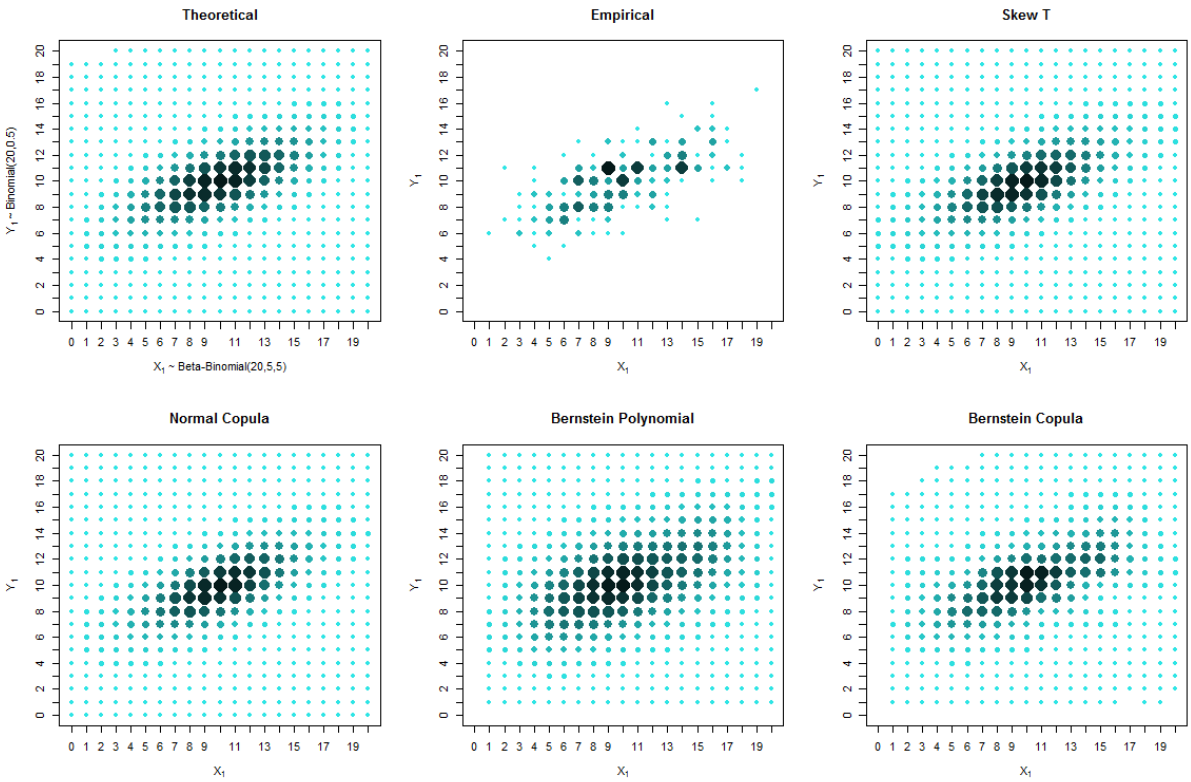

Figure 3: Estimates and theoretical probabilities for example 4.1.2

Example 4.1.3. Censured data and in both variables, with 600 observations for each variable, of which 300 are complete pairs

Example 4.1.3	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	7.32955	0.03035	0.13826	0.09162
Skew T	1.12383	0.00419	0.02221	0.00185
Normal Copula	**1.06365**	**0.00375**	**0.01891**	**0.00146**
Bernstein Polynomial	4.54073	0.01934	0.10051	0.03531
Bernstein Copula	3.54526	0.01377	0.06743	0.01963

Table 3: Distances between the estimates and theoretical probabilities for the example 4.1.3

In these examples, we can see, from figures 2, 3 and 4 and tables 1, 2 and 3, that the solutions based on elliptical distributions, skew t and normal, yield better estimates. It occurs because the theoretical probability mass function has an elliptical shape. However, in practical situations, we do not have knowledge about the real shape of the distribution. In this case, the empirical distribution could be a good base to study the estimates, despite the existence of many unobserved points that are estimated as zero. Comparing the estimates with the empirical distribution, our proposed solution seems to produce good results, specially in the presence of censured data.

