ABSTRACT

LU, LU. Multivariate Dependence using Nonparametric Copula Models. (Under the direction of Sujit K. Ghosh.)

In many practical scenarios (e.g., finance, system reliability, etc.), it is often of interest to measure the dependence among variables by modeling a multivariate distribution using a copula and test for some desired association properties like positive quadrant dependent (PQD) or negative quadrant dependent (NQD). Many parametric copula families have been used that allow for controlling the PQD/NQD property by a finite dimensional parameter (often just real-valued) and the problem reduces to the straightforward estimation and testing for fixed dimensional parameter using standard statistical methodologies (e.g., maximum likelihood). Chapter 2 extends such a line of work by dropping any parametric assumptions and provides a fully data-dependent automated approach to estimate a bivariate copula and test for PQD property. The estimator is shown to be large-sample consistent under a set of mild regularity conditions. Numerical illustrations based on simulated data are also presented to compare the performance of the proposed testing procedure with some available methods and applications to real case studies are also provided.

In chapter 3, by utilizing the so-called empirical checkerboard copula we build a hierarchical empirical Bayes model that enables the estimation of a smooth copula function for arbitrary dimensions. The proposed estimator built on the multivariate Bernstein polynomials is itself a genuine copula and the selection of its dimension-varying degrees is data-dependent. We call the proposed estimator the empirical checkerboard Bernstein copula(ECBC) and show that it provides a more accurate estimate of several multivariate dependence measures which can be obtained in closed form. We examine the asymptotic and finite-sample performance of the proposed estimator and compare it with some nonparametric estimators through simulation studies. An application to portfolio risk management is presented along with a quantification of estimation uncertainty.

Conditional copulas are useful tools for modeling the dependence between variables that is allowed to change with the value of covariates. Conditional dependence measures such as conditional Kendall's tau and Spearman's rho that can be expressed as functionals of the conditional copula are often used to evaluate the strength of dependence conditioning on the covariates. In general, semiparametric estimation methods of conditional copulas rely on an assumed parametric copula family where the copula parameter is assumed to be a function of the covariates. The functional relationship can be estimated nonparametrically using different techniques but it is required to choose an appropriate copula model from various candidate families. In chapter 4, by employing the empirical checkerboard Bernstein copula (ECBC) estimator we propose a fully nonparametric approach for estimating conditional copulas, which doesn't require any selection of parametric copula models. Closed-form estimates of the conditional dependence measures are derived directly from the proposed ECBC-based conditional copula estimator. We show the large-sample consistency of the proposed estimator. The finite-sample performance of the proposed estimator and comparison with semiparametric methods are investigated through simulation studies. An application to real case studies is also provided.

Multivariate Dependence using Nonparametric Copula Models

by
Lu Lu

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Statistics

Raleigh, North Carolina

2021

APPROVED BY:

Ryan Martin	Rui Song
Tao Pang	Sujit K. Ghosh Chair of Advisory Committee

DEDICATION

To my family.

BIOGRAPHY

Lu Lu was born and grew up in Dalian, China. In June 2016, she got her bachelor's degree in Statistics from Nankai University, Tianjin, China. In August 2016, she came to the U.S. and joined North Carolina State University to pursue a doctoral degree in Statistics. She will earn her Ph.D. in Statistics in 2021.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincerest gratitude to my advisor, Dr. Sujit K. Ghosh. His broad knowledge, wealth of ideas, and great enthusiasm have been deeply inspiring me in my research work. Throughout my PhD studies, He is always there providing me with great patience, persuasive arguments, and words of encouragement. The completion of this dissertation would not have been possible without his continuous support and guidance.

I would also like to thank Dr. Ryan Martin, Dr. Rui Song and Dr. Tao Pang for serving as my committee members. I really appreciate their valuable and insightful comments that help strengthen the work in this dissertation.

I am grateful to DGP Dr. Wenbin Lu, Lanakila Alexander, and all the other faculties and staff in the Department of Statistics at NC state who have greatly assisted me through my graduate school education here. Thanks also goes to all my peers for their helpful discussions.

Last but not least, I want to thank my parents and sister for their constant love and faith in me. They have always been extremely supportive in every decision I have ever made in my life. I also very much appreciate all my friends who have supported me to overcome different challenges in the past five years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

Understanding the natural dependence among a collection of two or more variables is one of the fundamental problems in statistics and has found many applications in the field of insurance, finance, system reliability, etc. among other areas. Pearson correlation coefficient (also referred to as Pearson's r) has been widely used to measure the linear relationship between two random variables. However, modern risk analysis tends to assess risks caused by multiple events and often calls for an understanding of stochastic dependence beyond simple linear correlation. For instance, in the management of large portfolios of stock assets, the main risk is the simultaneous drop of prices. Thus a better understanding of the dependence among financial assets is crucial for evaluating the risk of losses.

The interest in revealing the underlying dependence often involves two main questions: (i) what is the form of the dependence; and (ii) how strong is the dependence among vari-

ables. These create challenges in testing for specific dependence structures and estimating dependence measures, especially for large dimensions. Even in bivariate cases, there are a variety of positive or negative dependence properties, e.g. positive quadrant dependent (PQD), negative quadrant dependent (NQD), stochastically increasing (SI), stochastically decreasing (SD), likelihood ratio dependent (LRD), etc. Multivariate dependence concepts are often natural extensions of the corresponding bivariate counterparts that have straightforward interpretations, although the dependence structure could become much more complicated in higher dimensions (see Joe [1997]). On the other hand, many association measures, such as Spearman's rho, Kendall's tau, Gini's gamma, and Blomqvist's beta, are well established for measuring the strength of the bivariate dependence along with their multivariate generalizations; see Nelsen [1996], Schmid & Schmidt [2007], etc.

Copula models are useful tools in the analysis of multivariate data since by using the well-known Sklar's theorem, any multivariate joint distribution can be decomposed into its univariate marginal distributions and a copula function, which allows for capturing the arbitrary dependence structure between several random variables. Given a random vector $(X_1, ..., X_d)$ with joint cumulative distribution function (CDF) F and marginal CDFs F_j, $j = 1, ..., d$, by Sklar's theorem (Sklar [1959]), the CDF F can be expressed as

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)), \ \forall (x_1, ..., x_d) \in \mathbb{R}^d, \tag{1.1}$$

where $C(\cdot)$ denotes the copula function. If the marginal CDFs F_j, $j = 1, ..., d$ are continuous, then the copula C is unique and it is itself the joint CDF of a random vector $(U_1 = F_1(X_1), ..., U_d = F_d(X_d))$ having marginals as uniform distributions on $[0, 1]$.

From the perspective of joint distribution estimation, Sklar's Theorem allows us to separate the modeling of the dependence structure from the marginal distributions. Unlike Pearson's r that depends on the marginal distributions of random variables, most of the other dependence measures such as Spearman's rho and Kendall's tau can be expressed as

a function of copulas (see Nelsen [2007]). Moreover, many of the dependence properties like PQD and SI can also be stated in terms of copulas. Thus copula plays an important role in modeling the general dependence structure between multiple variables and it is critical to estimate copula in an accurate way.

Many parametric families have been proposed for modeling copulas and there has been previous work addressing the corresponding parametric estimation methods. Some commonly used copula families are Archimedean copulas, elliptical copulas, vine copulas, etc. (see Nelsen [2007] and Joe [2014]). In regard to frequentist methods based on an assumed parametric class, popular estimation methods for parametric copulas are based on the maximum likelihood principle either using the full likelihood or a two-stage maximum likelihood estimation (see Joe [2005]). In contrast, several Bayesian methods have also been utilized for estimating parametric copula based on the full likelihood and priors assigned to the set of parameters for the marginal distributions and the copula function; see Smith [2011], etc. However, no matter how sophisticated and flexible parametric models that we may use, it might still lead to biased copula estimations when the parametric model is misspecified and thus may not be able to capture complex dependence structures required in practice.

Recognizing some of the limitations of parametric copula models, a variety of nonparametric estimators have been proposed for copula estimation. Most of the available nonparametric estimators rely on the empirical methods (Deheuvels [1979]; Fermanian et al. [2004]; Genest et al. [2017]) or kernel-based methods (Gijbels & Mielniczuk [1990]; Chen & Huang [2007]); Omelka et al. [2009]). However, most of these estimators are valid copulas only asymptotically, meaning that they are not necessarily genuine copulas for finite samples. As a result, dependent measures (e.g. Spearman's rho, Kendall's tau, etc.) based on such estimated copula could take values outside of their natural range, which makes them problematic in practice. Besides, it is often difficult to obtain such nonparametric estimators under the constraint of a certain property, making them less attractive in

testing for a specific dependence structure.

In order to overcome the above-mentioned technical challenges, we introduce new nonparametric estimation methods for copulas. In Chapter 2, we focus on estimating bivariate copula and testing for the positive quadrant dependent (PQD) property, which is a particular form of stochastic dependence between two random variables. Based on the Bernstein copula (Sancetta & Satchell [2004]), we propose Anderson-Darling discrepancy-based criteria for copula estimation with or without the constraint of PQD property and develop a grid search algorithm that adaptively takes data into consideration to select the degrees of the Bernstein copula guided by the asymptotic theory. Our nonparametric copula estimator is shown to be a proper copula for any finite sample sizes and large-sample consistent under mild regularity conditions. While the unconstrained copula estimator is appropriate for the general estimation of bivariate copula, the PQD-constrained copula estimator is useful either when it is known that the underlying copula satisfies PQD or to approximate the p-values under the null hypothesis of PQD by using bootstrap-based methodology when testing for the PQD property of an unknown copula.

Chapter 3 extends the estimation of copulas to higher dimensions. In this chapter, we introduce a new nonparametric estimator for multivariate copula, which is constructed by extending the Bernstein copula using the so-called empirical checkerboard copula. For the selection of its dimension-varying degrees, we build a hierarchical empirical Bayes model that allows for automatic data-dependent estimation of such tuning parameters for arbitrary dimensions, which is shown to be practically useful. We call the proposed estimator the empirical checkerboard Bernstein copula (ECBC) and it is itself a genuine smooth copula. The large-sample consistency of ECBC is established under mild regularity conditions. Owing to the closed form of estimated copula function and its density, it can be shown that ECBC allows for straightforward estimation of various dependence measures always taking values within the proper parameter range.

In some circumstances, the dependence structure among variables can be influenced

by covariates and it is then of interest to explore the impact of the covariates on the dependence structure. Conditional copulas are useful tools for modeling the dependence among variables conditionally upon covariates and conditional dependence measures such as conditional Kendall's tau and Spearman's rho can be utilized to evaluate how the strength of dependence changes with the value of covariates. In Chapter 4, we provide a fully nonparametric approach for estimating conditional copulas based on the empirical checkerboard Bernstein copula (ECBC) estimator given in Chapter 3. Compared to semiparametric methods, it is not required to make any selection of copula families, making it easy to implement in practice. Like unconditional cases, most of the conditional dependence measures can be expressed as functionals of the conditional copula. The proposed ECBC-based conditional copula estimator immediately results in closed-form nonparametric estimates of the conditional dependence measures. We also present the large-sample consistency of the proposed conditional copula estimator in this chapter.

CHAPTER

2

NONPARAMETRIC ESTIMATION AND TESTING FOR POSITIVE QUADRANT DEPENDENT BIVARIATE COPULA

2.1 Introduction

In the field of finance and system reliability, it is often of interest to estimate the joint probability distribution of a pair of random variables (X, Y) and test for a desired dependency property. For example, if X and Y represent the insurance claims for losses to building and its contents, respectively, then it is likely higher values of Y lead to higher values of X than that without the knowledge of Y. Such a scenario can be expressed by the following prob-

abilistic expression: $\Pr[X > x | Y > y] \geq \Pr[X > x]$ for any x and y for which $\Pr[Y > y] > 0$. Notice that such a property is equivalent to $\Pr[X > x, Y > y] \geq \Pr[X > x]\Pr[Y > y]$ and leads the concept of positive quadrant dependent (PQD). Let the cumulative distribution function (CDF) (X, Y) be denoted by $H(x, y) = \Pr[X \leq x, Y \leq y]$ and $F(x) = \Pr[X \leq x] = H(x, \infty)$ and $G(y) = \Pr[Y \leq y] = H(\infty, y)$ denote the marginal CDF of X and Y, respectively. For the rest of the article, we assume that both marginal CDFs are absolutely continuous.

In a celebrated paper Sklar [1959] has that any bivariate CDF $H(x, y)$ with continuous marginal CDFs, $F(x)$ and $G(y)$ can be expressed uniquely as $H(x, y) = C(F(x), G(y))$, where $C(u, v)$ defined on the square $I^2 \equiv [0, 1]^2$ is known as the copula function satisfying the property that $C(0, v) = C(u, 0) = 0$ and $C(u, 1) = u, C(1, v) = v$ for any $(u, v) \in [0, 1]^2$, meaning that $C(u, v)$ is the joint CDF of the pair of random variables $(U = F(X), V = G(Y))$ having $Unif[0, 1]$ marginal distributions. It is to be noted that more general results of Sklar's theorem beyond bivariate case and also applicable to discrete valued random variables are available in Durante & Sempi [2015]. However, as stated above, the focus of this chapter is to explore dependency property of continuous valued bivariate random variables (X, Y). Thus, we assume the existence of a unique copula C corresponding to the joint CDF H satisfying the following relations

$$H(x, y) = C(F(x), G(y)) \ \forall \ x, y \in \mathbb{R}, \tag{2.1}$$

and

$$(F(X), G(Y)) \sim C. \tag{2.2}$$

The joint CDF H or equivalently the corresponding copula function C is said to satisfy the PQD property if the following condition is satisfied:

$$C(u, v) \geq uv \ \forall \ u, v \in [0, 1]. \tag{2.3}$$

Notice that in terms of the joint CDF H, the PQD property (2.3) is equivalent to $H(x,y) \geq F(x)G(y)$ for any $x, y \in \mathbb{R}$. From an intuitive point of view, two random variables are PQD if the probability that they are simultaneously small (or large) is at least as great as it would be were they independent, i.e.,

$$\begin{aligned} \mathbb{P}(U \leq u, V \leq v) \geq uv &\Leftrightarrow \mathbb{P}(U \leq u | V \leq v) \geq \mathbb{P}(U \leq u) \\ &\Leftrightarrow \mathbb{P}(X \leq x | Y \leq y) \geq \mathbb{P}(X \leq x) \end{aligned} \tag{2.4}$$

PQD is a useful concept since it has found many applications in the field of insurance, finance and reliability. For extensive discussion, see Denuit & Scaillet [2004] and Nelsen [2007].

Two of the primary goals of this chapter are (i) to estimate the copula function C from a random sample of n independent identically distributed (iid) observations $(X_i, Y_i) \overset{iid}{\sim} H(x,y)$ for $i = 1, 2, \ldots, n$ with or without knowing that C or equivalently H satisfies the PQD property; and (ii) to test if the samples were generated from a copula that satisfies the PQD property.

There have been several previous work addressing the estimation problem (i) as described above. Many parametric copula families have been used to model copula, in particular, copula satisfying PQD property. For example, all members of the Gumbel copula families have the property of PQD. Technically, any parametric copula family that includes at least a nonempty subset of PQD can be used to model a PQD copula where the parameter is estimated by the maximum likelihood method. However, it might lead to biased estimation if the parametric family is not flexible to capture all sets of dependence. Amblard & Girard [2002] consider a semiparametric family of copula, which is an extension of the Farlie-Gumbel-Morgenstern copula family, and the estimation procedure for PQD copula in such semiparametric family is developed by estimating the univariate generating function nonparametrically in Amblard & Girard [2005]. However, it is shown that the semiparametric approach has its limitations in modeling lager dependence and asymmetric

cases. In contrast, many nonparametric estimators of copula have been proposed such as kernel-based copula estimators (Scaillet & Fermanian [2002]), empirical Bernstein copula (Sancetta & Satchell [2004]), integrated version of the density mirror reflection estimator (Gijbels & Mielniczuk [1990]), local linear estimator (Chen & Huang [2007]), improvement of these two estimators (Omelka et al. [2009]), etc. See also Rémillard & Scaillet [2009] for other nonparametric estimators. However, most of these estimators are valid copulas only asymptotically and hard to be restricted to be PQD. Besides, Gijbels & Sznajder [2013] propose a PQD-constrained nonparametric copula estimator but there is a lack of theoretical support for its PQD property. In this chapter, we provide a nonparametric estimator for copula, which is guaranteed to be a proper copula for any finite sample size, by extending the Bernstein copula proposed in Sancetta & Satchell [2004]. Also, we give a PQD-constrained nonparametric copula estimator that is useful either when it is known that copula satisfies PQD or to approximate the p-values when testing for PQD property.

There has been relatively scarce work on methodologies to address the testing problem (ii) based on a fully nonparametric approach. Different approaches have been developed for testing PQD between two random variables. Scaillet [2005] proposes a Kolmogorov-Smirnov type test for PQD that is based on the distance between the independent copula and empirical copula estimator. Multiplier and bootstrap methods are applied to calculate the p-value. Both methods rely on an approximation of the asymptotic distribution rather than the finite-sample distribution of the test statistic. Gijbels et al. [2010] further extend this test by considering two more distance measures, various nonparametric copula estimators for the test statistic and drawing samples from the independent copula to approximate the distribution of test statistic under the null hypothesis of PQD. Since the null hypothesis is not independence, such resampling approach has its limitation in reflecting the dependence of the distribution of test statistic on the underlying copula. Two other resampling procedures built on parametric and nonparametric estimations of PQD copula are discussed in Gijbels & Sznajder [2013]. The idea of resampling from the PQD-constrained nonparametric copula

estimation is to first get an PQD copula estimator based on the empirical copula and then smooth it to obtain the partial derivative. However, the validity of the nonparametric PQD-constrained copula estimation, which is constructed by integrating the partial derivative, remains an unsolved problem. In contrast, our nonparametric estimation of PQD copula is a valid copula itself and easy to sample from, which automatically leads to a new test for PQD.

The rest of the chapter is organized as follows: in Section 2.2 we present models for general copula and copula satisfying the PQD property. In Section 2.3, we present a novel methodology based on minimizing a Anderson-Darling type goodness-of-fit criteria to estimate the true copula with or without the knowledge of PQD property. In Section 2.4, we present Bootstrap based methodology to test for the PQD property of an unknown copula. We then illustrate the comparative performance of the proposed methodologies in Section 2.5. Section 2.6 provides real case studies. Finally, we make some general comments in Section 2.7.

2.2 Models for Copula

Consider a random sample of independent and identically (i.i.d.) observations $(X_i, Y_i) \overset{i.i.d.}{\sim} H(x,y)$ for $i = 1,2,\ldots,n$, where $H(x,y) = C(F(x), G(y))$ is as defined in (2.1) an unknown joint CDF that is assumed to absolutely continuous with unique copula function $C(u,v)$. Our goal is to derive a sequence of models that will be able to approximate any copula with or without the PQD constraint. As the PQD is primarily a property defined for bivariate copula, we restrict our exposition to only the bivariate case but later in Section 2.7 we provide some ideas on possible generalizations to higher dimensions.

A Bernstein copula is from a family of copulas defined in terms of Bernstein polynomials (Sancetta & Satchell [2004]). It provides a flexible class of models that can be used to uniformly approximate any copula including those that are restricted to satisfy the PQD

property. For simplicity of notations, we consider a class of bivariate Bernstein copula defined on $[0,1]^2$ with degrees $m=(m_1,m_2)$ and parameter vector $\theta=(\theta_{11},\theta_{12},\ldots,\theta_{m_1,m_2})$, defined as

$$C_m(u,v;\theta)=\sum_{k_1=0}^{m_1}\sum_{k_2=0}^{m_2}\theta_{k_1,k_2}\binom{m_1}{k_1}u^{k_1}(1-u)^{m_1-k_1}\binom{m_2}{k_2}v^{k_2}(1-v)^{m_2-k_2}. \tag{2.5}$$

Next we present a set of sufficient conditions for the parameter coefficients $\theta_{k_1,k_2}, k_1 = 0,1,...,m_1, k_2 = 0,1,...,m_2$ ensuring that C_m as defined in eq. (2.5) is a valid copula and additional conditions ensuring the PQD property. Recall that a bivariate copula defined on $[0,1]^2$ must satisfy the following properties:

(a) $C(u,0)=C(0,v)=0,\ \forall u,v\in[0,1]$,

(b) $C(u,1)=u, C(1,v)=v,\ \forall u,v\in[0,1]$,

(c) $\Delta(u_1,u_2,v_1,v_2)\equiv C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1)\geq 0$,

$\forall u_1,u_2,v_1,v_2\in[0,1]$, such that $u_1\leq u_2, v_1\leq v_2$,

(c)′ $\frac{\partial^2 C(u,v)}{\partial u\partial v}\geq 0\ \forall u,v\in[0,1]$.

Notice that (c)′ is an equivalent condition for (c) if we assume that $C(u,v)$ is twice differentiable. The following theorem, which is an extension to PQD of the Lemma 2.1 in Segers et al. [2017], gives sufficient conditions on $\theta_{k_1k_2}$ that ensures C_m to be a PQD copula.

Theorem 2.1. *Consider the class of functions defined by equation (2.5) with parameter θ_{k_1,k_2}'s satisfying the following properties:*

(a) $\theta_{k_1,0}=\theta_{0,k_2}=0$, for $0\leq k_1\leq m_1, 0\leq k_2\leq m_2$,

(b) $\theta_{k_1,m_2}=\frac{k_1}{m_1}$ and $\theta_{m_1,k_2}=\frac{k_2}{m_2}$, for $0\leq k_1\leq m_1, 0\leq k_2\leq m_2$, and

(c) $\theta_{k_1+1,k_2+1}-\theta_{k_1,k_2+1}-\theta_{k_1+1,k_2}+\theta_{k_1,k_2}\geq 0$, for $0\leq k_1\leq m_1-1, 0\leq k_2\leq m_2-1$,

then C_m is a copula. Further, if

(d) $\theta_{k_1,k_2}\geq\frac{k_1}{m_1}\frac{k_2}{m_2}$, for $0\leq k_1\leq m_1, 0\leq k_2\leq m_2$.

Then C_m is a PQD copula for any integers $m_1,m_2>1$.

Notice that by using condition (a) above we could consider the reduced parameter vector

$\theta = (\theta_{11}, \theta_{12}, \ldots, \theta_{m_1,m_2})$ by dropping those with subscripts 0's, but we keep the indices with zero's just for notational convenience to be used later when defining copula density.

Proof of Theorem 2.1 is given in Section 2.8.1. Note that C_m is the independence copula satisfying $C_m(u, v; \theta) = uv$ for any $m_1, m_2 > 1$ when $\theta_{k_1,k_2} = (k_1/m_1)(k_2/m_2)$, i.e., condition (d) in Theorem 2.1 is satisfied with equality. If the condition (d) in Theorem 2.1 is dropped then C_m is a valid unconstrained copula and it can be used to uniformly approximate any arbitrary continuous bivariate copula $C(u, v)$ as $m_1, m_2 \to \infty$ (which we will abbreviate by writing $m \to \infty$ henceforth). The following theorem is the bivariate case of Lemma 3.2 in Segers et al. [2017].

Theorem 2.2. *For any continuous copula C, there exists a Bernstein copula C_m such that*

$$\sup_{u,v\in[0,1]} |C(u, v) - C_m(u, v; \theta)| \to 0 \text{ as } m_1 \wedge m_2 \to \infty$$

Proof of Theorem 2.2 is given in Section 2.8.2.

We will show that the Bernstein copula with appropriate choice of degrees $m = (m_1, m_2)$ and estimation of the coefficients $\theta_{k_1,k_2}, k_1 = 1, ..., m_1, k_2 = 1, ..., m_2$ is capable of approximating any arbitrary continuous copula.

2.3 Anderson-Darling Discrepancy Based Criteria for Copula Estimation

In previous section, we have established that the sequence of models spanned by $\{C_m(\cdot; \theta) : \theta$ satisfies (a)-(c) in Theorem 2.1$\}$ as $m \to \infty$ is dense in the space of all continuous copula functions. Moreover, if we add the condition (d), such a sequence of models is dense in the space of all continuous PQD copula functions. Thus, for statistical inference what remains is an empirical method to estimate the coefficients (satisfying the conditions (a)-(c) or (a)-(d) for PQD in Theorem 2.1) and choose a suitable sequence $m(n) = (m_1(n), m_2(n))$

such that $m(n) \to \infty$ and $n \to \infty$ for the large sample consistency of the estimated copula function.

As illustrated in the proof of Theorem 2.2, the coefficient of the Bernstein copula θ_{k_1,k_2} can possibly be estimated by the empirical copula $C_n(k_1/m_1, k_2/m_2)$ for fixed degrees $m = (m_1, m_2)$, which leads to a copula estimator called the empirical Bernstein copula (EBC) in Sancetta & Satchell [2004]. However, EBC is not guaranteed to be a valid copula, much less a valid PQD copula. Besides, it is not obvious how to determine $m = (m_1, m_2)$ as a function of the observed data. We, thus propose a new approach for the estimation procedure, where the estimation of the coefficients θ_{k_1,k_2} is obtained by minimizing an Anderson-Darling type goodness-of-fit measure between the empirical copula and Bernstein copula under the conditions given in Theorem 2.1. Such estimation method enables the estimated Bernstein copula to be a valid copula or a valid PQD copula for every finite sample size n and for any degrees (m_1, m_2). Furthermore, we develop an algorithm that takes the data into account to automatically choose the degrees of the Bernstein copula $m = (m_1, m_2)$.

Consider again an i.i.d. sample $(X_1, Y_1), ..., (X_n; Y_n) \sim H$ with marginals $X_i \sim F$ and $Y_i \sim G, i = 1, ..., n$. We define the empirical copula as

$$C_n(u, v) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbf{1}\{F_n(X_i) \leq u, G_n(Y_i) \leq v\}, \tag{2.6}$$

where $F_n(x) = \sum_{i=1}^{n} \mathbb{I}(X_i \leq x)/(n+1)$ and $G_n(y) = \sum_{i=1}^{n} \mathbb{I}(Y_i \leq y)/(n+1)$ are the marginal empirical CDFs. The modification $1/(n+1)$ instead of $1/n$ modifies the common empirical copula to be away from 1 in order to reduce potential problems at boundaries. For notational convenience, we define the following quantities

$$U_i \equiv U_{in} = F_n(X_i), \;\; V_i \equiv V_{in} = G_n(Y_i).$$

Following the work by Gijbels et al. [2010], these pseudo-observations (U_i, V_i)'s are approximately distributed as the vector $(F(X), G(Y)) \sim C$. Although these pseudo-observations are

no longer independently distributed, the approximate expressions for the mean and the variance of $C_n(u,v)$ are given by

$$E(C_n(u,v)) \approx \frac{n}{n+1} C(u,v) \text{ and } V(C_n(u,v)) \approx \frac{n}{(n+1)^2} C(u,v)(1-C(u,v)),$$

where the approximations are valid (uniformly) up to an error term of $O(\sqrt{\log\log n/n})$ (which, e.g. follows by using the Lemma 1 in [Janssen et al., 2012]). The above approximations motivate us to consider an Anderson-Darling type goodness-of-fit type discrepancy to estimate the parameter θ_m:

$$\begin{aligned} AD_n(\theta_m) &= \frac{1}{n}\sum_{i=1}^{n} \frac{(C_n(U_i,V_i) - \frac{n}{n+1} C_m(U_i,V_i,\theta_m))^2}{\frac{n}{(n+1)^2} C_n(U_i,V_i)(1-C_n(U_i,V_i)) + \xi_{m,n}} \\ &= \int_{I^2} \frac{(C_n(u,v) - \frac{n}{n+1} C_m(u,v,\theta_m))^2}{\frac{n}{(n+1)^2} C_n(u,v)(1-C_n(u,v)) + \xi_{m,n}} d\,C_n(u,v), \end{aligned} \tag{2.7}$$

where $\theta_m = (\theta_{0,0}, \theta_{0,1}, \ldots \theta_{0,m_2}, \theta_{1,0}, \ldots \theta_{m_1,m_2})$ is a $(m_1+1)\times(m_2+1)$ coefficient vector, $C_m(u,v,\theta_m)$ is the Bernstein copula depending on θ_m. We define $\xi_{m,n} = \max(\xi_m, 1/n^2) \to \xi_m > 0$ as $n \to \infty$ where $\xi_m = O(1)$ as $m \to \infty$ and $\liminf_{m\to\infty} \xi_m > 0$. The weight factor $\frac{n}{(n+1)^2} C_n(U_i,V_i)(1-C_n(U_i,V_i))$ in the denominator of the Anderson-Darling type distance $AD_n(\theta_m)$ inflates the criteria at the extreme values of the copula so we use the correction factor $\xi_{m,n}$ to ensure that the denominator remains positive. When the sample size n is relatively small, $\xi_{m,n}$ is expected to have a minor effect on the denominator and the weight factor plays a significant role. When the sample size n becomes relatively larger, the denominator is guaranteed to be positive with $\xi_{m,n} > 0$.

Next we consider the limiting value of the criteria in eq. (2.7) given by

$$\begin{aligned} AD(\theta_m) &= \int_{I^2} \frac{(C(u,v) - C_m(u,v,\theta_m))^2}{\xi_m} d\,C(u,v) \\ &= E_C\left[\frac{(C(U,V) - C_m(U,V,\theta_m))^2}{\xi_m}\right], \end{aligned} \tag{2.8}$$

where the last expectation is taken over $(U,V)\sim C(u,v)$. Now based on the uniform almost sure convergence of empirical distribution function we have $C_n(u,v)\to C(u,v)$ and by continuous mapping theorem (Theorem 7.7 on p.109, and its extensions Theorem 7.24 and Theorem 7.25 on p.117-118 of Kosorok [2007]) we claim that

$$AD_n(\theta_m)\xrightarrow{\mathrm{P}} AD(\theta_m)\ \ as\ n\to\infty, \tag{2.9}$$

for fixed $m=(m_1,m_2)$. Let us define

$$\hat{\theta}_m=\underset{\theta_m\in\Theta_m}{\operatorname{argmin}}\, AD(\theta_m). \tag{2.10}$$

and the estimation of the coefficients that minimizes the Anderson-Darling discrepancy, i.e,

$$\hat{\theta}_{m,n}=\underset{\theta_m\in\Theta_m}{\operatorname{argmin}}\, AD_n(\theta_m), \tag{2.11}$$

where $\Theta_m=\{\theta_m:\theta_{k_1,0}=\theta_{0,k_2}=0,\theta_{k_1,m_2}=\frac{k_1}{m_1},\theta_{m_1,k_2}=\frac{k_2}{m_2}$, for $0\le k_1\le m_1, 0\le k_1\le m_2, \theta_{k_1+1,k_2+1}-\theta_{k_1,k_2+1}-\theta_{k_1+1,k_2}+\theta_{k_1,k_2}\ge 0$ for $0\le k_1\le m_1-1$, $0\le k_1\le m_2-1\}$, which satisfies the conditions (a)-(c) in Theorem 2.1 in order to make the estimated Bernstein copula a valid copula. The following theorem gives the consistency of the unconstrained copula estimator $C_m(u,v,\hat{\theta}_{m,n})$.

Lemma 2.1. *For any $\epsilon>0$, assume that $\inf_{\|\theta_m-\hat{\theta}_m\|>\epsilon} AD(\theta_m)>AD(\hat{\theta}_m)$. Then for fixed $m=(m_1,m_2)$, we have*

$$\hat{\theta}_{m,n}\xrightarrow{P}\hat{\theta}_m\ \ as\ n\to\infty.$$

Theorem 2.3. *Assume that $\xi_m=O(1)$ as $m\to\infty$ and $\liminf_{m\to\infty}\xi_m>0$. Then for any $u,v\in[0,1]$, we have*

$$|C_m(u,v,\hat{\theta}_{m,n})-C(u,v)|\xrightarrow{p}0\ \ as\ n,m\to\infty.$$

Proofs of Lemma 2.1 and Theorem 2.3 is given in the Section 2.8.3 and 2.8.4. Notice that

with the PQD condition (d) given in Theorem 1, we can further obtain a PQD-constrained copula estimation when it is known to satisfy PQD property.

We separate the estimation procedure into two parts. The first part is to obtain the estimation of coefficients $\hat{\theta}_{m,n}$ for fixed $m=(m_1,m_2)$ by minimizing the Anderson-Darling type discrepancy $AD_n(\theta_m)$, where the degrees of the Bernstein copula $m=(m_1,m_2)$ depend on the sample size n as $m_1=\lceil n^{\alpha_1}\rceil, m_2=\lceil n^{\alpha_2}\rceil, 0<\alpha_1,\alpha_2<1$. The second part is to search for an appropriate choice of $m=(m_1,m_2)$. Since larger degrees of the Bernstein copula lead to lower biases but inflate computational inefficiency and increased variance, a relatively small $m=(m_1,m_2)$ that is sufficient for a practically 'good' copula estimation is desirable. So, we have used cross validation type criteria to select optimal m's based on a selected grid of integer pairs (see Steps 1-3 later for further details), which is admittedly bit ad hoc but the procedure is much more computationally efficient compared to full search over all possible integer pairs (especially when n is large).

For fixed $m=(m_1,m_2)$, let us denote

$$B_{k_1,k_2}(u,v)=\binom{m_1}{k_1}u^{k_1}(1-u)^{m_1-k_1}\binom{m_2}{k_2}v^{k_2}(1-v)^{m_2-k_2}. \tag{2.12}$$

and $B_m(u,v)=(B_{0,0}(u,v),B_{0,1}(u,v),\ldots B_{0,k_2}(u,v),B_{1,0}(u,v),\ldots B_{k_1,k_2}(u,v))$. Then the Bernstein copula $C_m(u,v,\theta_m)$ can be written as

$$C_m(u,v,\theta_m)=B_m(u,v)^T\theta_m. \tag{2.13}$$

and the Anderson-Darling type discrepancy that we want to minimize can be rewritten as

$$\begin{aligned} AD_n(\theta_m) &= \frac{1}{n}\sum_{i=1}^{n}\frac{(C_n(U_i,V_i)-\frac{n}{n+1}C_m(U_i,V_i,\theta_m))^2}{\frac{n}{(n+1)^2}C_n(U_i,V_i)(1-C_n(U_i,V_i))+\xi_{m,n}} \\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{(C_i-\frac{n}{n+1}B_i^T\theta_m)^2}{\frac{n}{(n+1)^2}C_i(1-C_i)+\xi_{m,n}} \\ &= \frac{1}{2}\theta_m^T D\theta_m - d^T\theta_m + \text{terms not involving } \theta_m \end{aligned} \tag{2.14}$$

where $C_i = C_n(U_i, V_i)$, $B_i = B_m(U_i, V_i)$ and

$$D = \frac{2n}{(n+1)^2}\sum_{i=1}^{n}\frac{B_i B_i^T}{\frac{n}{(n+1)^2}C_i(1-C_i)+\xi_{m,n}},\ d = \frac{2}{n+1}\sum_{i=1}^{n}\frac{C_i B_i}{\frac{n}{(n+1)^2}C_i(1-C_i)+\xi_{m,n}}.$$

This is a quadratic programming problem that minimizes $\frac{1}{2}\theta_m^T D\theta_m - d^T\theta_m$ subject to the conditions given in Theorem 1, which can be written in the form of $A^T\theta_m \geq b_0$. For example, if we want to estimate copula with PQD constraint when $m_1 = m_2 = 3$, $\theta_{00} = \theta_{0,1} = \theta_{0,2} = \theta_{0,3} = \theta_{1,0} = \theta_{2,0} = \theta_{3,0} = 0, \theta_{1,3} = \theta_{3,1} = 1/3, \theta_{2,3} = \theta_{3,2} = 2/3, \theta_{3,3} = 1$. Let

$$R = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, S = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/9 & 2/9 & 1/3 \\ 0 & 2/9 & 4/9 & 2/3 \\ 0 & 1/3 & 2/3 & 1 \end{bmatrix},$$

then A and b_0 take the form

$$A = \begin{bmatrix} R & S & \mathbf{0}_{3\times4} & \mathbf{0}_{3\times4} \\ \mathbf{0}_{3\times4} & R & S & \mathbf{0}_{3\times4} \\ \mathbf{0}_{3\times4} & \mathbf{0}_{3\times4} & R & S \\ & & & \\ & & I_{16\times16} & \\ & & & \end{bmatrix}, b_0 = \begin{bmatrix} \mathbf{0}_{9\times1} \\ \text{vec}(T) \end{bmatrix},$$

respectively. The estimation of θ_m can be obtained by applying the function `solve.QP` in R package `quadprog`.

Besides, we develop an algorithm for choosing data-based optimal degrees $m = (m_1, m_2)$. Janssen et al. [2014] show that a good (asymptotic) choice for m of the empirical Bernstein estimator is $m = n^\alpha$ with $\alpha \in [\frac{1}{3}, \frac{2}{3}]$. Our copula estimator is guaranteed to be a genuine copula for any degrees m_1, m_2, so we further explore $m_1 = \lceil n^{\alpha_1} \rceil$, $m_2 = \lceil n^{\alpha_2} \rceil$ with α_1, α_2 between

$0.3(<\frac{1}{3})$ and $0.7(>\frac{2}{3})$, and m_1 and m_2 are allowed to vary. The degrees are systematically searched on an equally spaced grid of powers of n (generating distinct integer values of m_1 and m_2) where the powers are increased by 0.05. Given a finite sample size n, we consider $m_1, m_2 \in \{\lceil n^a \rceil : a = 0.3, 0.35, 0.4, 0.45, ..., 0.7\}$. As a consequence, there are a total of $9 \times 9 = 81$ possible combinations of $m = (m_1, m_2)$, which can be denoted as entries of a 9×9 matrix M, i.e.,

$$M = \begin{bmatrix} (n^{0.3}, n^{0.3}) & (n^{0.3}, n^{0.35}) & (n^{0.3}, n^{0.34}) & \cdots & (n^{0.3}, n^{0.7}) \\ (n^{0.35}, n^{0.3}) & (n^{0.35}, n^{0.35}) & & & \vdots \\ (n^{0.4}, n^{0.3}) & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ (n^{0.7}, n^{0.3}) & \cdots & \cdots & \cdots & (n^{0.7}, n^{0.7}) \end{bmatrix},$$

Let m_{ij} denote the entry of M at location (i, j). Then the algorithm to obtain optimal value of $m = (m_1, m_2)$ can be summarized as the following three steps:

Step 1: Given a sample of size n, starting from the degrees $m = m_{11}$ at location $(1,1)$ in M, obtain the estimated coefficients $\hat{\theta}_{m,n}$ and calculate the Anderson-Darling based criteria

$$\begin{aligned} D_m(\hat{\theta}_{m,n}) &= \int_0^1 \int_0^1 \frac{(C_n(u,v) - C_m(u,v,\hat{\theta}_{m,n}))^2}{C_m(u,v,\hat{\theta}_{m,n})(1 - C_m(u,v,\hat{\theta}_{m,n})) + \xi_{m,n}} d\,u\,d\,v \\ &\approx \frac{1}{K^2} \sum_{k=1}^{K} \sum_{j=1}^{K} \frac{(C_n(u_k,v_j) - C_m(u_k,v_j,\hat{\theta}_{m,n}))^2}{C_m(u_k,v_j,\hat{\theta}_{m,n})(1 - C_m(u_k,v_j,\hat{\theta}_{m,n})) + \xi_{m,n}}, \end{aligned} \tag{2.15}$$

where $u_i = v_i = i/K$, $K = 100$ and $C_n(u,v)$ is the common empirical copula estimator. Notice that $D_m(\cdot)$ is calculated on a regular grid of values (u_k, v_j) and not necessarily at observed data points (U_i, V_i) to better approximate the integral.

Step 2: Repeat step 1 for m_{12}, m_{21} located at the second diagonal line in M and record the one with minimal value of $D_m(\hat{\theta}_{m,n})$ as m'. m' is compared with the m recorded in

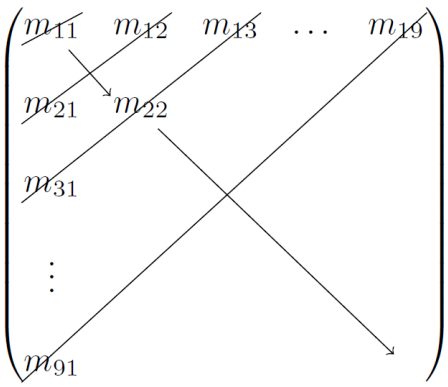


Figure 2.1 The algorithm for choosing degrees $m = (m_1, m_2)$.

the previous step and if

$$|D_m(\hat{\theta}_{m,n}) - D_{m'}(\hat{\theta}_{m',n})| \leq \epsilon D_m(\hat{\theta}_{m,n}), \tag{2.16}$$

where ϵ is a suitably chosen small number, then we stop searching and pick up m as the degrees of the Bernstein copula. In our numerical illustrations we have used $\epsilon = 0.01$ (i.e. we stop if the change in relative difference is only 1%).

Step 3: If condition (2.16) is not satisfied, repeat step 2 for m_{13}, m_{22}, m_{31} located at the third diagonal line and the next diagonal lines until we find a desirable choice of m that satisfies (2.16).

The searching procedure is shown in Figure 2.1. By doing so, we are able to choose a data-dependent parsimonious $m = (m_1, m_2)$ that is motivated by the asymptotic theory. In many practical scenarios (and simulated data settings) we have found that we the algorithm often stops with much fewer than all 81 tries.

2.4 Testing for PQD property

In previous sections we have developed methodologies to obtain nonparametric estimate of the copula when there is substantial background information to believe that the underlying

copula satisfies the PQD property. However, in many applications, we may not know for sure that such intrinsic property of a copula is truly valid. For that reason, we develop a methodology to test the hypothesis if the true underlying copula satisfies the PQD property by using a test of hypotheses framework. We consider testing the following null hypothesis for PQD:

$$H_0 : \forall\, u, v \in [0,1] \;\; C(u,v) \geq uv, \tag{2.17}$$

against the alternative

$$H_a : \exists\, u, v \in [0,1] \;\; C(u,v) < uv. \tag{2.18}$$

Admittedly, the alternative is not very sharp to develop a test for the above scenario and we would need some regularity conditions to separate the two hypotheses so that any test to be developed will have reasonable statistical power to make a decision (see assumption (A) stated later). Given a sample $(X_i, Y_i), i = 1, ..., n$, we denote the PQD-constrained and unconstrained copula estimators as $\hat{C}_{1,m,n}$ and $\hat{C}_{2,m,n}$, respectively. Then a Kolmogorov-Smirnov type test statistic based on discrepancy between the unconstrained estimated copula and independent copula is constructed as

$$\Delta_n = \sqrt{n} \max_{0 \leq u,v \leq 1} (uv - \hat{C}_{2,m,n}(u,v)). \tag{2.19}$$

Notice that $\Delta_n \leq 0$ under the null hypothesis of PQD, while $\Delta_n > 0$ under the alternative. So, we can construct a test that would reject the null hypothesis of PQD if

$$\Delta_n > c_{1-\alpha,n}, \tag{2.20}$$

where $c_{1-\alpha,n}$ is the $(1-\alpha)$-th quantile of the test statistic Δ_n under H_0 satisfying $\Pr[\Delta_n > c_{1-\alpha,n} | H_0] \leq \alpha$. The distribution of the test statistic Δ_n can be approximated by the bootstrap method. More succinctly, to approximate the sampling distribution of Δ_n under the null hypothesis of PQD, M samples each of size n are drawn from the PQD-constrained

estimated copula $\hat{C}_{1,m,n}$. For each bootstrap sample $\{(U_i^*, V_i^*), i = 1, ..., n\}$, we obtain copula estimator $\hat{C}^*_{2,m,n}$ without PQD constraint, and compute

$$\Delta_n^* = \sqrt{n} \max_{0 \le u,v \le 1} (uv - \hat{C}^*_{2,m,n}(u, v)). \tag{2.21}$$

Then, the critical value of the test $c_{1-\alpha,n}$ can be approximated by the $(1-\alpha)$-quantile of Δ_n^* based on its M bootstrapped values.

Following the derivation of the Bernstein copula density from the empirical Bernstein copula, see e.g. Bouezmarni et al. [2010] and Janssen et al. [2014], the corresponding copula density of our copula estimator can be written as

$$\frac{\partial^2 C_m(u, v)}{\partial u \partial v} = m_1 m_2 \sum_{k_1=0}^{m_1-1} \sum_{k_2=0}^{m_2-1} w_{k_1,k_2} \binom{m_1-1}{k_1} u^{k_1}(1-u)^{m_1-k_1-1} \binom{m_2-1}{k_2} v^{k_2}(1-v)^{m_2-k_2-1}, \tag{2.22}$$

where $w_{k_1,k_2} = \theta_{k_1+1,k_2+1} - \theta_{k_1,k_2+1} - \theta_{k_1+1,k_2} + \theta_{k_1,k_2}$. Since the sum of $w_{k_1,k_2}, k_1 = 0, ..., m_1-1, k_2 = 0, ..., m_2-1$ equals to 1, a sample $(U, V) \sim C_m(u, v)$ can be drawn as follows.

$$\begin{aligned} (k_1, k_2) &\sim w_{k_1,k_2},\ k_1 = 0, ..., m_1-1, k_2 = 0, ..., m_2-1 \\ U &\sim \text{Beta}(k_1+1, m_1-k_1) \\ V &\sim \text{Beta}(k_2+1, m_2-k_2) \end{aligned} \tag{2.23}$$

Thus, given a sample $(X_i, Y_i), i = 1, ..., n$, we can first obtain the PQD-constrained estimated copula $\hat{C}_{1,m,n}$, and then draw M samples from $\hat{C}_{1,m,n}$ to mimic the distribution of Δ_n under the null hypothesis of PQD.

Similar to Theorem 1 in Gijbels et al. [2010], Theorem 2.4 provides the consistency of the proposed testing procedure. Similar to the condition (C4) given in Gijbels et al. [2010], the following regularity assumption (A) is required to show the consistency of the testing procedure:

Assumption (A): Let μ_C be the measure associated with a copula C, λ_2 is the Lebesgue

measure on $[0,1]^2$, $I_0 = \{(u,v) \in [0,1]^2 : C(u,v) = uv\}$, *and* ∂I_0 *the boundary of the set* I_0. *Then* $\mu_C(\partial I_0) = \lambda_2(\partial I_0) = 0$.