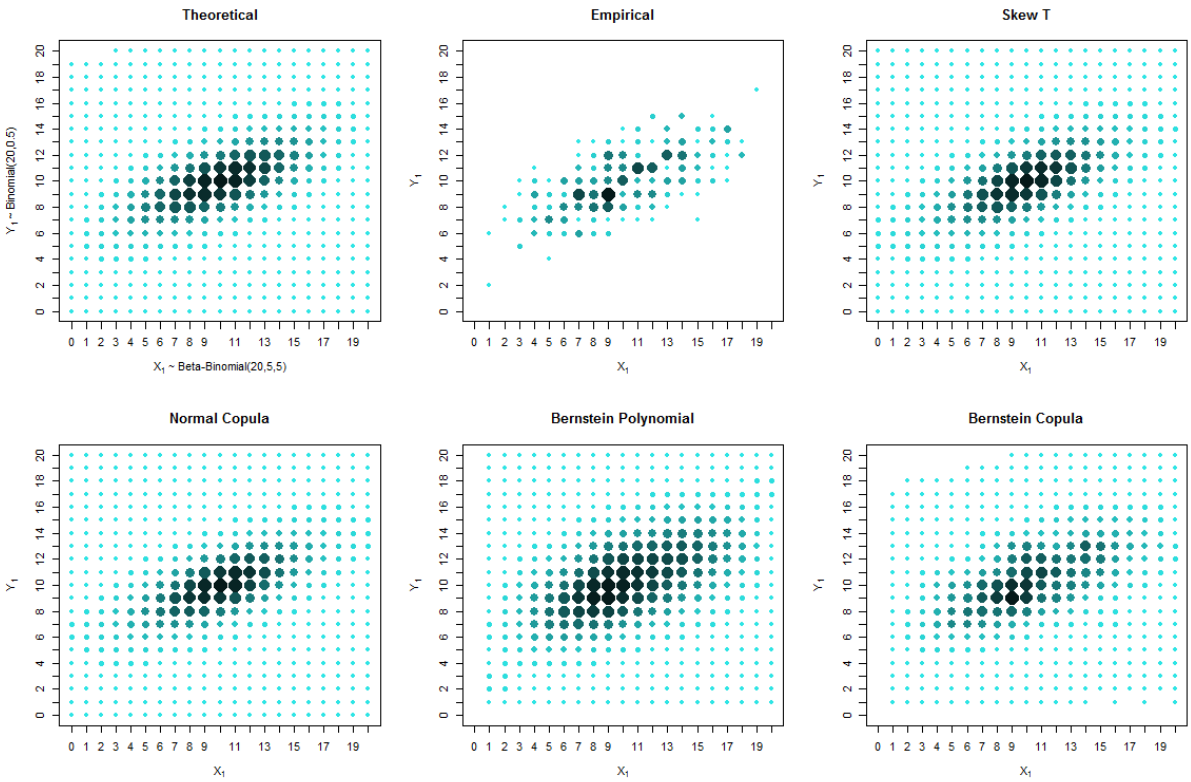

Figure 4: Estimates and theoretical probabilities for example 4.1.3

4.2 Simulated asymmetrical distribution

In this section, we present the simulated data for an asymmetrical distribution, with margins $X_2 \sim$ beta-binomial$(N_x = 20, \alpha = 0.85, \beta = 1.1)$ and $Y_2 \sim$ binomial$(N_y = 15, \pi = 0.6)$ and Gumbel copula with the parameter $\theta = 0.7$.