In addition to the Kolmogorov-Smirnov (KS) distance, we can also consider various other test statistics based on desired discrepancy measures, e.g., Cramér-von Mises (CvM) and Anderson-Darling (AD) discrepancies given in Gijbels et al. [2010]. The CvM type and AD type test statistics take the form

$$\Delta_n^{CvM} = n \int_{I^2} (uv - \hat{C}_{2,m,n}(u,v))_+^2 \, d\hat{C}_{2,m,n}(u,v) \tag{2.24}$$

and

$$\Delta_n^{AD} = n \int_{I^2} \frac{(uv - \hat{C}_{2,m,n}(u,v))_+^2}{uv(1-u)(1-v)} \, d\hat{C}_{2,m,n}(u,v), \tag{2.25}$$

separately, where $(\cdot)_+ = \max(\cdot, 0)$ and $\hat{C}_{2,m,n}$ is the unconstrained copula estimation.

Theorem 2.4. *Assume that the third order partial derivatives for* C *are bounded on* $(0,1)^2$. *If* $n^{\frac{1}{2}} \tilde{m}^{-1} \to 0$, *where* $\tilde{m} = \min(m_1, m_2)$, *then* $\sqrt{n}(\hat{C}_{2,m,n} - C)$ *converges weakly to a mean zero Gaussian process on* $(0,1)^2$. *Further if the assumption (A) holds, we have that*

(i) if H_0 *is true, then*

$$\lim_{n\to\infty} \mathbb{P}(\text{ reject } H_0) = \lim_{n\to\infty} \mathbb{P}(\Delta_n > c_{1-\alpha,n}) \leq \alpha,$$

(ii) if H_0 *is false, then*

$$\lim_{n\to\infty} \mathbb{P}(\text{ reject } H_0) = 1.$$

The proof of weak convergence of the process $\sqrt{n}(\hat{C}_{2,m,n} - C)$ relies on the proof of Theorem 2 and Remark 3 in Janssen et al. [2012]. Based on the weak convergence and regularity assumption (A), the consistency of the testing procedure can be shown following the proof of Theorem 1 in Gijbels et al. [2010]. Details are given in Section 2.8.5.

2.5 Numerical Illustrations using Simulated Data

2.5.1 Estimation of PQD Copula

We investigate the finite-sample behavior of our nonparametric estimation for PQD copula. Two copula families with a wide range of parameter values are considered. The first one is the Gumbel copula family defined as

$$C_G(u,v) = \exp(-((-\ln(u))^\theta + (-\ln(v))^\theta)^{1/\theta}), \tag{2.26}$$

where $\theta \in [1,\infty)$. All members of Gumbel copula family have the property of PQD.

The second one is the Clayton copula family defined as

$$C_C(u,v) = (\max\{u^{-\theta} + v^{-\theta} - 1, 0\})^{-1/\theta}, \tag{2.27}$$

where $\theta \in [-1,\infty)\backslash\{0\}$. $C_C(u,v)$ is a PQD copula when the parameter is greater than 0.

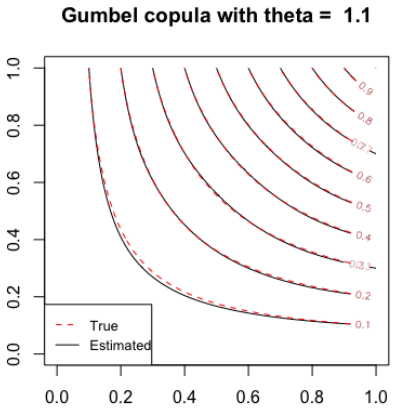
Simulations are based on $N = 100$ Monte Carlo (MC) replications of samples (each of size $n = 100$) drawn from either Gumbel or Clayton copula family. The m_1 and m_2 are chosen from the set of values $\{4,6,7,8,10,13,16,20,26\}$ for the sample size of $n = 100$ (see details as described in Section 3). ξ_m is set to be 0.001 such that the weight factor does not affect too much at the extreme values of the copula when the sample size n is relatively large. Thus, we have $\xi_{m,n} = \max(\xi_m, 1/n^2) = 0.001$. Note that ϵ is the threshold of the percentage change of the Anderson-Darling based criteria $D_m(\hat{\theta}_{m,n})$ in the algorithm for choosing degrees. To reduce computational burdens and to strike a balance between biases and variances, we select reasonably large values of $m = (m_1, m_2)$ as described in our algorithm that allows for bias-variance trade-of. By setting $\epsilon = 0.01$ we stop the algorithm when the percentage change falls below 1% in terms of the $D_m(\hat{\theta}_{m,n})$. Other thresholds can be used as well but we have found this to provide visually satisfactory results in our simulation

studies across many copula families and sample sizes (not all reported here). We consider $\theta = 1.1, 1.25, 1.5, 2$ for Gumbel copulas and $\theta = 0.25, 0.5, 1, 1.5$ for Clayton copulas to explore varying levels of dependence as measured by Kendall's tau or Spearman's rho which are one-to-one functions of the dependence parameter θ. The average of the estimated contours of the copula functions (across $N = 100$ MC replications) are given in Figure 2.2 and Figure 2.3. The contour plots of the true underlying copula are also included. The average of data-based selected m's denoted by $\bar{m}$ are also reported in each case. The contour plots reveal a very good match between the true copula and the average of estimated copula, indicating almost no biases uniformly across all contour lines.

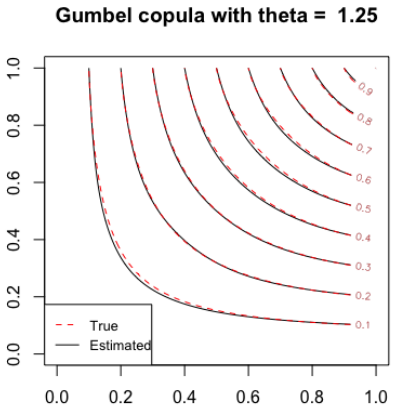
Moreover, the data-based selection of the degrees m's appear to be relatively robust across different simulated cases with slightly larger values of degrees being selected for larger dependence parameter θ. Thus, adaptive choice of degrees is a better choice than any ad-hoc predetermined choices of the degrees. This aspect can be considered as one of the major contributions of the proposed estimation method.

2.5.2 Testing for PQD

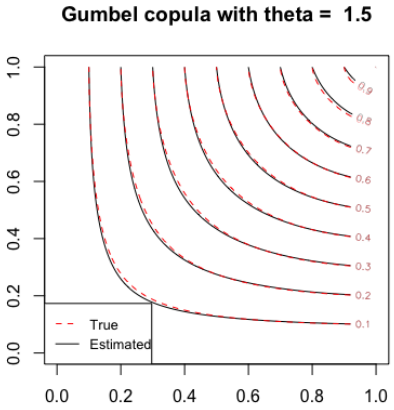
The finite-sample performance of the testing procedures is investigated through a Monte Carlo simulation study. We consider three types of test statistics including Kolmogorov-Smirnov (KS), Cramér-von Mises (CvM) and Anderson-Darling(AD) in this simulation study. The significance level is 0.05 and the sample size is $n = 200$. $M = 1000$ samples are drawn from the PQD-constrained copula estimator to approximate the p-value. Computation of the power and size is based on 1000 replicated samples drawn from the corresponding copula under different hypotheses. For the Kolmogorov-Smirnov type test statistic, the maximum is searched for on an equally spaced grid of points where the distance between two consecutive points is 0.05, which leads to a 20×20 grid of points. Simulation settings follow the set-up as in Gijbels & Sznajder [2013] so that we can compare the proposed testing procedure (referred to as adaptive AD Bernstein) with the nonparametrically built



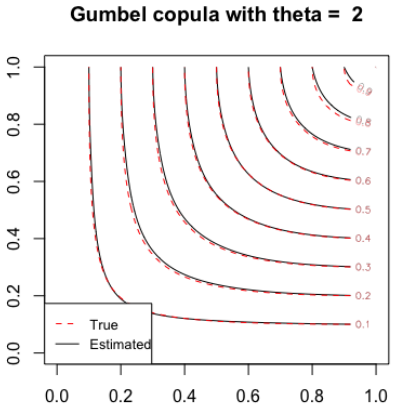
(a) $\bar{m}$=(9.12, 9.76)



(b) $\bar{m}$=(9.63, 10.67)

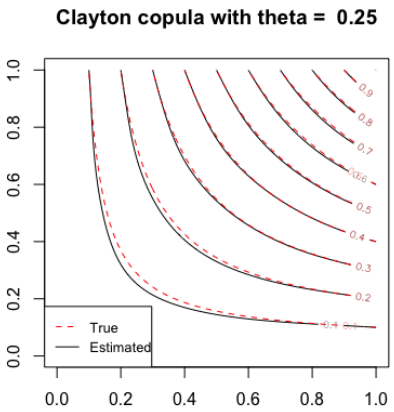


(c) $\bar{m}$=(11.10, 11.02)

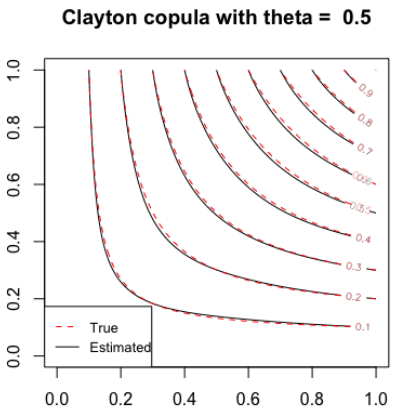


(d) $\bar{m}$=(11.65, 11.72)

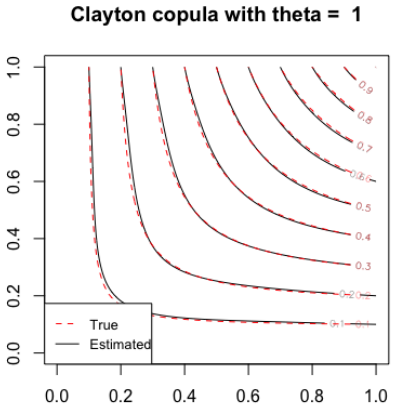
Figure 2.2 Proposed nonparametric estimates of copula with data (of size $n = 100$) generated from the Gumbel family with varying labels of the dependence parameter θ. The average of data-based selected degrees (across $N = 100$ Monte Carlo replicates) are also presented in each cases.



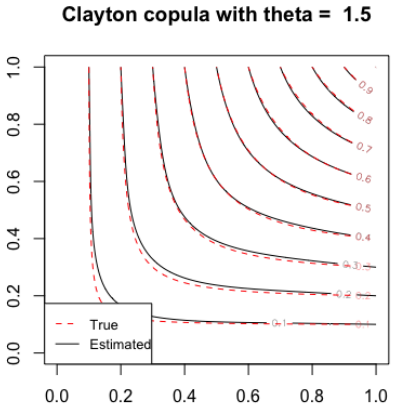
(a) $\bar{m}$=(7.97, 6.95)



(b) $\bar{m}$=(7.56, 7.94)



(c) $\bar{m}$=(11.15, 11.67)



(d) $\bar{m}$=(12.22, 10.72)

Figure 2.3 Proposed nonparametric estimates of copula with data (of size $n = 100$) generated from the Clayton family with varying labels of the dependence parameter θ. The average of data-based selected degrees (across $N = 100$ Monte Carlo replicates) are also presented in each cases.

testing procedure (referred to as GS Nonparametric) in that paper.

2.5.2.1 Size Study

This section aims to investigate the actual size of the tests. Besides the independent copula, the finite-sample performance for two other examples of PQD copulas are also evaluated. The first one is a Frank copula with parameter 0.5 and the second one is a Clayton copula with parameter 0.11. Both copulas give a theoretical Kendall's tau equal to 0.05.

Table 2.1 Nominal size of the tests based on $n = 200$ samples and 1000 Monte Carlo simulated data sets with a target size of 0.05.

Copula	Test Statistic	GS Nonparametric	adaptive AD Bernstein
Π	KS	0.047	0.048
	CvM	0.046	0.053
	AD	0.050	0.041
Frank	KS	0.011	0.010
$\theta = 0.5$	CvM	0.010	0.010
	AD	0.009	0.010
Clayton	KS	0.006	0.014
$\theta = 0.11$	CvM	0.004	0.010
	AD	0.004	0.010

The simulation results in Table 2.1 suggests that the proposed testing procedure with three different test statistics has low sizes under all of the considered situations. The significant level is approximately held for finite sample size when the true underlying copula is the Π copula. Our testing procedure (adaptive AD Bernstein) has larger sizes than the nonparametrically built testing procedure of Gijbels & Sznajder [2013] (GS Nonparametric) when the true model comes from Clayton family, but it still under the significant level.

2.5.2.2 Power Study

Several copula models are considered in the power study. These include models from classical copula families including Frank and Clayton, a family of mixtures of Frank copulas, an asymmetric copula and a special copula where the null hypothesis of PQD is violated at a subset of the unit square.

The first two copula examples are from classical copula families. One is a Frank copula with parameter -1 and the other is a Clayton copula with parameter -0.2. Both copulas characterize a mild negative dependence with the value of Kendall's tau equal to -0.11. They violate the PQD condition by simply being negative quadrant dependents (NQD), i.e., the whole underlying copulas are below the Π copula.

Since PQD is a global feature, it is also of interest to explore how the tests work when the PQD condition is violated locally. The next two examples are mixtures of two Frank copulas, which are neither PQD nor NQD. A mixture of Frank copulas can be written as

$$C^{mF}_{\theta_1,\theta_2,\gamma} = \gamma C^{F}_{\theta_1} + (1-\gamma) C^{F}_{\theta_2}. \tag{2.28}$$

The first parameter set is $(\theta_1, \theta_2, \gamma) = (9.5, -9.5, 0.5)$, which gives a theoretical Kendall's tau equal to zero. The second one is $(\theta_1, \theta_2, \gamma) = (10, -15, 0.6)$ with a theoretical Kendall's tau close to 0.11. Both examples suggest that Kendall's tau is no longer an appropriate assessment of the positive or negative dependence when it comes to copulas which have local PQD or NQD properties.

Next, an asymmetric copula is adopted to evaluate the appropriateness of the models in the resampling procedure. The asymmetric copula is written as

$$C(u,v) = uv + \theta_1(1-v^2)\sin(v\theta_2)u(1-u), \tag{2.29}$$

where the parameters is $(\theta_1, \theta_2) = (-0.12, 8.3)$. It is a locally NQD copula with a theoretical

Kendall's tau close to -0.02.

Finally, a special copula, referred to as copula A in Gijbels & Sznajder [2013], is taken into consideration. It equals to the Π copula on a large set of the unit square except for a small square region, $[0.15, 0.5] \times [0.5, 0.85]$, where the mass that attributes the Π copula to this region is redistributed uniformly along the second diagonal. This results in a NQD copula that is strictly below the Π copula in the indicated small square while equal to the Π copula in the rest regions.

Table 2.2 Nominal power of the tests based on $n = 200$ samples and 1000 Monte Carlo simulated data sets.

Copula	Test Statistic	GS Nonparametric	adaptive AD Bernstein
Frank	KS	0.608	0.733
$\theta = -1$	CvM	0.695	0.800
	AD	0.706	0.810
Clayton	KS	0.573	0.762
$\theta = -0.2$	CvM	0.723	0.830
	AD	0.795	0.876
mixture of Frank I	KS	0.457	0.629
$(\theta_1, \theta_2, \gamma) =$	CvM	0.499	0.638
$(9.5, -9.5, 0.5)$	AD	0.739	0.810
mixture of Frank II	KS	0.317	0.457
$(\theta_1, \theta_2, \gamma) =$	CvM	0.386	0.400
$(10, -15, 0.6)$	AD	0.642	0.638
asymmetric copula	KS	0.386	0.276
$(\theta_1, \theta_2) =$	CvM	0.320	0.295
$(-0.12, 8.3)$	AD	0.500	0.419
copula A	KS	0.242	0.371
	CvM	0.262	0.381
	AD	0.248	0.324

The simulation results of the power study are presented in Table 2.2, where the number is the proportion of the times out of 1000 that the null hypothesis is rejected. It can be

observed that for each of the three different test statistics, our testing procedure has higher power than that of Gijbels & Sznajder [2013] in most cases. "GS Nonparametric" outperforms "adaptive AD Bernstein" when the true copula is the asymmetric copula, though neither of two approaches has a desirable power performance. Compared to the first mixture of Frank copula with a theoretical Kendall's tau equal to 0, the second one has a theoretical Kendall's tau close to 0.11, making it more difficult to distinguish from the PQD. Thus, it is not surprising to see a decrease in the power of both approaches, but loss of the "adaptive AD Bernstein" is much higher.

Figure 2.4 shows the power of the tests as a function of the non-PQD proportion (1 - γ) of the mixture of two Frank copulas ($\theta_1 = 9.5, \theta_2 = -9.5$). Starting with PQD copula (Frank coupla with parameter 9.5) we mix PQD with non-PQD copula (Frank coupla with parameter -9.5) and plot the power curve against the proportion of non-PQD data. From the plot we can see AD type test statistic reacts faster to the infusion of some proportion of non-PQD data than the other two test statistics. KS type and CvM type test statistics have comparable power curves, while AD type test statistic has higher power for moderate values of non-PQD proportions.

Among three distance measures, the best results are almost always from the AD type test statistic, while the KS type test statistic gives the worst results. The only exception is the copula B, for which the CvM type test statistic has the best power performance. Overall, our proposed estimation and resampling approach combined with the AD type test statistic often gives the best power results.

2.6 Real Case studies

2.6.1 Insurance Claim Data

The Danish fire insurance claims is a data set describing 2167 fire insurance claims in Denmark from January 3, 1980 until December 31, 1990. There are three types of claims

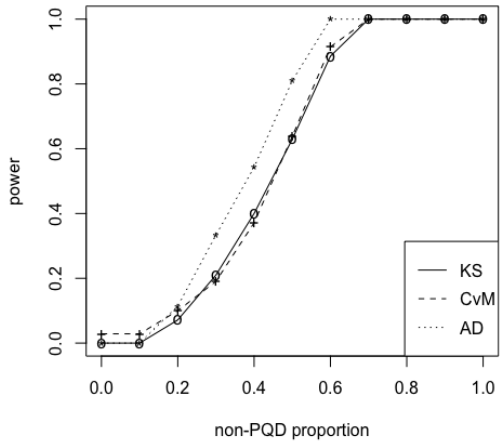


Figure 2.4 Power curve of the tests with three different test statistics as a function of the nonâĂŘPQD mixture proportion (1 - γ) of two Frank copulas ($\theta_1 = 9.5, \theta_2 = -9.5$) in (2.28).

including losses in buildings (B), contents (C) and profits (P). This data set is available in R and a detailed description is provided in Gijbels & Sznajder [2013]. In our case, 517 claims that are positive in all three variables are studied and pairwise copulas of three variables are estimated based on pseudo-observations.

The empirical values of Kendall's tau for (B, C), (B, P) and (C, P) are $0.117, 0.201, 0.462$, separately, which indicate possible PQD properties of each pair. The contour plots of two copula estimations, with or without PQD constraint, are given as (a), (c), and (e) in Figure 2.5. (b), (d) and (f) in Figure 2.5 are the contour plots of the difference between unconstrained copula estimation and independent copula $\hat{C}_{2,m,n}(u,v) - uv$ for each pair. For the pair of (B, C), there is a strong discrepancy between the PQD-constrained and unconstrained estimates in (a), suggesting that PQD might not be an inherent property between the buildings and contents claims. The negative components in (b) indicates some violation of PQD as well. The difference between the PQD-constrained and unconstrained estimates for (B, P) is not as large as that for (B, C) but still obvious. In contrast, two estimated copulas, one with PQD constraint and the other without, overlap for the pair of (C, P) in (e), which is a strong evidence for the PQD structure between the contents and profits claims. This is

also supported by the contour plot of $\hat{C}_{2,m,n} - uv$ that is uniformly positive over the unit cube in (e).

Conclusions from the contour plots can be consolidated by the testing results using proposed testing procedure with three different distances. Table 2.3 shows the p-values for testing PQD between pairwise Danish fire insurance claims. Though using different distances, all the tests suggest the same decision for each pair. For (B, C), there is a strong evidence to reject the null hypothesis of PQD since all the p-values equal to 0. We may also reject the null hypothesis for (B, P) because each test has a small p-value less than 0.05. For (C, P), we don't reject the null hypothesis of PQD. In this case, the test statistics are less (KS) or equal (CvM an AD) to 0, so there is perhaps no need to approximate the null distribution of test statistics by resampling and we conclude that the p-values are equal to 1.

Table 2.3 P-values for proposed PQD tests on pairwise Danish fire insurance claims.

Test Statistic	(B, C)	(B, P)	(C, P)
KS	0	0.005	1
CvM	0	0.006	1
AD	0	0.005	1

2.6.2 Stock Index Data

This data set consists of 252 observations on the stock indices including Dow Jones (DJ), NASDAQ 100 (NASDAQ) and FTSE 100 (FTSE) in 2015 (from January 2, 2015 to December 31, 2015). Pairwise scatterplots of the stock index data and the corresponding pseudo-observations are presented in Figure 2.6. From the plots we can see that the pseudo-observations of (DJ, FTSE) have a concentration around the positive diagonal, which indicates a strong positive dependence structure. The empirical value of Kendall's tau for the pair of (DJ, FTSE) is 0.601, suggesting a strong PQD property as well. In contrast, the

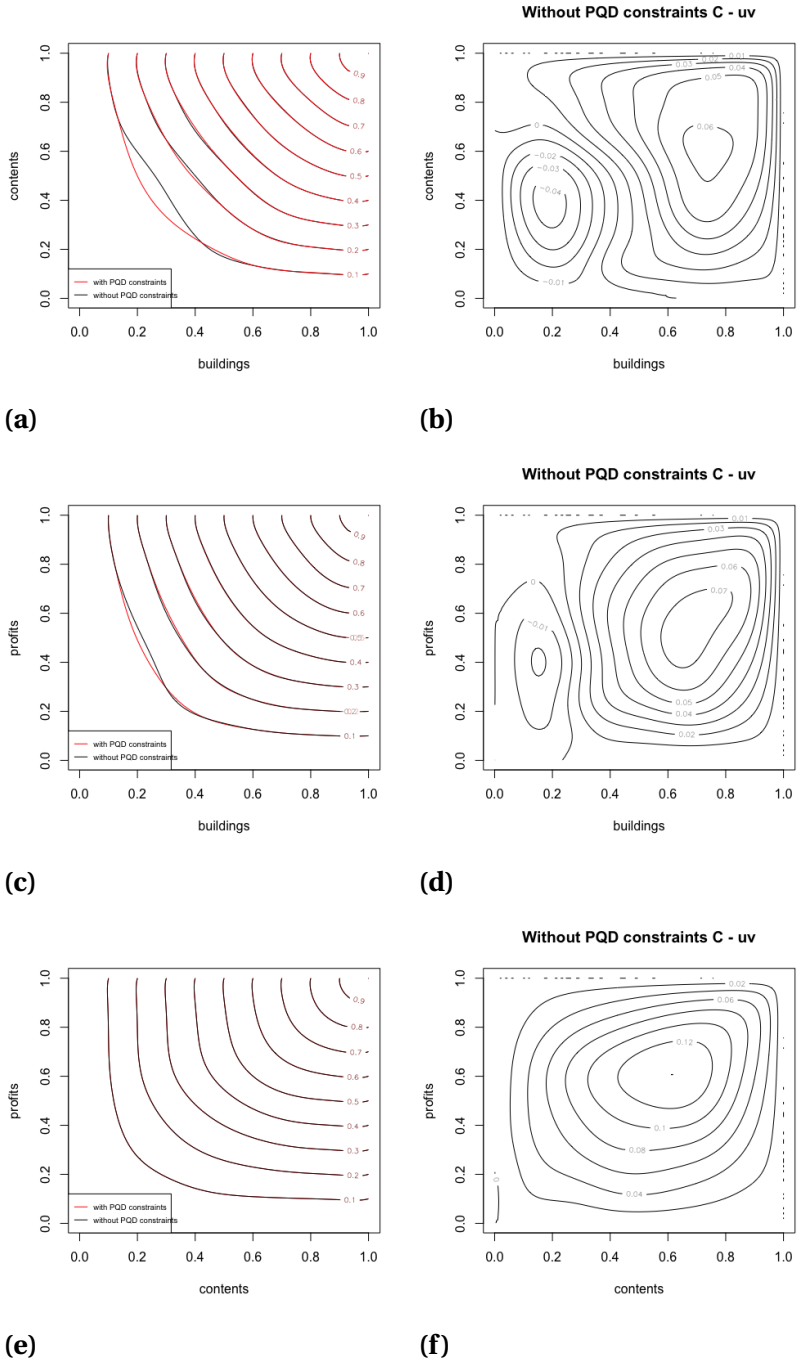
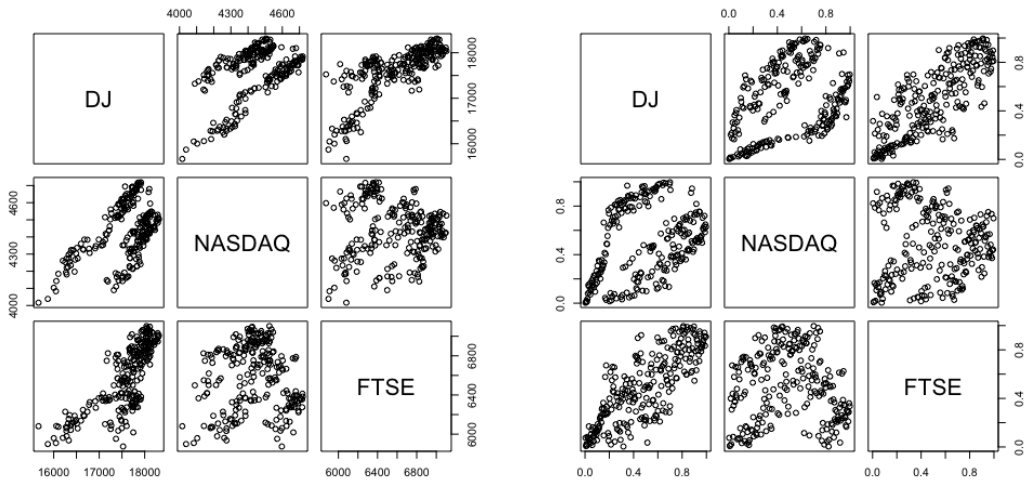


Figure 2.5 Left (a, c, e) : PQD-constrained and unconstrained estimates; Right (b, d, f): unconstrained copula estimate of $\hat{C}_{2,m,n}(u,v) - uv$ for pairwise Danish fire insurance claims.



(a) Observations of the stock index data **(b)** Pseudo-observations

Figure 2.6 Scatterplots of the stock index data (a) and the corresponding pseudo-observations (b).

dependence structure is less clear for (DJ, NASDAQ) with the empirical value of Kendall's tau equal to 0.266. The empirical value of Kendall's tau for (NASDAQ, FTSE) is −0.014, which suggests no positive dependence between NASDAQ 100 and FTSE 100.

P-values for testing PQD between pairwise stock indices are given in Table 2.4. PQD tests with different distance measures give the same conclusion for each pair of stock indices. There is a strong evidence to reject the null hypothesis of PQD for the pair of (NASDAQ, FTSE) since all tests have p-values equal to 0. With a significant level of 0.05, we reject the null hypothesis of PQD for (DJ, NASDAQ), though the evidence is less strong than that for (NASDAQ, FTSE). Whereas, there is no evidence to reject PQD for (DJ, FTSE), and this supports what we see from the plot of pseudo-observations.

Table 2.4 P-values for proposed PQD tests on pairwise stock index data.

Test Statistic	(DJ, FTSE)	(DJ, NASDAQ)	(NASDAQ, FTSE)
KS	1	0.034	0
CvM	1	0.035	0
AD	1	0.028	0

2.7 Concluding Remarks

In this chapter, we have expanded on the Bernstein copula by proposing a new nonparametric estimation method that minimizes an Anderson-Darling type distance between the empirical copula and Bernstein copula to approximate any continuous copula, especially for PQD copula. The proposed copula estimate remains a genuine copula for any degrees m_1 and m_2 of the Bernstein copula and for any sample size n. In addition, we have developed an algorithm that adaptively uses data to select the degrees of the Bernstein copula guided by the asymptotic theory, an aspect that has received very little attention in the literature. The codes written in R can be requested from the first author. Based on the nonparametric estimation of PQD-constrained and unconstrained copula, a test for the null hypothesis of PQD is constructed. As investigated in Denuit & Scaillet [2004] , it is straightforward to extend such estimation and testing methods to higher dimensions, which leads to the estimator and test for positive orthant dependent(POD) copula. A d-dimensional copula C is said to be positive lower orthant dependent (PLOD) if

$$C(u_1,...,u_d) \geq \prod_{i=1}^{d} u_i, \ \ \forall (u_1,...,u_d) \in [0,1]^d, \tag{2.30}$$

while it is said to be positive upper orthant dependent (PUOD) if

$$\bar{C}(u_1,...,u_d) \geq \prod_{i=1}^{d}(1-u_i),\ \ \forall (u_1,...,u_d) \in [0,1]^d. \tag{2.31}$$

where $\bar{C}$ denotes the survival copula associated with C, see Nelsen [2007]. If both (2.30) and (2.31) hold, then C is said to be POD.

The extension of estimation and testing for PLOD copula is immediate. In order to estimate and test for PUOD copula, we can model the survival copula $\bar{C}$ as

$$\bar{C}(u_1,...,u_d) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} \theta_{k_1,...,k_d} \prod_{i=1}^{d} \binom{m_i}{k_i} (1-u_i)^{k_i} u_i^{m_i-k_i}. \tag{2.32}$$

and the results of estimation and testing for PQD remain valid for PUOD.

However, the data-based selection of 'optimal' degrees becomes much more difficult when dimension $d \geq 3$, which remains to be solved and will require future work. Besides the PQD property, there are many other kinds of assumptions, such as symmetry, stochastically increasing (SI), positive function dependence (PFD), etc. Extension of the PQD to other properties is certainly possible if such property is expressed directly in terms of copula function or its derivatives, as such property can be achieved by corresponding constraints on θ's or ω's as described in this chapter. The constrained copula estimation under such conditions is still an open research area.

Finally, in this chapter, we have established the pointwise consistency of our proposed copula estimator (see Theorem 3) and the asymptotic normality of it (see Theorem 4). It would certainly be of interest to derive the asymptotic biases and uniform strong consistency of our copula estimator (with or without PQD), using similar methods as in Janssen et al. [2012] which has been used for the empirical Bernstein copula. However, such asymptotic analysis could be a topic for future contribution as it would require more in-depth analysis with possibly additional regularity conditions.

2.8 Technical Details

2.8.1 Proof of Theorem 2.1

Proof. Since $\theta_{k_1,0} = \theta_{0,k_2} = 0$, then we have

$$C_m(u,0) = \sum_{k_1=0}^{m_1} \theta_{k_1,0} \binom{m_1}{k_1} u^{k_1}(1-u)^{m_1-k_1} = 0,$$
$$C_m(0,v) = \sum_{k_2=0}^{m_2} \theta_{0,k_2} \binom{m_2}{k_2} v^{k_2}(1-v)^{m_2-k_2} = 0.$$

Since $\theta_{k_1,m_2} = \frac{k_1}{m_1}$ and $\theta_{m_1,k_2} = \frac{k_2}{m_2}$, then

$$C_m(u,1) = \sum_{k_1=0}^{m_1} \frac{k_1}{m_1} \binom{m_1}{k_1} u^{k_1}(1-u)^{m_1-k_1} = u,$$
$$C_m(1,v) = \sum_{k_2=0}^{m_2} \frac{k_2}{m_2} \binom{m_2}{k_2} v^{k_2}(1-v)^{m_2-k_2} = v.$$

The partial derivative of Bernstein polynomial can be expressed in the form

$$\frac{\partial C_m(u,v)}{\partial u} = m_1 \sum_{k_1=0}^{m_1-1} \sum_{k_2=0}^{m_2} (\theta_{k_1+1,k_2} - \theta_{k_1,k_2}) \binom{m_1-1}{k_1} u^{k_1}(1-u)^{m_1-k_1-1} \binom{m_2}{k_2} v^{k_2}(1-v)^{m_2-k_2},$$

then

$$\frac{\partial^2 C_m(u,v)}{\partial u \partial v} = m_1 m_2 \sum_{k_1=0}^{m_1-1} \sum_{k_2=0}^{m_2-1} (\theta_{k_1+1,k_2+1} - \theta_{k_1,k_2+1} - \theta_{k_1+1,k_2} + \theta_{k_1,k_2}) \binom{m_1-1}{k_1} u^{k_1}(1-u)^{m_1-k_1-1} \binom{m_2-1}{k_2} v^{k_2}(1-v)^{m_2-k_2-1}.$$

Since $\theta_{k_1+1,k_2+1}-\theta_{k_1,k_2+1}-\theta_{k_1+1,k_2}+\theta_{k_1,k_2}\geq 0$, then $\frac{\partial^2 C_m(u,v)}{\partial u \partial v}\geq 0$, thus C_m is a copula. Further, if $\theta_{k_1,k_2}\geq \frac{k_1}{m_1}\frac{k_2}{m_2}$, then

$$C_m(u,v)\geq \sum_{k_1=0}^{m_1}\sum_{k_2=0}^{m_2}\frac{k_1}{m_1}\frac{k_2}{m_2}\binom{m_1}{k_1}u^{k_1}(1-u)^{m_1-k_1}\binom{m_2}{k_2}v^{k_2}(1-v)^{m_2-k_2}=uv,$$

thus C_m is a PQD copula. □

2.8.2 Proof of Theorem 2.2

Proof. Let $\tilde{\theta}_{k_1,k_2}=C(\frac{k_1}{m_1},\frac{k_2}{m_2})$. Following the proof of Lemma 3.2 in Segers et al. [2017], we have

$$\sup_{u,v\in[0,1]}|C(u,v)-C_m(u,v)|\leq\Big(\frac{1}{m_1\wedge m_2}\Big)^{\frac{1}{2}}\to 0 \text{ as } m_1\wedge m_2\to\infty.$$

□

2.8.3 Proof of Lemma 2.1

Proof. Since we assume that $\inf_{||\theta_m-\hat{\theta}_m||>\epsilon} AD(\theta_m)>AD(\hat{\theta}_m)$, then for any $\epsilon>0$ and $||\theta_m-\hat{\theta}_m||>\epsilon$, there exists $\delta>0$ s.t. $AD(\theta_m)-AD(\hat{\theta}_m)>\delta$. Then,

$$\begin{aligned}
\mathbb{P}(||\hat{\theta}_{m,n}-\hat{\theta}_m||>\epsilon)&\leq\mathbb{P}(AD(\hat{\theta}_{m,n})-AD(\hat{\theta}_m)>\delta)\\
&=\mathbb{P}(AD(\hat{\theta}_{m,n})-AD_n(\hat{\theta}_{m,n})+AD_n(\hat{\theta}_{m,n})-AD(\hat{\theta}_m)>\delta)\\
&\leq\mathbb{P}(AD(\hat{\theta}_{m,n})-AD_n(\hat{\theta}_{m,n})+AD_n(\hat{\theta}_m)-AD(\hat{\theta}_m)>\delta)\\
&\leq\mathbb{P}(|AD(\hat{\theta}_{m,n})-AD_n(\hat{\theta}_{m,n})+AD_n(\hat{\theta}_m)-AD(\hat{\theta}_m)|>\delta)\\
&\leq\mathbb{P}(|AD(\hat{\theta}_{m,n})-AD_n(\hat{\theta}_{m,n})|+|AD_n(\hat{\theta}_m)-AD(\hat{\theta}_m)|>\delta)
\end{aligned}$$

Since $AD_n(\theta_m) \xrightarrow{P} AD(\theta_m)$ as $n \to \infty$, then we have $|AD(\hat{\theta}_{m,n}) - AD_n(\hat{\theta}_{m,n})| \xrightarrow{P} 0$, $|AD_n(\hat{\theta}_m) - AD(\hat{\theta}_m)| \xrightarrow{P} 0$ and $|AD(\hat{\theta}_{m,n}) - AD_n(\hat{\theta}_{m,n})| + |AD_n(\hat{\theta}_m) - AD(\hat{\theta}_m)| \xrightarrow{P} 0$. Thus,

$$\begin{aligned} 0 &\leq \lim_{n\to\infty} \mathbb{P}(||\hat{\theta}_{m,n} - \hat{\theta}_m|| > \epsilon) \\ &\leq \lim_{n\to\infty} \mathbb{P}(|AD(\hat{\theta}_{m,n}) - AD_n(\hat{\theta}_{m,n})| + |AD_n(\hat{\theta}_m) - AD(\hat{\theta}_m)| > \delta) \\ &= 0 \end{aligned}$$

i.e., $\hat{\theta}_{m,n} \xrightarrow{P} \hat{\theta}_m$ as $n \to \infty$. □

2.8.4 Proof of Theorem 2.3

Proof. Let $\tilde{\theta}_{k_1,k_2} = C(\frac{k_1}{m_1}, \frac{k_2}{m_2})$ and $\tilde{\theta}_m = (\tilde{\theta}_{0,0}, ..., \tilde{\theta}_{m_1,m_2})$. From Theorem 2 we know that

$$\sup_{u,v\in[0,1]} |C(u,v) - C_m(u,v,\tilde{\theta}_m)| \leq \Big(\frac{1}{m_1 \wedge m_2}\Big)^{\frac{1}{2}} \to 0 \text{ as } m_1 \wedge m_2 \to \infty.$$

Let $\tilde{m} = \min(m_1, m_2)$. Since $\xi_m = O(1)$, then we have

$$\begin{aligned} AD(\hat{\theta}_m) \leq AD(\tilde{\theta}_m) &= \int_{I^2} \frac{(C(u,v) - C_m(u,v,\tilde{\theta}_m))^2}{\xi_m} dC(u,v) \\ &\leq \sup_{u,v\in[0,1]} |C(u,v) - C_m(u,v,\tilde{\theta}_m)|^2 \int_{I^2} \frac{1}{\xi_m} dC(u,v) \\ &\leq \frac{1}{\xi_m}\Big(\frac{1}{4m_1} + \frac{1}{4m_2} + \Big(\frac{1}{4m_1 m_2}\Big)^{\frac{1}{2}}\Big) \\ &\leq \frac{1}{\xi_m \tilde{m}} \to 0 \text{ as } \tilde{m} \to \infty \end{aligned}$$

Note that $AD(\hat{\theta}_m) = \frac{1}{\xi_m} E_C(C(U,V) - C_m(U,V,\hat{\theta}_m))^2 \to 0$ and $\xi_m = O(1)$. Since C and C_m are continuous functions, then we have $|C_m(u,v,\hat{\theta}_m) - C(u,v)| \to 0$ as $\tilde{m} \to \infty$.

Finally we show that for any $u, v \in [0,1]$,

$$\begin{aligned}|C_m(u,v,\hat{\theta}_{m,n}) - C(u,v)| &\le |C_m(u,v,\hat{\theta}_{m,n}) - C_m(u,v,\hat{\theta}_m)| \\ &\quad + |C_m(u,v,\hat{\theta}_m) - C(u,v)| \\ &\xrightarrow{\text{p}} 0 \text{ as } n, \tilde{m} \to \infty\end{aligned}$$

□

2.8.5 Proof of Theorem 2.4

Proof. For simplicity, we denote the unconstrained copula estimation $\hat{C}_{2,m,n}$ as $\hat{C}_{m,n}$. Following the proof of Theorem 2 in Janssen et al. [2012], we first show the asymptotic normality of $\sqrt{n}(\hat{C}_{m,n} - C)$. Let $\hat{C}_m \equiv C_m(u,v,\hat{\theta}_m)$, where $\hat{\theta}_m$ is defined in (2.10). The following decomposition is used to show the asymptotic distribution of $\hat{C}_{m,n}$:

$$\sqrt{n}(\hat{C}_{m,n} - C) = \sqrt{n}(\hat{C}_{m,n} - \hat{C}_m) + \sqrt{n}(\hat{C}_m - C). \tag{2.33}$$

Let $\tilde{C}_m \equiv C_m(u,v,\tilde{\theta}_m)$, where $\tilde{\theta}_m = (\tilde{\theta}_{0,0}, ..., \tilde{\theta}_{m_1,m_2})$ and $\tilde{\theta}_{k_1,k_2} = C(\frac{k_1}{m_1}, \frac{k_2}{m_2})$. Assuming the third order partial derivatives for C are bounded on $(0,1)^2$, the convergence rate of the Bernstein representation $\tilde{C}_m$ is given as (see Janssen et al. [2012])

$$\begin{aligned}\tilde{C}_m - C &= \sum_{k_1=0}^{m_1}\sum_{k_2=0}^{m_2}\Big[\frac{1}{2}\Big(\frac{k_1}{m_1} - u\Big)^2 C_{uu}(u,v) + \frac{1}{2}\Big(\frac{k_2}{m_2} - v\Big)^2 C_{vv}(u,v) \\ &\quad + \Big(\frac{k_1}{m_1} - u\Big)\Big(\frac{k_2}{m_2} - v\Big)C_{uv}(u,v)\Big]P_{k_1,m_1}(u)P_{k_2,m_2}(v) + o\Big(\frac{1}{\tilde{m}}\Big) \\ &= \frac{1}{2m_1}u(1-u)C_{uu}(u,v) + \frac{1}{2m_2}v(1-v)C_{vv}(u,v) + o\Big(\frac{1}{\tilde{m}}\Big)\end{aligned}$$

where $\tilde{m} = \min(m_1, m_2)$, $P_{k_1,m_1}(u) = \binom{m_1}{k_1}u^{k_1}(1-u)^{m_1-k_1}$ and $P_{k_2,m_2}(v) = v^{k_2}(1-v)^{m_2-k_2}$. Then

we have $\tilde{C}_m - C = O(\tilde{m}^{-1})$. From the definition of $\hat{\theta}_m$, we can obtain that

$$\begin{aligned} AD(\hat{\theta}_m) \leq AD(\tilde{\theta}_m) &= \int_{I^2} \frac{(C-\tilde{C}_m)^2}{\xi_m} dC \\ &\leq \sup_{u,v\in[0,1]} (C-\tilde{C}_m)^2 \int_{I^2} \frac{1}{\xi_m} dC \\ &= \frac{1}{\xi_m} \sup_{u,v\in[0,1]} (C-\tilde{C}_m)^2. \end{aligned}$$

Since $\xi_m = O(1)$, then we have

$$\tilde{m}^2 AD(\hat{\theta}_m) = \frac{1}{\xi_m} \tilde{m}^2 E_c (C-\hat{C}_m)^2 \leq \frac{1}{\xi_m} \tilde{m}^2 \sup_{u,v\in[0,1]} (C-\tilde{C}_m)^2 = O(1).$$

Thus, the order of $\hat{C}_m - C$ cannot be higher than $\tilde{m}^{-1}$. Under the condition of $n^{\frac{1}{2}}\tilde{m}^{-1} \to 0$, it is easy to show the convergence of the second term in (2.33)

$$\sqrt{n}(\hat{C}_m - C) = \frac{\sqrt{n}}{\tilde{m}} \tilde{m}(\hat{C}_m - C) \to 0.$$

Next we show the asymptotic normality of first term in (2.33). From the weak convergence of the empirical copula process C_n we know that $C_n = C + O(\frac{1}{\sqrt{n}})$. Let

$$\begin{aligned} y_i &\equiv \frac{C_n(U_i,V_i)}{\sqrt{\frac{n}{(n+1)^2} C_n(U_i,V_i)(1-C_n(U_i,V_i)) + \xi_{m,n}}} \\ x_i &\equiv \frac{\frac{n}{n+1} B_m(U_i,V_i)}{\sqrt{\frac{n}{(n+1)^2} C_n(U_i,V_i)(1-C_n(U_i,V_i)) + \xi_{m,n}}}, \end{aligned}$$

then we have $y_i = \frac{C(U_i,V_i)}{\sqrt{\xi_m}} + O(\frac{1}{\sqrt{n}})$ and $x_i = \frac{B_m(U_i,V_i)}{\sqrt{\xi_m}} + o(1)$. The Anderson-Darling type distance can be rewritten as

$$AD_n(\theta_m) = \frac{1}{n}\sum_{i=1}^{n} (y_i - x_i^T\theta_m)^2 = (y - X\theta_m)^T (y - X\theta_m),$$

where

$$y = (y_1, ..., y_n)^T = \Big(\frac{C(U_1, V_1)}{\sqrt{\xi_m}} + O\Big(\frac{1}{\sqrt{n}}\Big), ..., \frac{C(U_n, V_n)}{\sqrt{\xi_m}} + O\Big(\frac{1}{\sqrt{n}}\Big)\Big),$$
$$X = (x_1, ..., x_n)^T = \Big(\frac{B_m(U_i, V_i)}{\sqrt{\xi_m}} + o(1), ..., \frac{B_m(U_i, V_i)}{\sqrt{\xi_m}} + o(1)\Big)^T.$$

Then the estimation of coefficients $\hat{\theta}_{m,n}$ that minimizes Anderson-Darling type distance takes the form

$$\hat{\theta}_{m,n} = \underset{\theta_m \in \Theta}{\operatorname{argmin}} (y - X\theta_m)^T (y - X\theta_m) \text{ subject to } A^T \theta_m \geq b_0.$$

Based on the Theorem 1 in Liew [1976], the inequality constrained least squares (ICLS) estimator $\hat{\theta}_{m,n}$ reduces to the OLS estimator $\hat{\theta}^* = (X^T X)^{-1} X^T y$ for a sufficient large sample size $n \geq n_0$ if none of the ICLS estimator is bounded, i.e., $A^T \theta_m \gg b_0$. Then we have

$$\hat{\theta}_{m,n} = (X^T X)^{-1} X^T y = (\frac{1}{n} X^T X)^{-1} \frac{1}{n} X^T y.$$

Notice that

$$\frac{1}{n} X^T X = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T \to \frac{1}{\xi_m} E(B_m(U, V) B_m(U, V)^T) \text{ as } n \to \infty,$$

and

$$\begin{aligned} \frac{1}{n} X^T y = \frac{1}{n} \sum_{i=1}^{n} y_i x_i &= \frac{1}{n} \sum_{i=1}^{n} \Big(\frac{C(U_i, V_i)}{\sqrt{\xi_m}} + O\Big(\frac{1}{\sqrt{n}}\Big)\Big)\Big(\frac{B_m(U_i, V_i)}{\sqrt{\xi_m}} + o(1)\Big) \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{C(U_i, V_i) B_m(U_i, V_i)}{\xi_m} + o_p\Big(\frac{1}{\sqrt{n}}\Big). \end{aligned}$$

By CLT we have

$$\sqrt{n}\Big(\frac{1}{n} \sum_{i=1}^{n} \frac{C(U_i, V_i) B_m(U_i, V_i)}{\xi_m} \Big) \xrightarrow{d} N(\mu, V) \text{ as } n \to \infty,$$

where $\mu = E(\frac{1}{\xi_m} B_m(U, V) C(U, V)) = E(\frac{1}{\xi_m} B_m(U, V) B_m(U, V)^T \hat{\theta}_m) =$
$\frac{1}{\xi_m} E(B_m(U, V) B_m(U, V)^T) \hat{\theta}_m$ and $V = Var(\frac{1}{\xi_m} B_m(U, V) C(U, V)) =$

$\frac{1}{\xi_m} Var(B_m(U,V)B_m(U,V)^T\hat{\theta}_m)$. Thus, by using Slutsky's theorem, we obtain the asymptotic normality of $\sqrt{n}(\hat{\theta}_{m,n}-\hat{\theta}_m)$. Since we have $\hat{C}_{m,n}-C=(\hat{\theta}_{m,n}-\hat{\theta}_m)^T B_m$, then the asymptotic normality of the first term in (2.33) is also obtained.