Example 4.2.1. 600 complete pairs of observations

Example 4.2.1	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	6.17032	0.02767	0.12761	0.06377
Skew T	5.47625	0.03287	0.14429	0.07724
Normal Copula	5.76598	0.03293	0.14785	0.08020
Bernstein Polynomial	5.41325	**0.02436**	0.11969	0.05380
Bernstein Copula	**5.07634**	0.02519	**0.11842**	**0.05068**

Table 4: Distances between the estimates and theoretical probabilities for the example 4.2.1

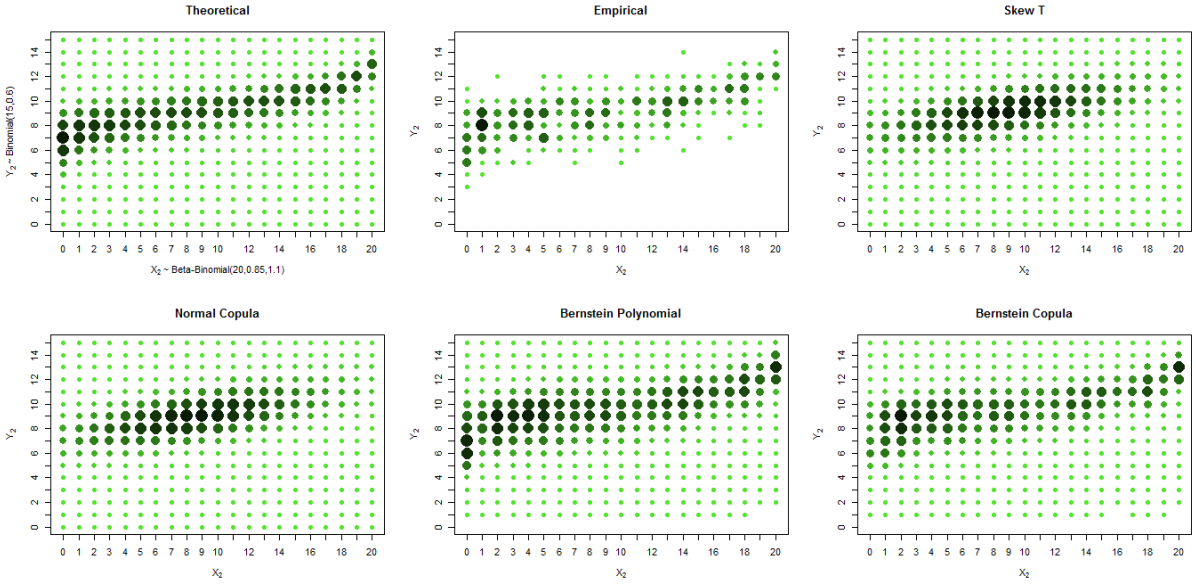

Figure 5: Estimates and theoretical probabilities for example 4.2.1

Example 4.2.2. Censured data only in one marginal, with 1000 observations in one marginal and 200 in another

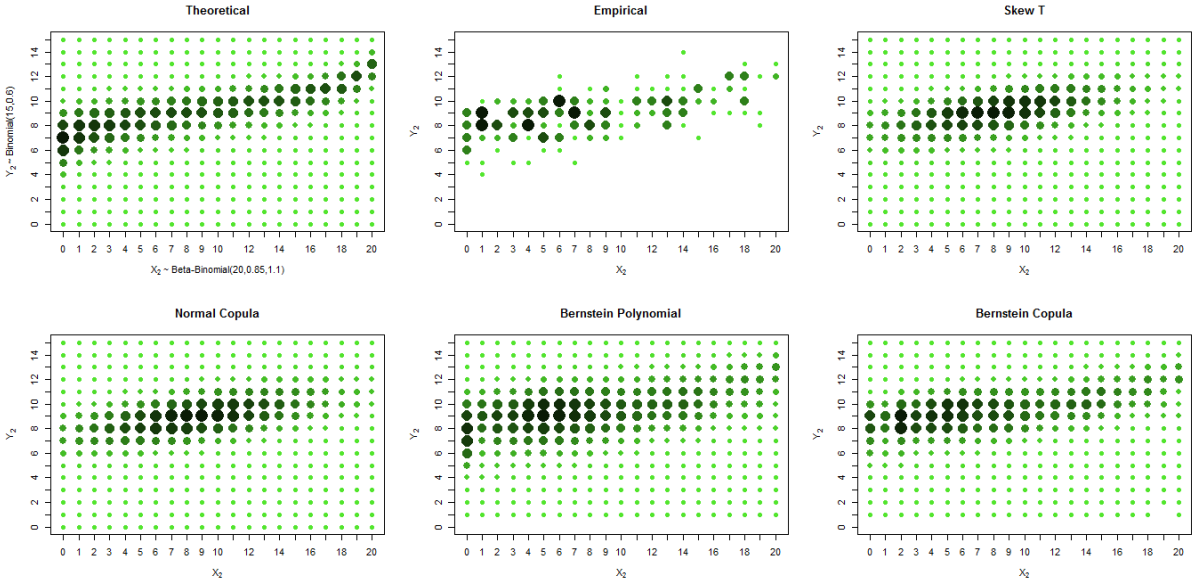

Figure 6: Estimates and theoretical probabilities for example 4.2.2

Example 4.2.2	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	8.77534	0.03906	0.19104	0.14363
Skew T	5.23773	0.03130	0.13494	0.07356
Normal Copula	5.07437	0.02892	0.12727	0.06549
Bernstein Polynomial	5.65580	0.02626	0.13135	0.06562
Bernstein Copula	**4.86027**	**0.02558**	**0.11172**	**0.05379**

Table 5: Distances between the estimates and theoretical probabilities for the example 4.2.2

Example 4.2.3. Censured data and in both variables, with 600 observations for each variable, of which 300 are complete pairs

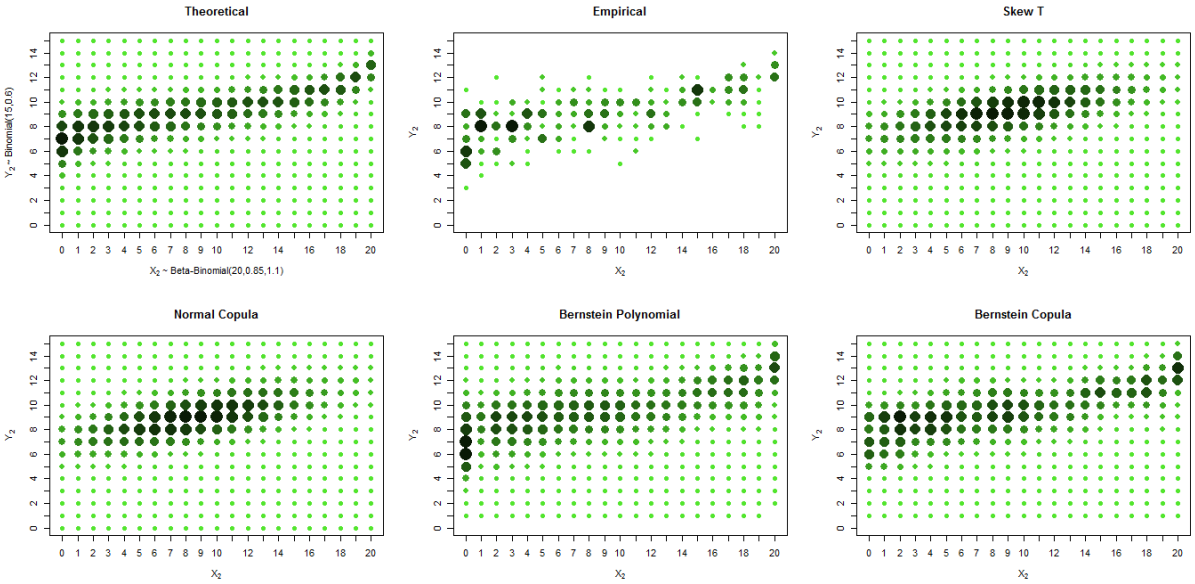

Figure 7: Estimates and theoretical probabilities for example 4.2.3

Example 4.2.3	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Empirical	7.32005	0.03284	0.15576	0.09786
Skew T	4.79233	0.02760	0.12321	0.05917
Normal Copula	5.06253	0.02819	0.12855	0.06325
Bernstein Polynomial	5.09522	0.02253	0.11486	0.05028
Bernstein Copula	**4.35547**	**0.01957**	**0.09863**	**0.03595**

Table 6: Distances between the estimates and theoretical probabilities for the example 4.2.3

In the case of an asymmetrical distribution, our proposed solution presents a better estimation in the three described cases: with no censured data, with missing data in one variable and with missing data in both variables. This can be seen graphically (figures 5, 6 and 7) and through distances (tables 4, 5 and 6). It is possible to note graphically

that the probabilities of the smaller and the greater values of X_2 are well estimated by our method.