Again, by Slutsky's theorem, the limiting distribution of $\sqrt{n}(\hat{C}_{m,n}-C)$ is a normal distribution with zero mean. Such asymptotic normality result can be strengthened to weak convergence of $\sqrt{n}(\hat{C}_{m,n}-C)$ in $(0,1)^2$(see Janssen et al. [2012]). Thus, we can follow the proof of Theorem 1 in Gijbels et al. [2010] with the empirical copula estimator $\hat{C}_n$ replaced by our copula estimator $\hat{C}_{m,n}$ to obtain the consistency of the proposed testing procedure. For more details, see Gijbels et al. [2010]. □

CHAPTER

3

NONPARAMETRIC ESTIMATION OF MULTIVARIATE COPULA USING EMPIRICAL BAYES METHODS

3.1 Introduction

Copula models are useful tools for the analysis of multivariate data since by using the well-known Sklar's theorem, any multivariate joint distribution can be decomposed into its univariate marginal distributions and a copula function, which allows for capturing the arbitrary dependence structure between several random variables. As a result, copulas have been widely used in the field of finance, insurance, system reliability, etc. among many

other application areas. See e.g. Nelsen [2007], Jaworski et al. [2010] and Joe [2014] for more details about copulas and their applications.

Given a random vector $(X_1, \ldots, X_d)$ with joint cumulative distribution function (CDF) F and continuous marginal CDFs F_j, $j = 1, \ldots, d$, by Sklar's theorem (Sklar [1959]), the CDF F can be expressed uniquely as $F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$, where $C(\cdot)$ denotes the copula function. A copula is itself the joint CDF of a random vector $(U_1 = F_1(X_1), \ldots, U_d = F_d(X_d))$ having its marginals as uniform distributions on $[0,1]$, henceforth denoted by $Unif[0,1]$. It is to be noted that the original results in Sklar [1959] are also applicable to discrete-valued random variables. However, the focus of this paper is modeling continuous-valued multivariate random vectors (e.g., daily return values of stock prices). Thus, for the rest of the paper, we assume the marginal CDFs F_j, $j = 1, \ldots, d$ are absolutely continuous.

As copula plays an important role in capturing the general dependence structure between multiple variables, it is critical to estimate copula in an accurate way, especially in higher dimensions where the dependence structure becomes much more complicated, often illusive and may even be supported on a lower-dimensional manifold. One of theThe primary goals of this paper is to estimate a smooth copula function C from a random sample of n independent identically distributed (iid) observations $(X_{i1}, \ldots, X_{id}) \overset{iid}{\sim} F(x_1, \ldots, x_d)$ for $i = 1, 2, \ldots, n$.

Many parametric families have been proposed for modeling multivariate copulas and there has been previous work addressing the corresponding parametric estimation methods. For detailed discussions, see, e.g. Joe [2005], Nelsen [2007], McNeil & Nešlehová [2009], Žežula [2009], Smith [2011] and Joe [2014], etc. However, no matter how sophisticated and flexible the parametric models that we may use, they might still lead to biased copula estimates when the parametric model is misspecified and thus may not be able to capture complex dependence structures required in practice. Compared to standard multivariate copulas, vine copula models allow for more flexibility in capturing complex dependency structures using appropriate vine tree structures by choosing bivariate copula families

for each node of pair copulas from a vast array of parametric bivariate copulas. But it is often challenging to obtain estimates of multivariate dependence measures which involve high-dimensional integrals that are often algebraically intractable by using vine copulas.

Thus, recognizing some of the above-mentioned limitations of parametric copula models, a variety of nonparametric estimators have been proposed for multivariate copula estimation. Most of the available nonparametric estimators rely on the empirical methods, e.g., the empirical copula and its multilinear extension, the empirical multilinear copula (Deheuvels [1979]; Fermanian et al. [2004]; Genest et al. [2017]), or kernel-based methods such as local linear estimator (Chen & Huang [2007]), mirror reflection estimator (Gijbels & Mielniczuk [1990]) and improvements of these two estimators (Omelka et al. [2009]). See also Scaillet & Fermanian [2002] and Rémillard & Scaillet [2009] for other nonparametric copula estimators. However, except for the empirical multilinear copula, most of these estimators are valid copulas only asymptotically, meaning that they are not necessarily genuine copulas for finite samples. Moreover, multivariate dependent measures (e.g. Spearman's rho, Kendall's tau, etc.) based on such estimated copula could take values outside of their natural range, thus making them unattractive in practice. On the other hand, there has been recent work on Bayesian nonparametric methods for estimating general d-dimensional copula and among many others, a noteworthy Bayesian nonparametric model is based on an infinite mixture of multivariate Gaussian or the skew-normal copulas proposed by [Wu et al., 2014]. The infinite mixture models provide a lot of flexibility in modeling various dependence structures but those typically lack simple (analytic) expressions of dependence measures making them harder to compute in practice.

The primary focus of this chapter is the nonparametric estimation of multivariate copulas for any arbitrary dimensions that are genuine copulas for any finite sample size and are uniformly consistent as the sample size gets large. We consider an extension of the Bernstein copula (Sancetta & Satchell [2004]), which is a family of copulas defined in terms of multivariate Bernstein polynomials. One of the primary advantages of the

Bernstein copula is that it provides a class of nested models that are able to uniformly approximate any multivariate copula with minimal regularity conditions. A simple case of the Bernstein copula is the empirical Bernstein copula, which is a nonparametric copula estimator proposed by Sancetta & Satchell [2004]. The asymptotic properties of the empirical Bernstein copula are well studied in Janssen et al. [2012] and its application in testing independence is described in Belalia et al. [2017]. The application of the Bernstein copula to the modeling of dependence structures of non-life insurance risks is provided in Diers et al. [2012], among many other applications.

However, the empirical Bernstein copula has two main drawbacks that could prevent us from obtaining accurate copula estimation for small samples: (i) the empirical Bernstein copula is not necessarily a valid copula itself, which is a common disadvantage for most nonparametric copula estimators; and (ii) the degrees of Bernstein polynomials are often set to be equal to an integer across different dimensions, which limits the flexibility of the Bernstein copula and thus might not be appropriate for large dimensions.

In order to address the above-described problem (i), Segers et al. [2017] show that the empirical Bernstein copula is a genuine copula if and only if all the polynomial degrees are divisors of the sample size, and further propose a new copula estimator called the empirical beta copula, which can be seen as a special case of the empirical Bernstein copula when the degrees of Bernstein polynomials are all set equal to the sample size. The empirical beta copula is a valid copula itself and has been shown to outperform some classical copula estimators in terms of bias and variance. But it always has a larger variance compared to the empirical Bernstein copula with smaller polynomial degrees. It is surprising that much less attention has been given to the problem (ii), and even for equally set degrees, there has been limited work on the data-dependent choice of degrees in the literature. Janssen et al. [2012] recommend an optimal choice of the equal degrees in the bivariate case by minimizing the asymptotic mean squared error. Nevertheless, such a choice requires the knowledge of the first- and second-order partial derivatives, which might not

be easy to estimate in practice. Burda & Prokhorov [2014] put priors on the polynomial degrees, however, their priors don't rely on data or sample size and they use multivariate Bernstein density instead of Bernstein copula density. The Dirichlet process assigned as the prior for the copula doesn't guarantee the copula estimate to be a valid copula itself. Besides, the number of weights grows exponentially as the dimension increases, leading to computational inefficiency of MCMC methods for larger dimensions. To the best of our knowledge, Lu & Ghosh [2020] first develop a data-dependent grid search algorithm for the selection of polynomial degrees which has shown superior empirical estimation properties for small to moderate sized samples. But the methodology is limited to bivariate cases and extension to larger dimensions remained challenging.

For the purpose of addressing the two problems as described above, we introduce a new nonparametric estimator for multivariate copula that we call the empirical checkerboard Bernstein copula (ECBC), which is constructed by extending the Bernstein copula allowing for varying degrees of the polynomials. It is shown to be a genuine copula for any number of degrees and any finite sample size. Furthermore, we develop an empirical Bayesian method that takes the data into account to automatically choose the degrees of the proposed estimator using its posterior distribution thereby accounting for the uncertainty of such tuning parameter selection. As shown in Segers et al. [2017], larger degrees of the Bernstein copula lead to a larger variance of the estimation, so a choice of degrees that is relatively small compared to the sample size but sufficient for a good copula estimation is desirable. The degrees are allowed to be dimension-varying within the Bayesian model, which provides much more flexibility and accuracy, especially in higher dimensions.

It is especially noteworthy that while the focus of the chapter is to estimate the copula function, it is straightforward to get a closed-form estimate of the corresponding copula density by taking derivatives of the ECBC. However, direct estimation of a closed-form copula function has many advantages compared to first estimating a copula density, e.g., it is often easier to differentiate than to integrate for higher dimensions. Besides, for those

copulas which are not absolutely continuous such as Marshall-Olkin copulas (Embrechts et al. [2001]) having support on a possibly lower-dimensional manifold, the direct estimation of the copula density could be difficult. Owing to the closed form of the estimated copula function and its density, it can be shown that the proposed ECBC allows for straightforward estimation of various dependence measures.

3.2 An Empirical Bayes Nonparametric Copula Model

Suppose we have i.i.d. samples $(X_{i1},\ldots,X_{id}) \sim F(x_1,\ldots,x_d), i = 1,\ldots,n$, where F is a cumulative distribution function and F_j is the absolutely continuous marginal CDF of the j-th component. By Sklar 's theorem (Sklar [1959]), there exists a unique copula $C(\cdot)$ such that

$$F(x_1,\ldots,x_d) = C(F_1(x_1),\ldots,F_d(x_d)),\ \ \forall\,(x_1,\ldots,x_d) \in \mathbb{R}^d, \tag{3.1}$$

and

$$(F_1(X_1),\ldots,F_d(X_d)) \sim C. \tag{3.2}$$

The Bernstein copula is a family of copulas defined in terms of Bernstein polynomials and it was first introduced by Sancetta & Satchell [2004]. It is a flexible model that can be used to uniformly approximate any copula. The Bernstein polynomial with degrees $(m_1,\ldots,m_d)$ of a function $C : [0,1]^d \to \mathbb{R}$ is defined as

$$B_m(C)(\mathbf{u}) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} C\left(\frac{k_1}{m_1},\ldots,\frac{k_d}{m_d}\right) \prod_{j=1}^{d} \binom{m_j}{k_j} u_j^{k_j} (1-u_j)^{m_j-k_j}, \tag{3.3}$$

and $B_m(C)$ is called the Bernstein copula when C is a copula.

A general estimation for the Bernstein copula of an unknown copula C is the empirical Bernstein copula (Sancetta & Satchell [2004]) $B_m(C_n)$, where C_n is the rank-based empirical

copula. We denote the empirical Bernstein copula as

$$C_{m,n}(\mathbf{u}) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} \hat{\theta}_{k_1,\ldots,k_d} \prod_{j=1}^{d} \binom{m_j}{k_j} u_j^{k_j} (1-u_j)^{m_j-k_j}, \tag{3.4}$$

where

$$\hat{\theta}_{k_1,\ldots,k_d} = C_n\left(\frac{k_1}{m_1}, \ldots, \frac{k_d}{m_d}\right) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} \mathbb{I}\left(F_{nj}^*(X_{ij}) \leq \frac{k_j}{m_j}\right). \tag{3.5}$$

where $\mathbb{I}(\cdot)$ denotes indicator function and slightly modified empirical marginal distribution functions are defined as

$$F_{nj}^*(x_j) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbb{I}(X_{ij} \leq x_j), \text{ for } j = 1, \ldots, d, \tag{3.6}$$

where the modification $1/(n+1)$ instead of $1/n$ modifies the standard empirical marginal distribution to be away from 1 in order to reduce potential problems at boundaries.

However, the empirical Bernstein copula $C_{m,n}$ is not guaranteed to be a valid copula for finite samples as the empirical copula C_n is not necessarily a genuine copula. Segers et al. [2017] show that the empirical Bernstein copula $C_{m,n}$ is a copula if and only if all the degrees $m_1, ..., m_d$ are divisors of n. In order to obtain a valid copula estimation for any degrees, we replace the empirical copula C_n with the empirical checkerboard copula $C_n^{\#}$, which is a simple multilinear extension of the empirical copula defined as

$$C_n^{\#}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} \min(\max((n\,u_j - R_{i,j}^{(n)} + 1), 0), 1), \tag{3.7}$$

where $R_{i,j}^{(n)}$ is the rank of X_{ij} among $X_{1j}, \ldots, X_{nj}$, see e.g. Li et al. [1997] and Carley & Taylor [2002] for more details. Notice that the main difference between the empirical copula C_n and the empirical checkerboard copula $C_n^{\#}$ is that $C_n^{\#}$ is a genuine copula, so we can obtain

a valid copula estimation $C^{\#}_{m,n}$ taking the form

$$C^{\#}_{m,n}(\mathbf{u}) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} \tilde{\theta}_{k_1,\ldots,k_d} \prod_{j=1}^{d} \binom{m_j}{k_j} u_j^{k_j} (1-u_j)^{m_j-k_j}, \tag{3.8}$$

where

$$\tilde{\theta}_{k_1,\ldots,k_d} = C^{\#}_n\left(\frac{k_1}{m_1},\ldots,\frac{k_d}{m_d}\right) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} \min(\max\left(\left(n\frac{k_j}{m_j} - R^{(n)}_{i,j} + 1\right), 0), 1\right), \tag{3.9}$$

and we call the proposed *empirical checkerboard Bernstein copula (ECBC)*.

Unlike the empirical Bernstein copula, the ECBC is a genuine copula for any degrees $m_1, m_2, \ldots, m_d \in \mathbb{Z}^+$ and any fixed sample size n. It is known that Bernstein polynomials with smaller values of degrees m_j's may lead to biased estimates while unnecessary larger degrees of Bernstein polynomials will necessarily lead to larger variances. So it is critical to choose the proper degrees of the ECBC based on a given sample. In order to do that, we develop an empirical Bayes method for choosing 'optimal' degrees $(m_1, m_2, \ldots, m_d)$, where m_j's are allowed to be different for different $j = 1, \ldots, d$ and also depend on the random sample of observations.

As illustrated in Sancetta & Satchell [2004], using partial derivatives of (3.8) with respect to each u_j and rearranging, we can obtain the density corresponding to ECBC as follows:

$$\begin{aligned} c^{\#}_{m,n}(\mathbf{u}) &= \sum_{k_1=0}^{m_1-1} \cdots \sum_{k_d=0}^{m_d-1} \tilde{w}_{k_1,\ldots,k_d} \prod_{j=1}^{d} m_j \binom{m_j-1}{k_j} u_j^{k_j} (1-u_j)^{m_j-k_j-1} \\ &= \sum_{k_1=0}^{m_1-1} \cdots \sum_{k_d=0}^{m_d-1} \tilde{w}_{k_1,\ldots,k_d} \prod_{j=1}^{d} Beta(u_j, k_j+1, m_j-k_j), \end{aligned} \tag{3.10}$$

where

$$\tilde{w}_{k_1,\ldots,k_d} = \sum_{l_1=0}^{1} \cdots \sum_{l_d=0}^{1} (-1)^{d+l_1+\ldots+l_d} C^{\#}_n\left(\frac{k_1+l_1}{m_1},\ldots,\frac{k_d+l_d}{m_d}\right). \tag{3.11}$$

Clearly, the Bernstein copula is a mixture of independent Beta distributions leading to a tensor product form. For notational convenience, let us denote

$$U_{ij} \equiv F^*_{nj}(X_{ij}) \;\; i = 1,\ldots,n, j = 1,\ldots,d, \tag{3.12}$$

Following the work by Gijbels et al. [2010], the pseudo-observations $(U_{i1},\ldots U_{id}), i = 1,\ldots,n$ can be treated as samples from $(F_1(U_{i1}),\ldots,F_d(U_{id})) \sim C$. We then use this approximation to build an empirical Bayesian hierarchical model:

$$U_{ij}|L_{ij} \overset{ind}{\sim} Beta(L_{ij}+1, m_j - L_{ij}), \;\; i = 1,\ldots,n, j = 1,\ldots,d, \tag{3.13}$$

$$L_{ij} = \lfloor m_j V_{ij} \rfloor, \tag{3.14}$$

where $\lfloor a \rfloor$ is the 'floor' function denoting the largest integer not exceeding the value a and

$$(V_{i1},\ldots,V_{id}) \overset{i.i.d}{\sim} C^{\#}_n(\cdot), \;\; i = 1,\ldots,n, \tag{3.15}$$

i.e., $(V_{i1},\ldots,V_{id}), i = 1,\ldots,n$ are samples from the empirical checkerboard copula $C^{\#}_n$. It then follows that

$$\begin{aligned} \tilde{w}_{k_1,\ldots,k_d} &= \Pr(L_{i1} = k_1,\ldots,L_{id} = k_d) \\ &= \Pr\left(\frac{k_1}{m_1} \le V_{i1} < \frac{k_1+1}{m_1},\ldots,\frac{k_d}{m_d} \le V_{id} < \frac{k_d+1}{m_d}\right). \end{aligned} \tag{3.16}$$

Based on the proposition 1 in Genest et al. [2017], V_{ij} can be drawn using the following hierarchical scheme:

$$\begin{aligned} \pi_i &\overset{i.i.d}{\sim} DisUnif\{1,2,\ldots,n\}, \;\; i = 1,\ldots,n, \\ \Lambda_{ij} &\overset{i.i.d}{\sim} Unif(0,1) \;\; i = 1,\ldots,n, j = 1,\ldots,d. \\ V_{ij} &= (1-\Lambda_{ij})F_{n,j}(X_{\pi_i j}-) + \Lambda_{ij}F_{n,j}(X_{\pi_i j}), \end{aligned} \tag{3.17}$$

where $F_{n,j}(x_j) = \frac{1}{n}\sum_{i=1}^{n} I(X_{ij} \le x_j)$ and $DisUnif$ denotes the discrete uniform distribution,

i.e., $\Pr[\pi_i = j] = \frac{1}{n}$ for $j = 1, \ldots, d$. Assuming that there are no ties in the pseudo samples $U_{1j}, \ldots, U_{nj}$ (owing to absolute continuity of marginal distributions or breaking it by random assignment in practice), we can equivalently represent the V_{ij}'s more conveniently as

$$\begin{aligned} V_{ij} &= (1-\Lambda_{ij})\frac{R^{(n)}_{\pi_i,j}-1}{n} + \Lambda_{ij}\frac{R^{(n)}_{\pi_i,j}}{n} \\ &= \frac{R^{(n)}_{\pi_i,j}-1+\Lambda_{ij}}{n}. \end{aligned} \tag{3.18}$$

Next, to account for the uncertainty in the estimation of the degrees m_j's, we propose to introduce a sample-size dependent empirical prior distribution on the degrees $m_1, \ldots, m_d$ and obtain posterior estimates by Markov Chain Monte Carlo (MCMC) methods. This would not only allow for the almost automatic adaptive estimation of the degrees (based on the observed data) but also allow for quantifying the uncertainty of this crucial tuning parameter vector. Notice that the idea of putting priors on the polynomial degrees has also been adopted by Burda & Prokhorov [2014]. However, their priors don't rely on data or sample size and they use multivariate Bernstein density instead of Bernstein copula density, i.e., the weights are belonged to a simplex without any more constraints. A Dirichlet process with a baseline of uniform distribution on $[0,1]^d$ is assigned as the prior for the copula C in (3.3), which doesn't guarantee C to be a valid copula. In order to avoid the construction of priors under constraints, we use the empirical estimates for the coefficients of the Bernstein copula instead of assigning priors to them.

Motivated by the asymptotic theory of the empirical Bernstein estimator, e.g., as in Janssen et al. [2014], we propose the hierarchical shifted Poisson distributions as the prior distribution for $m_1, \ldots m_d$:

$$m_j \mid \alpha_j \overset{ind}{\sim} Poisson(n^{\alpha_j}) + 1, \;\; j = 1, \ldots, d. \tag{3.19}$$

and

$$\alpha_j \overset{i.i.d}{\sim} Unif\Big(\frac{1}{3},\frac{2}{3}\Big),\ \ j=1,\ldots,d. \tag{3.20}$$

The following theorem provides the large-sample consistency of the ECBC using the same set of assumptions as required for the large-sample consistency of the empirical checkerboard copula.

Theorem 3.1. *Given the empirical priors distribution of $m_1,\ldots,m_d$ as in (3.19) and (3.20), and assuming the regularity conditions for the consistency of the empirical checkerboard copula, the proposed ECBC is consistent in the following sense:*

$$E(||C^{\#}_{m,n}-C||)=E\left(\sup_{\mathbf{u}\in[0,1]^d}|C^{\#}_{m,n}(\mathbf{u})-C(\mathbf{u})|\right)\overset{a.s.}{\to}0\ \ as\ \ n\to\infty.$$

where the expectation is taken with respect to the empirical prior distribution.

The proof of Theorem 3.1 is outlined in section 3.6.1. Notice that in the above result the a.s. convergence is with respect to the empirical *marginal* distribution of the data integrating out the conditional empirical distribution of data (given the m_j's) weighted by the empirical prior distribution of the tuning parameters m_j's. This is not the usual notion of posterior consistency but rather the notion can be viewed as using integrated likelihood approach (Berger et al. [1999]) with respect to the empirical marginal distribution obtained by integrating the priors given by equations (3.19) and (3.20).