4.3 4.3 Real Data

In this section, we present the estimates for the real data mentioned in the introduction.

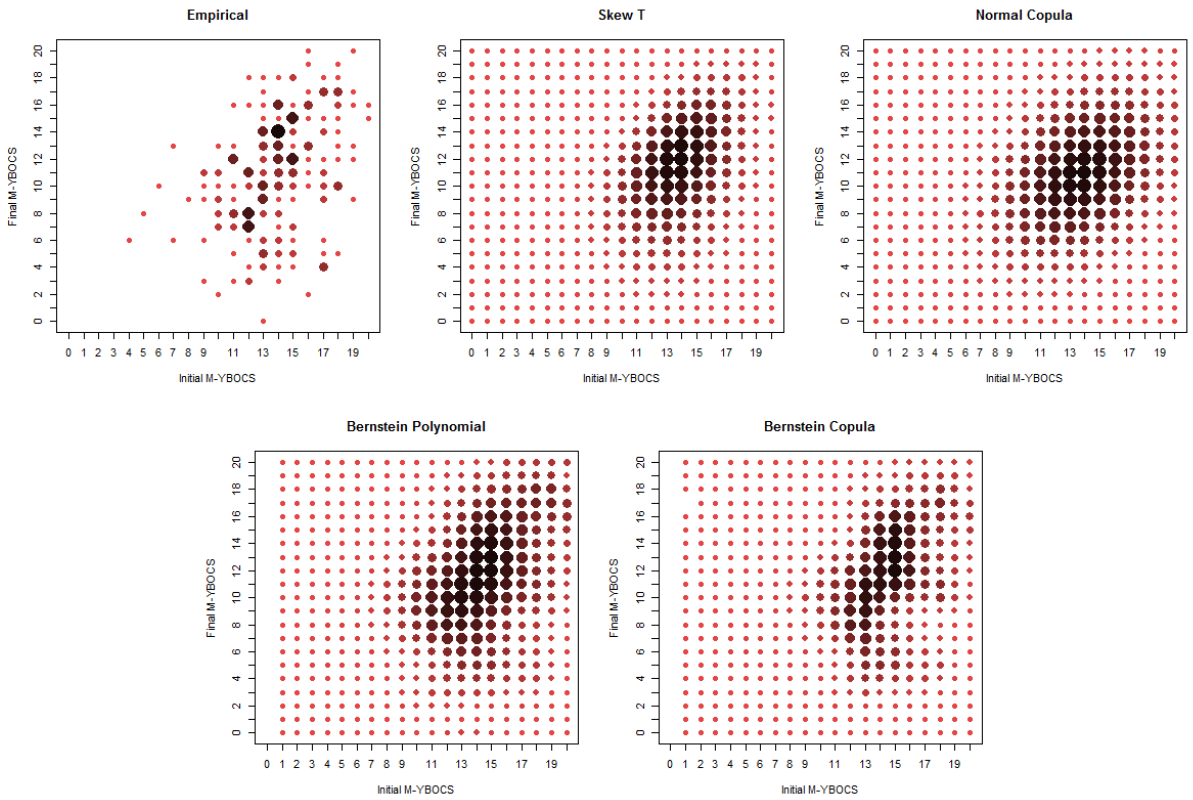

Figure 8: Estimates of probabilities for example 4.3

Example 4.3	Aitchison	Euclidean	Total Variation	Kullback-Leibler
Skew T	11.56219	0.03941	0.22684	0.19899
Normal Copula	12.07092	0.04143	0.24701	0.21760
Bernstein	11.35361	0.03933	0.22910	0.19125
Bernstein Copula	**10.81475**	**0.03703**	**0.21020**	**0.17184**

Table 7: Distances between the estimates and empirical probabilities for example 4.3

The theoretical distribution is not known in applications. In this case, we compare the estimates with the empirical distribution. The proposed solution presents smaller distances. Moreover, it can be seen graphically that our method captures the asymmetrical nature of the data.