It is to be noted that the joint posterior distribution of $(m_1,\ldots m_d)$ may not preserve necessarily an exchangeable structure as the above prior. Using the empirical Bayes hierarchical structure of the above-proposed model it can be shown that efficient MCMC methods can be utilized to draw approximate samples from the path of a geometrically ergodic Markov Chain with posterior distribution as its stationary distribution. By generating a sufficiently large number of MCMC samples we can estimate the marginal posterior mode of the discrete-valued parameter m_j's as final estimates. Let $m_{j1},\ldots,m_{jK}$ denote K MCMC samples of m_j, $j=1,\ldots,d$ and for each j, let $\tilde{m}_{j1}<\tilde{m}_{j2}<\cdots<\tilde{m}_{jD_j}$ denote the

distinct values among these MCMC samples. Then the (marginal) posterior mode of m_j is estimated by

$$\hat{m}_j = \underset{m_{jb}, b=1,\ldots,D_j}{\operatorname{argmax}} \sum_{a=1}^{K} \mathbb{I}(m_{ja} = \tilde{m}_{jb}), \;\; j = 1, \ldots, d. \tag{3.21}$$

The final estimate of the smooth copula based on the proposed ECBC is then given by

$$C^{\#}_{\hat{m},n}(\mathbf{u}) = \sum_{k_1=0}^{\hat{m}_1} \cdots \sum_{k_d=0}^{\hat{m}_d} \tilde{\theta}_{k_1,\ldots,k_d} \prod_{j=1}^{d} \binom{\hat{m}_j}{k_j} u_j^{k_j} (1-u_j)^{\hat{m}_j - k_j} \tag{3.22}$$

where

$$\tilde{\theta}_{k_1,\ldots,k_d} = C^{\#}_n\left(\frac{k_1}{\hat{m}_1}, \ldots, \frac{k_d}{\hat{m}_d}\right). \tag{3.23}$$

It is to be noted that other posterior estimates (e.g., posterior mean when it exists or coordinate-wise posterior median or some version of multivariate posterior median) can also be used but for simplicity (and the requirement that these posterior estimates of m_j's be necessarily integer-valued) we have chosen to use posterior mode based on the marginal posterior distributions of m_j's. Through many numerical illustrations we show the easy applicability of this choice in various examples.

3.2.1 Multivariate Dependence Estimation

In higher dimensions, it is often of interest to evaluate the strength of dependence among variables. This is often done using copulas since most dependence measures can be expressed as a function of copulas. Spearman's rank correlation coefficient (Spearman's rho) is one of the most widely used dependence measures. For a bivariate copula C, Spearman's rho can be written as

$$\rho = 12 \int_0^1 \int_0^1 C(u,v)\,du\,dv - 3 = 12 \int_0^1 \int_0^1 (C(u,v) - uv)\,du\,dv. \tag{3.24}$$

A multivariate extension of Spearman's rho given in Nelsen [1996] takes the form

$$\rho_d = \frac{\int_{I^d} C(\mathbf{u})d\mathbf{u} - \int_{I^d} \Pi(\mathbf{u})d\mathbf{u}}{\int_{I^d} M(\mathbf{u})d\mathbf{u} - \int_{I^d} \Pi(\mathbf{u})d\mathbf{u}} = \frac{d+1}{2^d-(d+1)}\Big(2^d \int_{I^d} C(\mathbf{u})d\mathbf{u} - 1\Big). \tag{3.25}$$

Compared to vine copulas that rely on pair copulas and complex tree structures, one of the advantages of our copula estimator is that it is straightforward to obtain an estimate of multivariate Spearman's rho as

$$\hat{\rho}_d = \frac{d+1}{2^d-(d+1)}\left(2^d \sum_{k_1=0}^{\hat{m}_1} \cdots \sum_{k_d=0}^{\hat{m}_d} \tilde{\theta}_{k_1,\ldots,k_d} \prod_{j=1}^{d} \binom{\hat{m}_j}{k_j} B(k_j+1, \hat{m}_j - k_j + 1) - 1\right). \tag{3.26}$$

where B is the beta function.

It can be shown that the multivariate Spearman's rho is bounded by

$$\frac{2^d-(d+1)!}{d!(2^d-d-1)} \leq \rho_d \leq 1, \tag{3.27}$$

where the lower bound approaches to zero as dimension increases. Since our copula estimator is a genuine copula, the estimate of multivariate d-dimensional Spearman's rho $\hat{\rho}_d$ can avoid taking values out of the parameter space, which might be an issue for estimates built on other nonparametric copula estimators, e.g. the empirical copula (see [Pérez & Prieto-Alaiz, 2016]).

Similar to Spearman's rho, Kendall's tau is another common dependence measure and has its multivariate version as well, which is given by Nelsen [1996] as

$$\tau_d = \frac{1}{2^{d-1}-1}\left(2^d \int_{I^d} C(\mathbf{u})dC(\mathbf{u}) - 1\right). \tag{3.28}$$

By applying (3.10) and (3.11), it is also easy to obtain an estimate of multivariate Kendall's

tau based on our copula estimator as

$$\hat{\tau}_d = \frac{1}{2^{d-1}-1}\Big(2^d \sum_{k_1=0}^{\hat{m}_1-1} \cdots \sum_{k_d=0}^{\hat{m}_d-1} \sum_{l_1=0}^{\hat{m}_1} \cdots \sum_{l_d=0}^{\hat{m}_d} \tilde{w}_{k_1,\ldots,k_d} \tilde{\theta}_{l_1,\ldots,l_d} \prod_{j=1}^{d} \hat{m}_j \binom{\hat{m}_j - 1}{k_j} \binom{\hat{m}_j}{l_j} B(k_j + l_j + 1, 2\hat{m}_j - k_j - l_j) - 1\Big). \tag{3.29}$$

Thus, using our proposed ECBC copula not only are we able to obtain a fully non-parametric estimate of any copula function in closed form (once the tuning parameters m_j, $j = 1, \ldots, d$ are estimated by their posterior modes) but also we are able to derive the closed-form expression of estimates of the popular multivariate measures of dependence for any arbitrary dimension $d \geq 2$.

Moreover, although we only illustrate the use of multivariate extensions of Kendall's tau and Spearman's rho as possible measures of multivariate dependence, any other multivariate notion of dependence measures that are suitable functionals of the underlying copula can also be computed using our closed-form expression of the ECBC estimator. This is particularly advantageous compared to even some of the flexible yet complicated parametric copula family (e.g., Archimedian, multivariate Gaussian or t, etc.) for which it may require high-dimensional numerical integration to compute multivariate versions of Spearman's rho as given in (3.25) and/or Kendall's tau given in (3.28). For vine copulas, it is particularly challenging to obtain estimates of these multivariate measures of dependence as such high-dimensional integrals are often algebraically and even numerically intractable say for dimension $d \geq 5$, whereas for ECBC even when a new measure of dependence is created as a functional of the copula that may be more complicated than those defined in eq.s (3.25) and (3.28), we can easily obtain a large number of Monte Carlo (MC) samples from the ECBC and use MC-based approximation to estimate such new measures of multivariate dependence (we illustrate such a case in our real case study involving portfolio risk optimization in Section 3.4).

3.3 Numerical Illustrations using Simulated Data

3.3.1 Finite-Sample Performance for Bivariate Cases

We investigate the finite-sample performance of the ECBC through a Monte Carlo simulation study. Samples from the true copula are generated using the package `copula` in R (Hofert et al. [2014]). In order to visualize the results using contour plots, we first restrict our illustration to bivariate copulas. Three copula families with various parameters and an asymmetric copula are considered. The first four examples are the bivariate Frank copulas

$$C_F(u,v) = -\frac{1}{\theta}\log\Big(1+\frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1)}\Big),$$

with parameter θ equal to $-2,-1,1$ and 2, which reflects a wide range of dependence from negative to positive. The next two examples are the Clayton copula with parameter 1

$$C_C(u,v) = (\max\{u^{-1}+v^{-1}-1,0\})^{-1},$$

and the Gumbel copula with parameter 2

$$C_G(u,v) = \exp(-((-\ln(u))^2+(-\ln(v))^2)^{1/2}).$$

The value of Kendall's tau is 0.33 for Clayton copula with parameter 1 and 0.5 for Gumbel copula with parameter 2. Both cases have a moderate positive dependence.

The next example is the independent copula $C(u,v) = uv$. Finally, we consider an asymmetric copula

$$C_a(u,v) = uv - 0.12(1-v^2)\sin(8.3v)u(1-u),$$

In the simulation study, $n = 100$ samples are drawn from the true copula for each

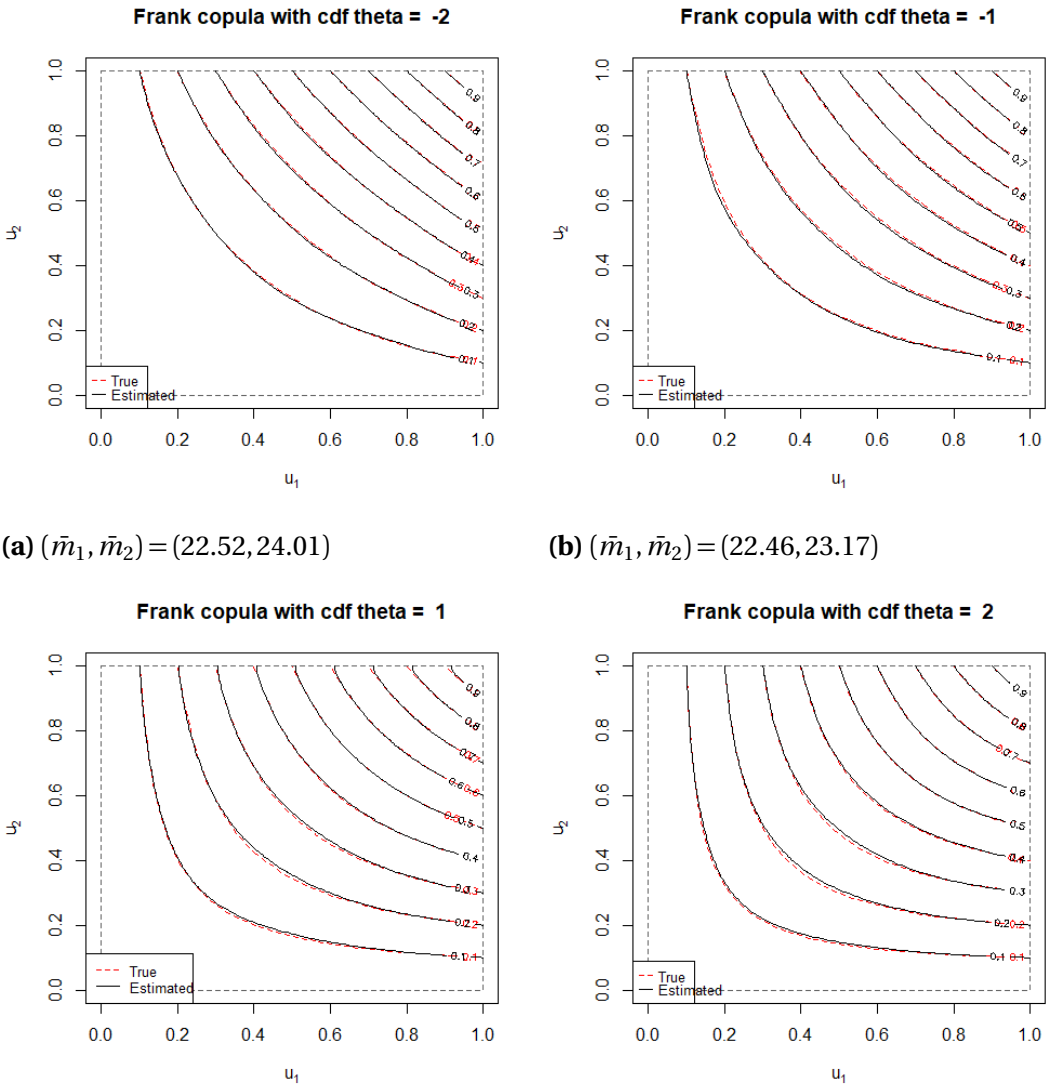
replicate (of size $n = 100$) and there are $N = 100$ replicates. Degrees of the ECBC are estimated by posterior modes by obtaining 5000 MCMC samples following 2000 burn-in samples of two chains for each of the N replicated data sets generated from a chosen true copula model. It is to be noted that it takes about 2, 11, and 60 minutes to run 7000 iterations for two chains by using MacOS with 16GB of RAM for $d = 2$, $d = 10$, and $d = 50$, respectively, when data are drawn from multivariate Frank copula. Convergence of MCMC runs was monitored based on preliminary runs using standard diagnostics available in R packages `rjags` and `coda`. We show the results for eight copulas in Figure 3.1 and 3.2. The contour plot of the true underlying copula and the empirical MC average of N copula estimates are given for comparison and $(\bar{m}_1, \bar{m}_2)$ represents the MC mean of the posterior modes of degree parameters.

We can see from the contour plots that the average of the estimated copula is extremely close to the underlying true copula across all different dependence structures irrespective of the assumed parametric models. This illustrates that the proposed ECBC has a robust performance in estimating various true copulas.

3.3.2 Accuracy of Multivariate Dependence Measures ($d = 3$)

To assess the finite sample performance of the estimate of multivariate Spearman's rho $\hat{\rho}_d$, we conduct Monte Carlo simulations for $d = 3$ copulas. We consider the independent copula and Clayton copulas with parameter value $\{0.5, 1, 2\}$, respectively. For each copula model, $N = 100$ Monte Carlo replicates are generated with size $n = 100$. For each replicate, we compute the proposed estimator $\hat{\rho}_d$ in (3.26) and the estimator based on empirical copula $\tilde{\rho}_d$

$$\begin{aligned} \tilde{\rho}_d &= \frac{d+1}{2^d-(d+1)}\Big(2^d \int_{I^d} C_n(\mathbf{u}) d\mathbf{u} - 1\Big) \\ &= \frac{d+1}{2^d-(d+1)}\Big(\frac{2^d}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}(1-U_{ij}) - 1\Big). \end{aligned} \tag{3.30}$$



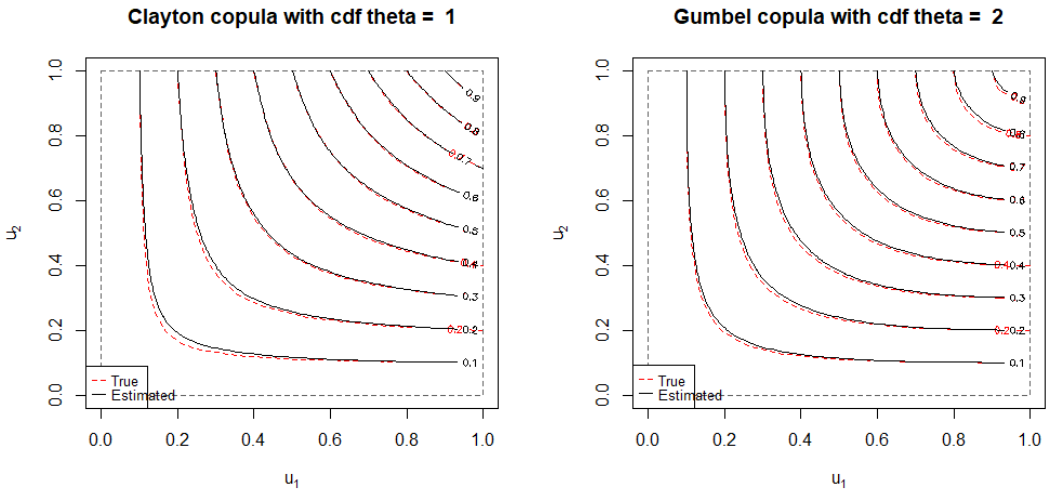
(a) $(\bar{m}_1, \bar{m}_2) = (22.52, 24.01)$

(b) $(\bar{m}_1, \bar{m}_2) = (22.46, 23.17)$

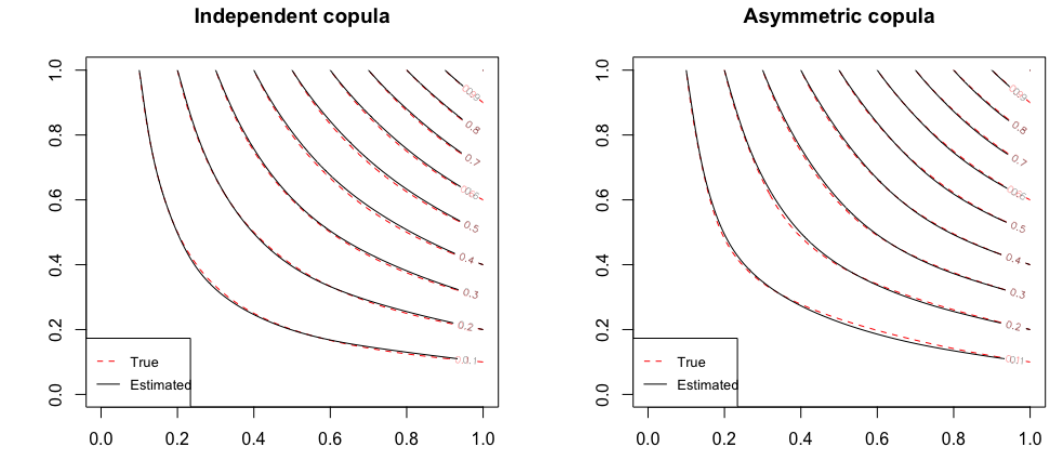
(c) $(\bar{m}_1, \bar{m}_2) = (23.29, 24.83)$

(d) $(\bar{m}_1, \bar{m}_2) = (22.82, 22.17)$

Figure 3.1 Estimation of Frank copulas using the ECBC with empirical Bayesian method for choosing proper degrees when sample size $n = 100$.



(a) $(\bar{m}_1, \bar{m}_2) = (23.94, 21.69)$ **(b)** $(\bar{m}_1, \bar{m}_2) = (22.58, 22.16)$



(c) $(\bar{m}_1, \bar{m}_2) = (22.34, 20.87)$ **(d)** $(\bar{m}_1, \bar{m}_2) = (21.08, 22.92)$

Figure 3.2 Estimation of various copulas using the ECBC with empirical Bayesian method for choosing proper degrees when sample size $n = 100$.

Finally, for each estimator we compute the mean, bias, variance and mean square error (MSE) over all replicates.

Table 3.1 shows the results on the estimation of multivariate Spearman's rho for four different copulas. An approximated value of the true multivariate Spearman's rho ρ_d can be obtained by numerical integration since there is no analytical expression as a function of the parameter (see [Pérez & Prieto-Alaiz, 2016]), which is also given in Table 3.1. The corresponding boxplots for the two estimates based on $N = 100$ replicates along with a horizontal line for true multivariate Spearman's rho ρ_d are shown in Figure 3.3.

From the results we can see our estimator $\hat{\rho}_d$ outperforms $\tilde{\rho}_d$ with respect to variance and MSE. In terms of bias, $\hat{\rho}_d$ tends to underestimate and have a larger bias as strength of dependence increases, but there is not a clear superiority of one estimator over the other. As shown in Figure 3.3 (c) and (d) where there is a moderate or strong dependence in trivariate cases, $\tilde{\rho}_d$ can take values out of parameter space $[-2/3, 1]$ (3% and 12% of $\tilde{\rho}_d$ are taking values greater than 1 in (c) and (d), respectively), which can be problematic in measuring dependence.

Table 3.1 Comparison of two estimates of multivariate Spearman's rho based on $N = 100$ replications of size $n = 100$ generated from the independent copula and Clayton copulas with different parameter values when dimension $d = 3$.

Copula	ρ_d	Estimator	Mean	Bias	Variance	MSE
Independent	0	$\hat{\rho}_d$	0.007	0.007	0.008	0.009
		$\tilde{\rho}_d$	-0.015	-0.015	0.014	0.015
Clayton	0.308	$\hat{\rho}_d$	0.294	-0.014	0.007	0.008
$\theta = 0.5$		$\tilde{\rho}_d$	0.287	-0.021	0.028	0.029
Clayton	0.504	$\hat{\rho}_d$	0.477	-0.027	0.007	0.008
$\theta = 1$		$\tilde{\rho}_d$	0.519	0.015	0.039	0.040
Clayton	0.717	$\hat{\rho}_d$	0.680	-0.037	0.004	0.006
$\theta = 2$		$\tilde{\rho}_d$	0.732	0.015	0.050	0.051

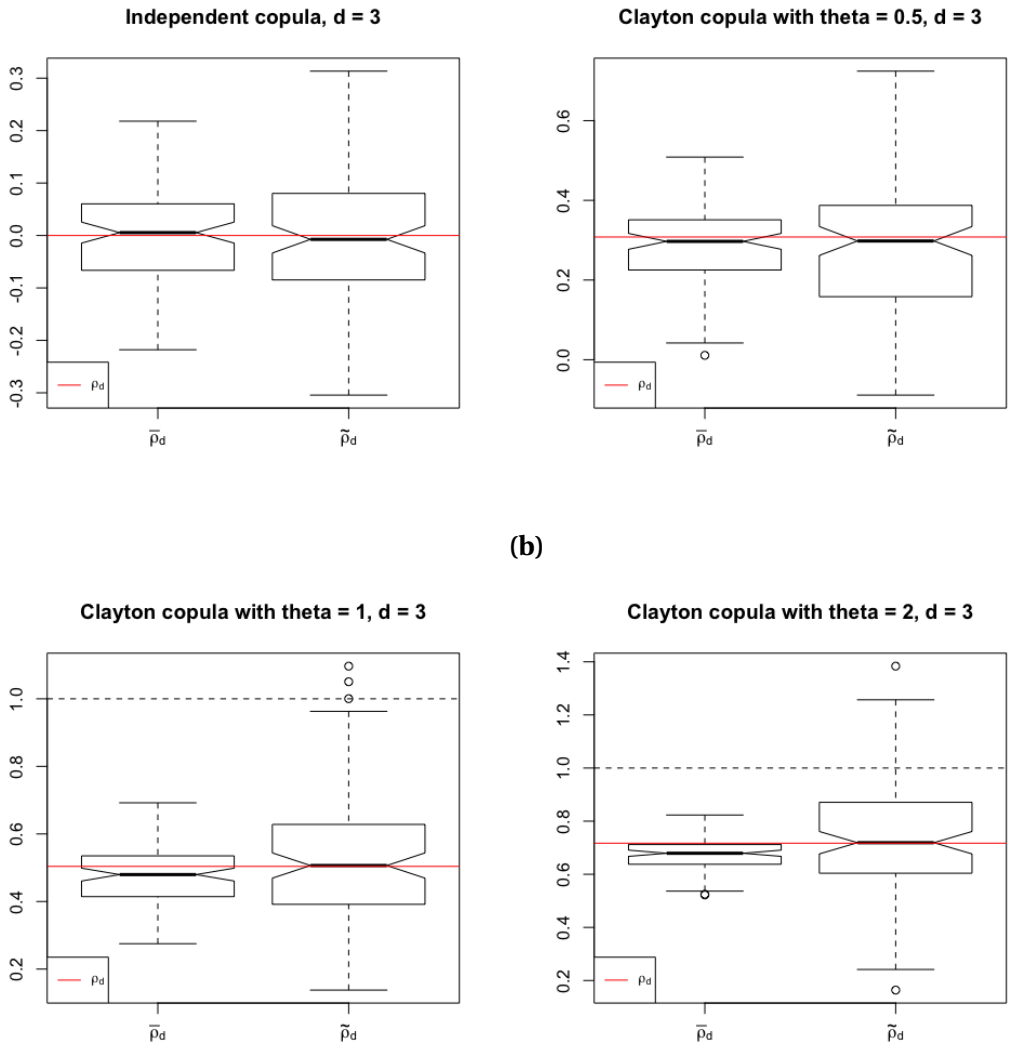


Figure 3.3 Boxplots of the two estimators based on $N = 100$ replicates and a horizontal line of true multivariate Spearman's rho ρ_d for each of four trivariate copulas.