5 Conclusions

In this work, a new approach to the problem of estimating discrete bivariate distributions is presented. The procedure, which basically consists of the estimation of both the marginals and the copula by means of Bernstein copulas, aims to cover three important points: the handling of discrete bivariate data in the presence of marginal missing values (using all available information, including observations of only one of the variables); the possibility of obtaining positive estimates for non-observed cells, yielding "smoother" estimated discrete distributions; the consideration of more general dependence structures between the relevant random variables owing to its less restrictive nonparametric nature. The generalization to k-dimensional random vectors, k $>$ 2, is straightforward and easy for computational implementation.

The new method was applied to simulated data in a few examples and on the basis of usual measures of distance (between the estimated and the theoretical distributions) it performed better than some existing solutions in cases of asymmetrical models, mainly in situations of censured data. The innovation was also applied to data sampled from adults with primary obsessive-compulsive disorder diagnostic. A few questions concerning the new procedure were not addressed here though: i) the study of asymptotic properties with the increasing of sample size n and/or the increasing of the polynomials degree m; ii) under a Bayesian decision-theoretic approach, the existence of a formal justification for the new procedure; iii) a more rational way to choose m, that could depend on n, under this approach; iv) the incorporation of prior knowledge, maybe as in Petrone (1999a,b) and Petrone and Wasserman (2002), although they worked in univariate Bayesian perspective. These queries are the goals of forthcoming articles.

References

J. Aitchison. *The Statistical Analysis of Compositional Data*. Blackburn Press, Caldwell, NJ, USA, 2003. ISBN 9781930665781.

J. Aitchison. The single principle of compositional data analysis, continuing fallacies, confusions and misunderstandings and some suggested remedies. Technical report, Universitat de Girona. Departament d'Informàtica i Matemàtica Aplicada, 2008.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. Washington DC, 1994.

G. J. Babu and Y. P. Chaubey. Smooth estimation of a distribution and density function on a hypercube using Bernstein polynomials for dependent random vectors. *Statistics & Probability Letters*, 76(9):959–969, 2006.

G. J. Babu, A. J. Canty, and Y. P. Chaubey. Application of Bernstein polynomials for smooth estimation of a distribution and density function. *Journal of Statistical Planning and Inference*, 105(2):377–392, 2002.

D. Berg. Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance*, 15(7-8):675–701, 2009.

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

M. D. Branco and D. K. Dey. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79(1):99–113, 2001. ISSN 0047-259X. doi: 10.1006/jmva.2000.1960. URL `http://www.sciencedirect.com/science/article/pii/S0047259X00919602`.

R. A. DeVore and G. G. Lorentz. *Constructive Approximation*. Springer, New York, 1993. ISBN 3-540-50627-6.

J. B. Diniz, V. Fossaluza, C. Belotto-Silva, R. G. Shavitt, and C. A. B. Pereira. The use of Yale-Brown Obsessive-Compulsive Scale: new views of an old measure. *European Neuropsychopharmacology*, 21(Supplement 3):S531 – S532, 2011. ISSN 0924-977X. doi: 10.1016/S0924-977X(11)70865-3. URL `http://www.sciencedirect.com/science/article/pii/S0924977X11708653`.

U. U. dos Anjos, F. H. Ferreira, N. V. Kolev, and B. M. V. Mendes. *Modeling Dependences via Copulas (in Portuguese)*. 16º SINAPE. Associação Brasileira de Estatística, São Paulo, Brazil, 2004.

V. Durrleman, A. Nikeghbali, and T. Roncalli. Copulas approximation and new families. Technical report, Groupe de Recherche Opérationnelle, Crédit Lyonnais, 2000a.

V. Durrleman, A. Nikeghbali, and T. Roncalli. Which copula is the right one? Technical report, Groupe de Recherche Opérationnelle, Crédit Lyonnais, 2000b.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2):199–213, 2009.

M. G. Genton and N. M. R. Loperfido. Generalized skew-elliptical distributions and their quadratic forms. *Annals of the Institute of Statistical Mathematics*, 57: 389–401, 2005. ISSN 0020-3157. URL `http://dx.doi.org/10.1007/BF02507031`. 10.1007/BF02507031.