3.3.3 Estimation of Tuning parameters of ECBC

We now illustrate one of the primary advantages of our proposed empirical Bayes estimate of the ECBC that allows for data-dependent automatic selection of dimension-varying degree parameters m_j's. We further explore the special case of choosing equal degrees $m_1 = \ldots = m_d = m$ by using the following prior distribution

$$m \mid \alpha \sim Poisson(n^{\alpha}) + 1, \tag{3.31}$$

$$\text{and } \alpha \sim Unif\Big(\frac{1}{3}, \frac{2}{3}\Big) \tag{3.32}$$

For our empirical illustration, we consider three true bivariate copulas and one trivariate copula to explore the comparative performance of choosing dimension-varying degrees compared to setting them equal across all dimensions. The first example is the Farlie-Gumbel-Morgenstern (FGM) copula $C_a(u,v) = uv(1-a(1-u)(1-v))$ with parameter $a = -1$. The next two choices are the independent copula (e.g., FGM with $a = 0$) and the Gaussian copula with positive dependence (correlation $\rho = 0.5$). The last one is the trivariate t-copula with degrees of freedom 4 and pairwise dependence $\rho_{12} = -0.2, \rho_{13} = 0.5$ and $\rho_{23} = 0.4$. Again samples of size $n = 100$ are obtained for each four cases and repeated $N = 100$ times for MC evaluation. For each sample, we choose the degrees of the ECBC by computing the posterior modes using our proposed empirical Bayesian method.

Figure 3.4 presents the scatterplot of estimated values of (m_1, m_2) or (m_1, m_2, m_3) for each chosen true copula model. From the plots we can observe that for bivariate copulas, posterior estimates of m_1 and m_2 are significantly different in most cases without any prior restrictions of equality. In fact, posterior probability of choosing equal $m_1 = m_2$ is only about 0.08, 0.08 and 0.13 for FGM copula, for independent copula, and for Gaussian copula, respectively indicating against forcing $m_1 = m_2$ as is popularly done in the literature. For the trivariate t-copula case, the posterior probability of $m_1 = m_2 = m_3$ is 0, decisively suggesting that equality assumption is sub-optimal in general and particular as the dimension in-

creases. We have conducted further studies with dimensions (not shown here for limitation of space) and the conclusions remain very similar.

In order to further compare the performances of copula estimators with flexible degrees vs. equal degrees, we fit our Bayesian models with two different settings of priors: (i) (flexible) the original priors given in (3.19) and (3.20) where degrees are allowed to vary in different dimensions; (ii) (equal) modified priors given in (3.31) and (3.32) where degrees are set to be equal; using the same data set generated from the three bivariate copulas. Following Segers et al. [2017], we consider three global performance measures: the integrated squared bias, the integrated variance, and the integrated mean squared error. Given a copula estimator $\hat{C}_n$, the performance measures are defined as

$$\text{integrated squared bias: } ISB = \int_{[0,1]^d} \left[E[\hat{C}_n(\mathbf{u}) - C(\mathbf{u})]\right]^2 d\mathbf{u}, \tag{3.33}$$

$$\text{integrated variance: } IV = \int_{[0,1]^d} E\left[[\hat{C}_n(\mathbf{u}) - E(\hat{C}_n(\mathbf{u}))]^2\right] d\mathbf{u}, \tag{3.34}$$

$$\text{integrated mean squared error: } IMSE = \int_{[0,1]^d} E\left[[\hat{C}_n(\mathbf{u}) - C(\mathbf{u})]^2\right] d\mathbf{u}. \tag{3.35}$$

We compute the performance measures by applying the computation method described in Segers et al. [2017], which relies on Monte Carlo simulation to get a Monte Carlo estimate of each performance measure. Table 3.2 presents the results for the four cases, where the first three are bivariate ($d = 2$) and the last one is trivariate ($d = 3$). The standard errors of the Monte Carlo estimates are not reported in the table as they are negligibly small. We can see that the copula estimators with flexible degrees perform better than those with equal degrees in terms of IV and IMSE in all cases. As the difference in IMSE is dominated by the IV term, the choice of flexible degrees leads to smaller uncertainty, and hence smaller IMSE while the biases remain relatively unaffected. It is also interesting to observe that estimated degrees are far smaller than sample sizes indicating the empirical beta copula may not have optimal performance, which we next explore.

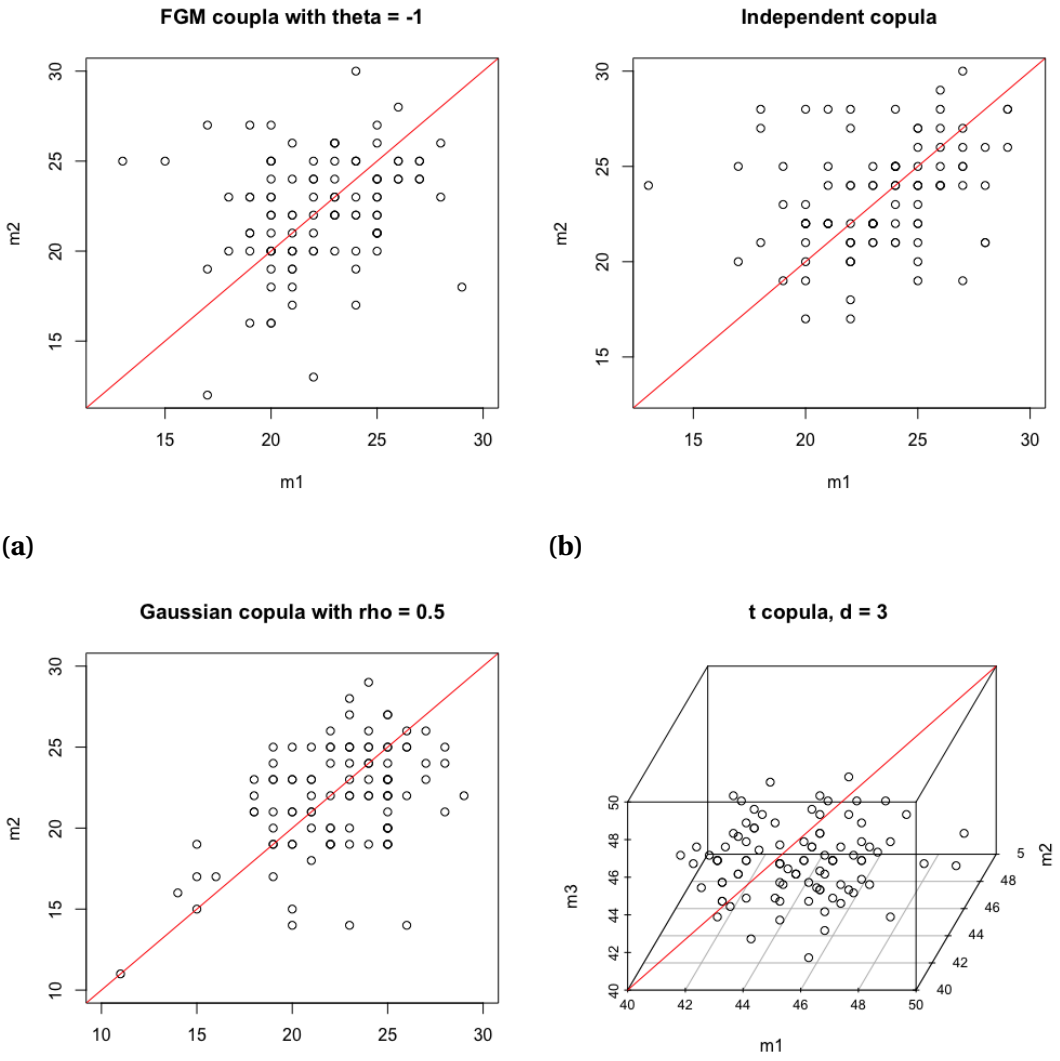


Figure 3.4 Choice of dimension-varying degrees obtained by applying the proposed empirical Bayesian method based on $N = 100$ replications when sample size $n = 100$.

Table 3.2 Comparison of copula estimators with flexible degrees vs. equal degrees using three performance measures computed by Monte Carlo simulation based on $N = 100$ replications when sample size $n = 100$.

Copula	Choice of degrees	ISB ($\times 10^{-4}$)	IV ($\times 10^{-4}$)	IMSE ($\times 10^{-4}$)
FGM	Flexible	0.10	1.28	1.38
$\theta = -1$	Equal	0.09	1.37	1.46
Independent	Flexible	0.13	1.89	2.02
	Equal	0.20	2.16	2.36
Gaussian	Flexible	0.36	0.71	1.07
$\rho = 0.5$	Equal	0.30	0.83	1.13
t	Flexible	0.11	2.67	2.78
$d = 3$	Equal	0.19	2.79	2.98

3.3.4 Comparison with the Empirical Bernstein Copulas

In this section, we compare the finite-sample performance of the ECBC with other non-parametric copula estimators only as it has been already demonstrated that parametric models based methods lead to biased estimates under model misspecification. First, we consider the empirical beta copula introduced by Segers et al. [2017], which is a special case of the empirical Bernstein copula where the degrees of the polynomials are set equal to the sample size. The empirical beta copula is a genuine copula and has been shown to outperform the classical empirical copula and the empirical checkerboard copula in terms of bias and variance. For bivariate cases, we also include the empirical Bernstein copula with degrees as suggested in Janssen et al. [2012] into the comparison. By setting degrees $m_1 = m_2 = m$ and minimizing the asymptotic pointwise mean squared error with respect to m, Janssen et al. [2012] suggested the choice of m in the bivariate case as

$$m_0(u_1, u_2) = \Big(\frac{4b^2(u_1, u_2)}{V(u_1, u_2)}\Big)^{2/3} n^{2/3} \tag{3.36}$$

where

$$
\begin{aligned}
b^2(u_1,u_2) &= \frac{1}{2}\sum_{j=1}^{2} u_j(1-u_j)C_{u_ju_j}(u_1,u_2) \\
\text{and } V(u_1,u_2) &= \sum_{j=1}^{2} C_{u_j}(u_1,u_2)(1-C_{u_j}(u_1,u_2))\Big(\frac{u_j(1-u_j)}{\pi}\Big)^{1/2}
\end{aligned}
$$

and C_{u_j} and $C_{u_ju_j}$ are the first-order and second-order partial derivatives of C, respectively, with respect to u_j, $j=1,2$. Note that even we use the integer part $\lfloor m_0(u_1,u_2)\rfloor$ in practice, it is not necessarily a divisor of n, meaning the empirical Bernstein copula with $m=\lfloor m_0(u_1,u_2)\rfloor$ is not guaranteed to be a genuine copula.

We consider the same copula models as in Section 3.3.3. The choice of degrees $m_0(u_1,u_2)$ suggested by Janssen et al. [2012] is not defined for the independent copula as $C_{u_ju_j}=0$ and it is restricted to bivariate cases, so we only take the empirical Bernstein copula into consideration for the bivariate FGM and Gaussian copulas. All results in Table 3.3 are based on $N=100$ MC replications each of sample sizes $n=25,50,100$. We compare the performance of the ECBC with flexible degrees (referred to as flexible ECBC), the empirical beta copula (referred to as Beta), and the empirical Bernstein copula with $m=m_0(u_1,u_2)$ (referred to as Bernstein) using the same performance measures as in Section 3.3.3.

Table 3.3 indicates that the ECBC with flexible degrees outperforms the empirical beta copula in terms of variance and mean square error in all cases. Compared to the empirical Bernstein copula with $m=m_0(u_1,u_2)$, the ECBC with flexible degrees has a smaller bias but the ordering with respect to mean square error is not clear between these two copula estimators. For small samples, the empirical beta copula seems to have the largest variance while the bias of the empirical Bernstein copula with $m=m_0(u_1,u_2)$ is shown to be the largest even though it uses optimal "true" degree given in (3.36).

Table 3.3 Comparison of the ECBC with flexible degrees (referred to as flexible ECBC), the empirical beta copula(referred to as Beta) and the empirical Bernstein copula with $m = m_0(u_1, u_2)$ (referred to as Bernstein) using three performance measures computed by Monte Carlo simulation based on $N = 100$ replications for sample size $n = 25, 50, 100$.

Copula	Estimator	ISB ($\times 10^{-4}$)			IV ($\times 10^{-4}$)			IMSE ($\times 10^{-4}$)		
		n =			n =			n =		
		25	50	100	25	50	100	25	50	100
FGM	flexible ECBC	0.29	0.07	0.10	3.46	2.72	1.28	3.75	2.79	1.38
$\theta = -1$	Beta	0.30	0.09	0.02	5.24	3.27	2.17	5.54	3.36	2.19
	Bernstein	0.83	0.51	0.07	1.52	1.33	1.08	2.35	1.84	1.15
Independent	flexible ECBC	0.68	0.31	0.13	2.37	2.33	1.89	3.05	2.64	2.02
	Beta	0.02	0.03	0.01	5.11	4.02	2.25	5.13	4.05	2.26
	Bernstein	NA	NA	NA	NA	NA	NA	NA	NA	NA
Gaussian	flexible ECBC	1.72	0.79	0.36	2.51	1.67	0.71	4.23	2.46	1.07
$\rho = 0.5$	Beta	0.26	0.16	0.20	4.78	3.03	1.81	5.04	3.19	2.01
	Bernstein	6.82	3.09	1.27	2.28	1.42	1.06	9.10	4.51	2.33
t	flexible ECBC	0.41	0.30	0.11	6.76	4.26	2.67	7.17	4.56	2.78
$d = 3$	Beta	0.27	0.15	0.04	7.55	5.05	3.07	7.82	5.20	3.11
	Bernstein	NA	NA	NA	NA	NA	NA	NA	NA	NA

3.4 Application to Portfolio Risk Management

Copulas have been widely used in portfolio optimization and risk measurement as they are powerful tools to model the dependence among different assets in a portfolio. The proposed ECBC is capable of estimating multivariate copula and it is straightforward to sample from the estimated copula, so it can be applied to find optimal weights and estimate risk measures for a portfolio with a variety of assets.

We now illustrate the use of ECBC for portfolio risk allocation using real data consisting of a d asset values. Value at risk (VaR) and conditional value at risk (CVaR) (also called expected shortfall (ES)) are common measures of risk in the field of risk management (see e.g. Jorion [2007] and Uryasev [2000]). Assume that X is the return of a portfolio or asset (daily log-return of a portfolio of stocks or individual stocks, with positive indicating profit and negative values representing loss) with distribution function $F_X(\cdot) = \Pr[X \leq x]$. The VaR of X at the level of $\alpha \in (0,1)$ is defined as

$$VaR_\alpha(X) = -\inf\{x \in \mathbb{R} : F_X(x) > \alpha\},$$

while the CVaR (ES) of X is defined as

$$CVaR_\alpha(X) = E(-X | X \leq -VaR_\alpha(X)).$$

Notice that, if we consider the corresponding loss of the same portfolio represented by $Y = -X$, then we have $VaR_\alpha(X) = VaR_\alpha(Y) = F_Y^{-1}(1-\alpha)$ and $CVaR_\alpha(X) = CVaR_\alpha(Y) = E(Y | Y \geq VaR_\alpha(Y))$.

Mean-CVaR portfolio optimization is a popular portfolio optimization technique introduced by Rockafellar & Uryasev [2000]. The advantage of Mean-CVaR portfolio optimization is that it calculates VaR and minimizes CVaR simultaneously where the optimization can be formulated as a linear programming problem.

Let $\mathbf{x} \in \mathbb{R}^d$ denote a realized return value of d assets in a portfolio, and $\mathbf{v} \in \mathscr{S}_d = \{\mathbf{v} \in \mathbb{R}^d : v_j \geq 0,\ \forall j,\ \sum_{j=1}^d v_j = 1\}$, denote the portfolio weights to be determined within the d-dimensional simplex $\mathscr{S}_d$. The key to the approach in Rockafellar & Uryasev [2000] is the auxiliary function for CVaR taking the form of

$$H_\alpha(\mathbf{v},\gamma) = \gamma + \frac{1}{\alpha}\int_{l(\mathbf{v},\mathbf{x})\geq\gamma} (l(\mathbf{v},\mathbf{x}) - \gamma) dF(\mathbf{x}), \tag{3.37}$$

where $l(\mathbf{v},\mathbf{x}) = -\mathbf{v}^T\mathbf{x}$ is a linear loss function and $F(\mathbf{x})$ is the joint distribution function of daily (random) return vector $\mathbf{X}$ which we will estimate using our proposed ECBC based empirical Bayes method. It has been shown in Theorem 1 of Rockafellar & Uryasev [2000] that for any weights $\mathbf{v}$, $H_\alpha(\mathbf{v},\gamma)$ is convex as a function of γ and is equal to $CVaR_\alpha(\mathbf{v})$ at the minimum point. Moreover, $VaR_\alpha(\mathbf{v})$ would be the left endpoint of $\arg\min_\gamma H_\alpha(\mathbf{v},\gamma)$. Moreover, Minimizing $CVaR_\alpha(\mathbf{v})$ with respect to $\mathbf{v}$ is equivalent to minimizing $H_\alpha(\mathbf{v},\gamma)$ with respect to $(\mathbf{v},\gamma)$ (e.g., see Theorem 2 of Rockafellar & Uryasev [2000] for details). To numerically approximate the integral in (3.37), it is often good enough to generate M samples from $F(\cdot)$ or its estimate, which can be done by using the proposed empirical Bayes method based on the ECBC. However, a relatively less answered question in finance is how large should we choose M for accurate estimation as the integral in (3.37) depends on sampling the tail part of $F(\cdot)$ or its estimate. The empirical estimate of $H_\alpha(\mathbf{v},\gamma)$ based on generating $\mathbf{x}_k \overset{iid}{\sim} F$ or $\hat{F}$ can be written as

$$F_\alpha(\mathbf{v},\gamma) = \gamma + \frac{1}{\alpha M}\sum_{k=1}^{M}(-\mathbf{v}^T\mathbf{x}_k - \gamma)_+ \tag{3.38}$$

where $(\cdot)_+ = \max(\cdot, 0)$.

We claim that in order to achieve an accuracy of $\epsilon > 0$ for the MC approximation, it is sufficient to generate M MC samples such that

$$\sqrt{\frac{2\log\log M}{M}\lambda_{max}} \leq \epsilon \ \ i.e.\ \frac{M}{\log\log M} \geq \frac{2\lambda_{max}}{\epsilon^2} \tag{3.39}$$

where $\Sigma = Var_F(\mathbf{X})$ and λ_{max} is the largest eigenvalue of Σ. Details are given in section 3.6.2. Notice that Σ and hence λ_{max} can be easily estimated from the observed return values without any modeling assumption as long as $n > d$, however sparse methods are necessary for large size portfolios when $n \leq d$. Next it can be shown that minimizing (3.38) is equivalent to minimizing

$$F_\alpha(\mathbf{v},\gamma) = \gamma + \frac{1}{\alpha M}\sum_{k=1}^{M} z_k \;\; s.t. \;\; z_k \geq 0, z_k + \mathbf{v}^T\mathbf{x}_k + \gamma \geq 0 \tag{3.40}$$

Thus, along with the linear constraints on the weights $\mathbf{v}$, it can be formulated as a linear programming problem and can be solved using standard convex optimization methods. Conveniently, R function `BDportfolio_optim` within the package `PortfolioOptim` can be used for this purpose. Following the algorithm in Semenov & Smagulov [2017], simulated return values can be obtained using estimated ECBC for portfolio optimization. The complete algorithm is summarized below:

Step 1. Transform assets' historical data X_{tj} to pseudo-observations U_{tj} and estimate copula using our proposed method.

$$U_{tj} = F^*_{Tj}(X_{tj}), \;\; t = 1,\ldots,T, j = 1,\ldots,d, \tag{3.41}$$

$$F^*_{Tj}(x_j) = \frac{1}{T+1}\sum_{t=1}^{T} I(X_{tj} \leq x_j), \;\; j = 1,\ldots,d. \tag{3.42}$$

Step 2. Generate a sample of pseudo-observations $(U^*_{k1},\ldots,U^*_{kd}), k = 1,\ldots,M$ from the estimated ECBC using empirical Bayes method and transform simulated pseudo-observations to univariate quantiles.

$$X^*_{kj} = F^{*-1}_{Tj}(U^*_{kj}), \;\; k = 1,\ldots,M, j = 1,\ldots,d. \tag{3.43}$$

Step 3. Calculate optimal weights $v_j, j = 1,\ldots,d$ using simulated data $(X^*_{k1},\ldots,X^*_{kd})$, $k = 1,\ldots,M$, and the corresponding VaR and CVaR, which are by-products of the portfolio

optimization, by solving the linear programming problem given in (3.40).

Our copula estimator is useful to find optimal weights and estimate risk measures as sampling from the estimated copula is straightforward. Considering the Bernstein copula density given in (3.10) and (3.11), we can obtain samples $(U_1, \dots U_d) \sim C_m$ as follows

$$\begin{aligned} (k_1, \dots, k_d) &\sim \tilde{w}_{k_1,\dots,k_d}, \;\; k_j = 0, \dots, m_j - 1, \, j = 1, \dots, d \\ U_j &\sim \text{Beta}(k_j + 1, m_j - k_j), \;\; j = 1, \dots, d \end{aligned} \tag{3.44}$$

3.4.1 A Portfolio of 3 Stocks

We investigate the time series of daily closing stock prices of three companies: Amazon.com, Inc. (AMZN), Facebook, Inc. (FB), and Alphabet Inc. (GOOGL) for the time period from January 1, 2018, to December 31, 2019. This data set consists of 502 observations and can be obtained using R package `quantmod`. Larger portfolios can be considered but for ease of illustration we restrict to $d = 3$ case only. See the supplementary document for an extended illustration with a portfolio of size $d = 10$.

Suppose we want to find an optimal portfolio of stocks including AMZN, FB, and GOOGL that minimizes the expected shortfall of the portfolio. First, we convert the price series P_{tj} to log-returns X_{tj}

$$X_{tj} = log \frac{P_{tj}}{P_{(t-1)j}}, \;\; t = 1, \dots, T, \, j = 1, \dots, d, \tag{3.45}$$

resulting in $T = 501$ log-return values for $d = 3$ assets. Then we follow Steps 1-3 as above to obtain the optimal portfolio weights and the corresponding VaR and CVaR. Following Semenov & Smagulov [2017], we set the minimum weight to be limited by $v_j \geq 0.05$, $j = 1, \dots, d$ to avoid corner portfolio cases.

In Step 1, the posterior mode estimator of the degrees of the proposed ECBC are $(m_1, m_2, m_3) = (166, 170, 171)$. The largest eigenvalue of the covariance matrix is $\lambda_{max} \approx 8 \times 10^{-4}$, so we are able to find the value of M that is sufficient for a given accuracy ϵ from

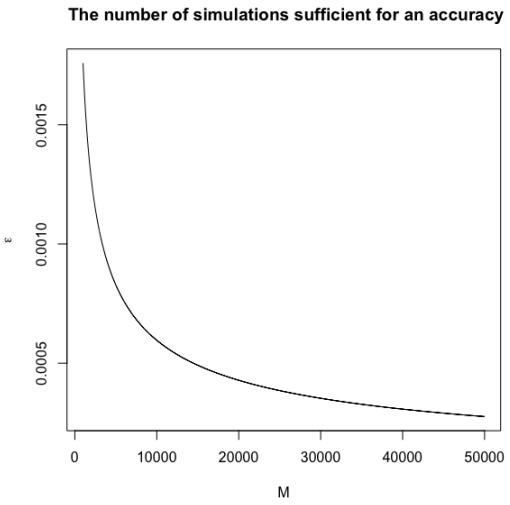


Figure 3.5 The value of M sufficient for the accuracy ϵ.

the relationship in (3.39). The value of M as a function of the accuracy ϵ is shown in Figure 3.5.