W. K. Goodman, L. H. Price, S. A. Rasmussen, C. Mazure, P. Delgado, G. R. Heninger, and D. S. Charney. The Yale-Brown Obsessive-Compulsive Scale: II. Validity. *Archives of general psychiatry*, 46(11):1012, 1989a.

W. K. Goodman, L. H. Price, S. A. Rasmussen, C. Mazure, R. L. Fleischmann, C. L. Hill, G. R. Heninger, and D. S. Charney. The Yale-Brown Obsessive-Compulsive Scale: I. Development, use, and reliability. *Archives of general psychiatry*, 46(11):1006, 1989b.

C. Heitzinger, A. Hössinger, and S. Selberherr. On smoothing three-dimensional Monte Carlo ion implantation simulation results. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(7):879–883, 2003.

C. Heitzinger, A. Hössinger, and S. Selberherr. An algorithm for smoothing three-dimensional Monte Carlo ion implantation simulation results. *Mathematics and Computers in Simulation*, 66(2):219–230, 2004.

H. Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, London, 1997. ISBN 0-412-07331-5.

R. A. Johnson and D. W. Wichern. *Applied multivariate statistical analysis*, volume 4. Prentice Hall, Upper Saddle River, NJ, 2002.

T. Kulpa. On approximations of copulas. *International Journal of Mathematics and Mathematical Sciences*, 22(2):259–269, 1999.

X. Li, P. Mikusiński, H. Sherwood, and M.D. Taylor. *Distributions with Given Marginals and Moment Problems*, chapter On approximation of copulas, pages 107–116. Kluwer Academic Publishers, 1997. ISBN 0-792-34573-8.

X. Li, P. Mikusiński, and M. D. Taylor. Strong approximations of copulas. *Journal of Mathematical Analysis and Applications*, 255:608–623, 1998. ISSN 0161-1712.

G. G. Lorentz. *Bernstein Polynomials*. Chelsea Pub Co, New York, second edition, 1986. ISBN 0-8284-0323-6.

R. B. Nelsen. *An Introduction to Copulas*. Springer Verlag, New York, second edition, 2006. ISBN 0-387-28659-4.

V G. Pawlowsky, J. J. Egozcue, and R. T. Delgado. Lecture notes on compositional data analysis. Technical report, Universitat de Girona, 2007.

C. A. B. Pereira, C. B. Silva, and J. B. Diniz. *Clínica Psiquiátrica*, chapter 147: Estatística em Psiquiatria. Editora Manole, 2011.

S. Petrone. Random Bernstein polynomials. *Scandinavian Journal of Statistics*, 26(3): 373–393, 1999a.

S. Petrone. Bayesian density estimation using Bernstein polynomials. *The Canadian Journal of Statistics*, 27(1):105–126, 1999b.

S. Petrone and L. Wasserman. Consistency of Bernstein polynomial posteriors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(1):79–100, 2002. ISSN 1467-9868.

P. Rakonczai and A. Zempléni. Copulas and goodness of fit tests. *Recent Advances in Stochastic Modeling and Data Analysis*, pages 198–206, 2007.

A. Sancetta. Nonparametric estimation of multivariate distributions with given marginals. *Cambridge Working Papers in Economics*, 2004.

A. Sancetta and S. E. Satchell. Bernstein approximations to the copula function and portfolio optimization. *Cambridge Working Papers in Economics*, 2001.

A. Sancetta and S. E. Satchell. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, 20(3):535–562, 2004.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statistique Univ. Paris*, 8:229–231, 1959.

M. D. Taylor. Bernstein polynomials and n-copulas. *Arxiv preprint arXiv:0903.1000*, 2009.

P. K. Trivedi and D. M. Zimmer. *Copula Modeling: an Introduction for Practitioners*, volume 1 of *Foundations and Trends in Econometrics*. Now Pub, Hanover, MA, USA, 2007. ISBN 978-1-60198-020-5.

A. W. Van Der Vaart and J. A. Wellner. *Weak convergence and empirical processes*. Springer Verlag, New York, USA, 1996. ISBN 0-387-94640-3.