We set $M = 10000$ (adequate for an accuracy $\epsilon \approx 6 \times 10^{-4}$) and repeat Step 2-3 $N = 100$ times to quantify estimation uncertainty. For each replicate we conduct portfolio optimization at the level of $\alpha \in \{0.10, 0.05, 0.01\}$ as popularly used. As a result, we are able to obtain the distribution of optimal weights (Figure 3.6) and risk measures (Figure 3.7) using simulated data from the estimated copula.

From the boxplots of optimal weights in Figure 3.6 we can see that GOOGL has a much higher weight than AMZN and FB in the mean-CVaR optimal portfolio across different levels. The optimal weight of AMZN tends to increase as the level decreases.

Table 3.4 summarizes the MC average and standard deviations of estimated optimal weights, VaR, and CVaR obtained from simulated data using the proposed ECBC.

Also, by applying mean-CVaR portfolio optimization to the historical log-return data X_{tj}, we can get estimates of optimal weights and risk measures as well. Table 3.5 shows the empirical weights, VaR, and CVaR of optimal portfolio at the level of 0.10, 0.05 and 0.01, respectively.

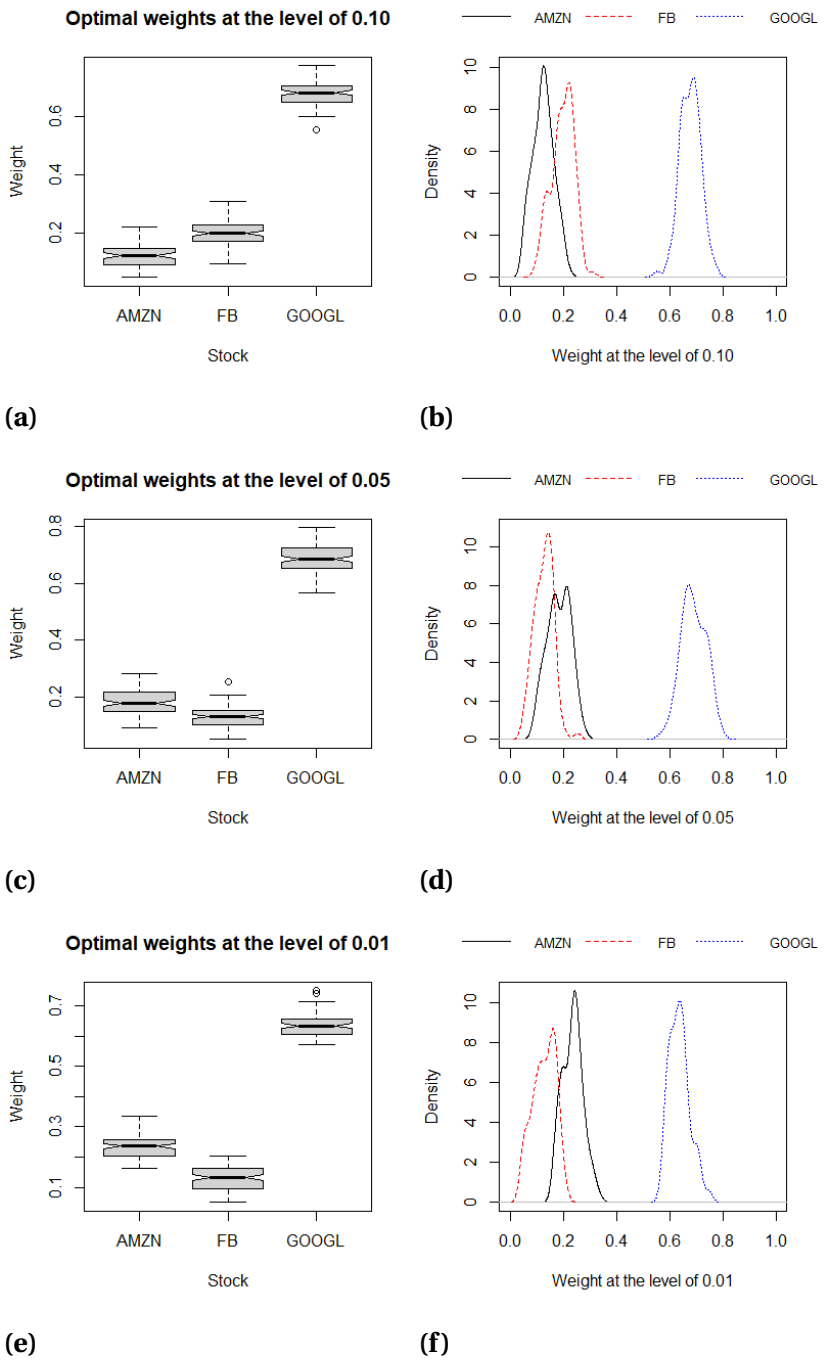
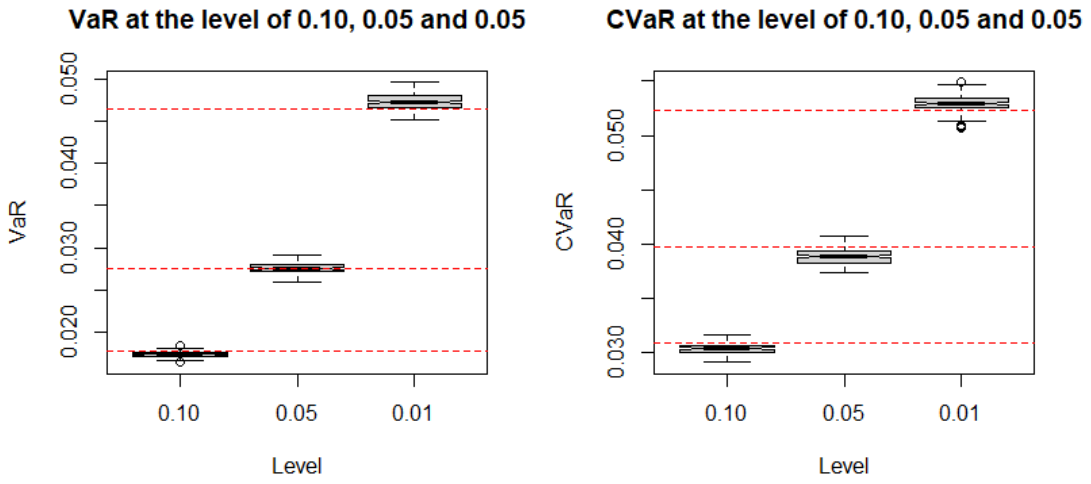


Figure 3.6 Distribution of optimal weights obtained from simulated data using the estimated copula at the level of 0.10, 0.05, and 0.01.



(a) **(b)**

Figure 3.7 Distribution of VaR and CVaR obtained from simulated data using the estimated copula at the level of 0.10, 0.05, and 0.01. Dashed lines indicate empirical estimates using historical data.

Table 3.4 Means (and SDs) of estimated optimal weights, VaR, and CVaR obtained from simulated data at the level of 0.10, 0.05, and 0.01.

Level	Weights			VaR	CVaR
	AMZN	FB	GOOGL	$\times 10^{-2}$	$\times 10^{-2}$
0.10	0.125 (0.039)	0.198 (0.042)	0.677 (0.039)	1.741 (0.036)	3.034 (0.057)
0.05	0.183 (0.044)	0.127 (0.036)	0.690 (0.046)	2.759 (0.066)	3.882 (0.070)
0.01	0.235 (0.038)	0.128 (0.042)	0.637(0.038)	4.729 (0.093)	5.296 (0.082)

Table 3.5 Optimal weights, VaR, and CVaR obtained from historical data at the level of 0.10, 0.05, and 0.01.

Level	Weights			VaR	CVaR
	AMZN	FB	GOOGL	$\times 10^{-2}$	$\times 10^{-2}$
0.10	0.156	0.089	0.755	1.783	3.093
0.05	0.104	0.143	0.753	2.754	3.972
0.01	0.234	0.207	0.559	4.649	5.236

Although the optimal weights obtained from historical data are different from the means of those obtained from simulated data using the proposed copula estimator, the estimated risk measures from two different methods seem to be fairly close. However, we are able to quantify the uncertainty for all the estimates by repeatedly sampling from the estimated copula. Semenov & Smagulov [2017] conduct a similar stability study to report the means and SDs of VaR and CVaR, but they used predetermined weights based on historical data and didn't report the distribution of optimal weights obtained from simulated data. Our copula estimator has shown good performance for relatively small samples and operationally we can generate as many samples as we want from the estimated copula, thus the copula-based method would be more reliable when there is not sufficient historical data. Besides, compared to the empirical estimates, it is possible to estimate VaR and CVaR for much smaller values of levels using the copula-based method.

3.4.2 A Portfolio of 10 Stocks

As an example with moderately large dimension, we consider the time series of daily closing stock prices of 10 top Nasdaq companies: AMZN, FB, GOOGL, AAPL, MSFT, INTC, CSCO, NFLX, CMCSA, and ADBE from January 1, 2018, to December 31, 2019, which consists of 502 observations and $T = 501$ log-return values for $d = 10$ assets. The minimum weight is set to be limited by 0.01 to avoid corner portfolio cases. The degrees of the proposed ECBC are estimated by the posterior mode $(209, 206, 208, 206, 208, 209, 211, 210, 209, 208)$.

Similar to the settings when $d = 3$, we set $M = 10000$ (adequate for an accuracy $\epsilon \approx 9 \times 10^{-4}$) and repeatedly sample $N = 100$ times from the estimated copula. Portfolio optimization is conducted at the level of $\alpha \in \{0.10, 0.05, 0.01\}$ for each replicate. The distributions of optimal weights and risk measures are shown in Figure 3.8 and Figure 3.9. In terms of computational time, for a given data set of size $T = 501$ it takes about 20 minutes to run 3000 iterations (2000 MCMC samples following 1000 burn-in) for each of two chains to obtain posterior estimates of the m_j's using MacOS with 16GB of RAM for

$d = 3$ and about 50 minutes for $d = 10$. Notice that computation of the optimal weights and its empirical distribution takes only about 45 and 113 seconds for dimensions $d = 3$ and $d = 10$, respectively. The MCMC codes will be realized on the author's GitHub page for broader dissemination.

3.5 Concluding Remarks

In this chapter, we propose the empirical checkerboard Bernstein copula, which is a nonparametric multivariate copula estimator. It can be considered as an advancement of the empirical Bernstein copula since it is a valid copula with any polynomial degrees for any sample size. For automatic data-dependent dimension-varying degree selections, we further developed an empirical Bayesian method that has been shown to be practically useful. While the proposed copula estimator is shown to be large-sample consistent, it also has a good finite-sample performance. Moreover, it has a beneficial effect on measuring the strength of dependence for large dimensions because the estimates derived from the proposed copula are always within the proper range.

As sampling from the estimated copula is quite straightforward, it is applicable to portfolio optimization and risk measurement where estimation is often done with simulations generated from copulas. We investigate the number of simulations that are good enough to achieve any given accuracy, which has been apparently out of reach in the literature. Furthermore, we are able to provide uncertainty quantification for all the estimates in portfolio risk management.

Under the hierarchical structure of the proposed empirical Bayes model, MCMC methods have been shown to work reasonably fast for relatively large dimensions ($d \leq 50$) with a moderate sample size ($n = 100$). To speed up the MCMC methods for very large sample sizes, it would be of interest to explore some scalable MCMC methods such as divide-and-conquer approaches and subsampling approaches (see e.g. Robert et al. [2018], Quiroz et al.

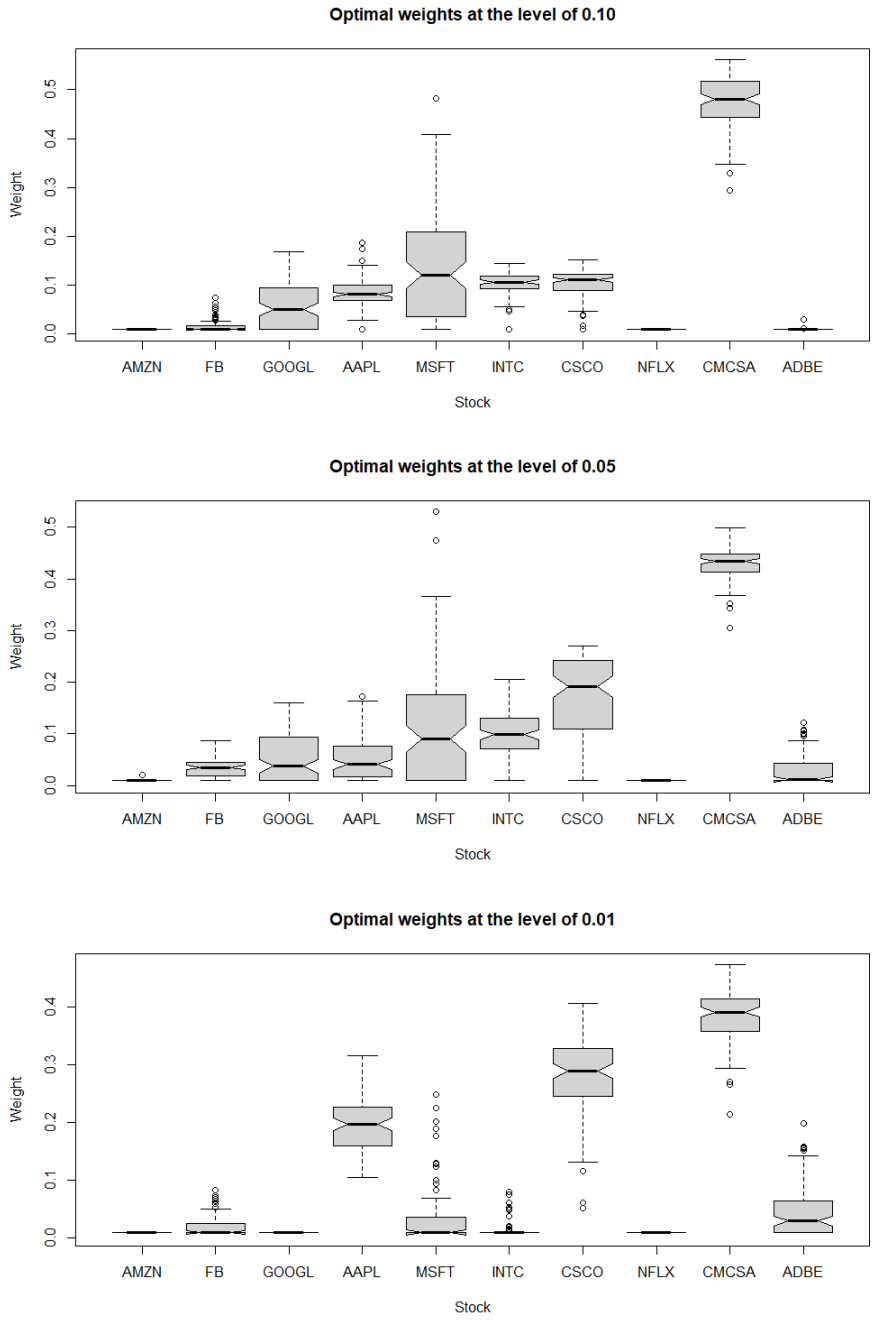
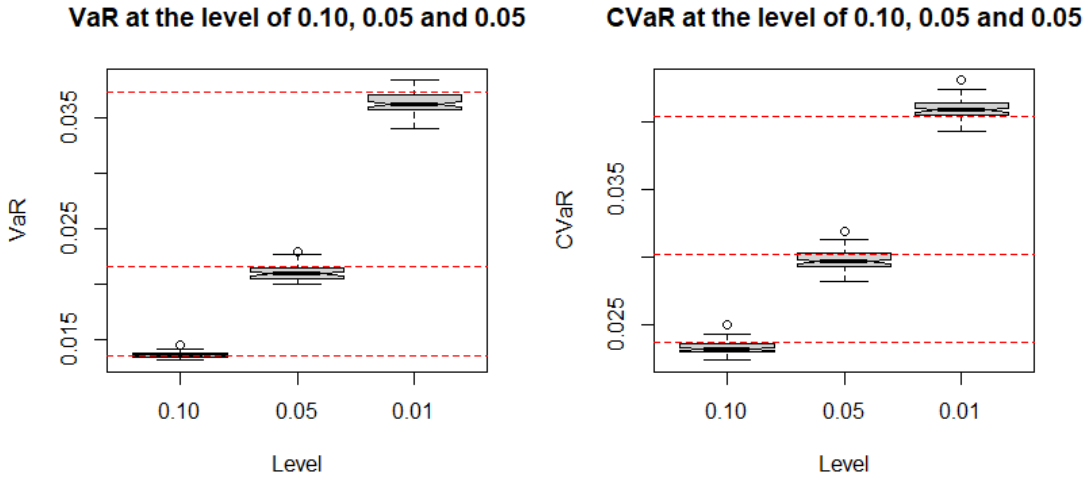


Figure 3.8 Distribution of optimal weights obtained from simulated data using the estimated copula at the level of 0.10, 0.05, and 0.01, for a portfolio of $d = 10$ stocks



(a) **(b)**

Figure 3.9 Distribution of VaR and CVaR obtained from simulated data using the estimated copula at the level of 0.10, 0.05, and 0.01 for a portfolio of $d = 10$ stocks. Dashed lines indicate empirical estimates using historical data.

[2018], etc.).

3.6 Technical Details

3.6.1 Proof of Theorem 3.1

Proof. We denote the ECBC as $B_m(C_n^{\#})$ for simplicity. Also, the empirical Bernstein copula and the Bernstein copula are denoted as $B_m(C_n)$ and $B_m(C)$, respectively. Let $||g|| = \sup_{\mathbf{u}\in[0,1]^d} g(\mathbf{u})$ denote the supremum norm of a function $g(\cdot)$ defined on d-dimensional square $[0,1]^d$. Using the familiar triangle inequality we have

$$||B_m(C_n^{\#}) - C|| \leq ||B_m(C_n^{\#}) - B_m(C_n)|| + ||B_m(C_n) - B_m(C)|| + ||B_m(C) - C||$$

First, under the assumption that the marginal CDFs are continuous, it follows from the Remark 2 in Genest et al. [2017] that

$$||C_n^{\#} - C_n|| \leq \frac{d}{n}.$$

Next, notice that

$$\begin{aligned}
&||B_m(C_n^{\#}) - B_m(C_n)|| \\
&= \sup_{\mathbf{u}\in[0,1]^d} \left| \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} \left(C_n^{\#}\left(\frac{k_1}{m_1}, \ldots, \frac{k_d}{m_d}\right) - C_n\left(\frac{k_1}{m_1}, \ldots, \frac{k_d}{m_d}\right)\right) \prod_{j=1}^{d} \binom{m_j}{k_j} u_j^{k_j} (1-u_j)^{m_j-k_j} \right| \\
&\leq \max_{0\leq k_1\leq m_1, \ldots 0\leq k_d\leq m_d} \left| C_n^{\#}\left(\frac{k_1}{m_1}, \ldots, \frac{k_d}{m_d}\right) - C_n\left(\frac{k_1}{m_1}, \ldots, \frac{k_d}{m_d}\right)\right| \\
&\leq ||C_n^{\#} - C_n|| \leq \frac{d}{n}
\end{aligned}$$

In above the second inequality follows from the fact that since $\binom{m_j}{k_j} u_j^{k_j}(1-u_j)^{m_j-k_j}, k_j = 0, 1, \ldots, m_j$ are binomial probabilities, $\sum_{k_j=0}^{m_j} \binom{m_j}{k_j} u_j^{k_j}(1-u_j)^{m_j-k_j} = 1$ for any $u_j \in [0,1]$ and for any $j = 1, \ldots, d$.

Next by using the Lemma 1 in Janssen et al. [2012] and equation (3) in Kiriliouk et al. [2019], we obtain

$$\begin{aligned}
||C_n - C|| &\leq ||C_n - F_n(F_{n1}^{-1}(u_1), \ldots, F_{nd}^{-1}(u_d))|| \\
&\quad + ||F_n(F_{n1}^{-1}(u_1), \ldots, F_{nd}^{-1}(u_d)) - F(F_1^{-1}(u_1), \ldots, F_d^{-1} d(u_d)|| \\
&\leq \frac{d}{n} + O(n^{-1/2}(\log\log n)^{1/2}) \quad a.s. \\
&= O(n^{-1/2}(\log\log n)^{1/2}) \quad a.s..
\end{aligned}$$

Hence, it now follows that

$$||B_m(C_n) - B_m(C)|| \leq ||C_n - C|| \leq O(n^{-1/2}(\log\log n)^{1/2}), \quad a.s.$$

Also, by using Lemma 3.2 in Segers et al. [2017] we have

$$||B_m(C)-C|| \leq \sum_{j=1}^{d} \frac{1}{2\sqrt{m_j}}$$

Thus, combining the above inequalities that are satisfied almost surely (a.s.) for every fixed m_j's we obtain

$$\begin{aligned} ||B_m(C_n^{\#})-C|| &\leq ||B_m(C_n^{\#})-B_m(C_n)|| + ||B_m(C_n)-B_m(C)|| + ||B_m(C)-C|| \\ &\leq \frac{d}{n} + \sum_{j=1}^{d} \frac{1}{2\sqrt{m_j}} + O(n^{-1/2}(\log\log n)^{1/2}) \;\; a.s. \end{aligned}$$

Next we consider the proposed empirical priors on the degrees $m_1,\ldots,m_d$ to be

$$m_j|\alpha_j \overset{ind}{\sim} Poisson(n^{\alpha_j})+1 \text{ and } \alpha_j \overset{iid}{\sim} Unif\left(\frac{1}{3},\frac{2}{3}\right) \text{ for } j=1,\ldots,d.$$

We make use of the following simple Lemma:

Lemma: Suppose $M \sim Poisson(\lambda)$, *then* $E[\frac{1}{\sqrt{M+1}}] \leq \sqrt{\frac{1-e^{-\lambda}}{\lambda}}$

Proof: By Jensen's inequality for the square-root function, it follows that

$$E\left[\frac{1}{\sqrt{M+1}}\right] \leq \sqrt{E\left[\frac{1}{M+1}\right]} = \sqrt{\frac{1}{\lambda}\sum_{m=0}^{\infty}\frac{\lambda^{m+1}e^{-\lambda}}{(m+1)!}} = \sqrt{\frac{1-e^{-\lambda}}{\lambda}}$$

Notice that as $\alpha_j \sim Unif(\frac{1}{3},\frac{2}{3})$, $\Pr(\alpha_j > \frac{1}{3}) = 1$ and conditioning on α_j, we then have by the above Lemma, $E\left(\sqrt{\frac{1}{m_j}}|\alpha_j\right) \leq \sqrt{\frac{1-e^{-n^{\alpha_j}}}{n^{\alpha_j}}} \to 0$ as $n \to \infty$. Thus, taking expectation with respect to the prior distribution, it follows that

$$\begin{aligned} E(||B_m(C_n^{\#})-C||) &\leq E(||B_m(C_n^{\#})-B_m(C_n)||) + E(||B_m(C_n)-B_m(C)||) + E(||B_m(C)-C||) \\ &\leq \frac{d}{n} + \sum_{j=1}^{d}\frac{1}{2}E\left(\frac{1}{\sqrt{m_j}}\right) + O(n^{-1/2}(\log\log n)^{1/2}) \to 0 \; as \; n \to \infty \;\; a.s. \end{aligned}$$

□

3.6.2 Choice of M

Proof. By the Law of the Iterated Logarithm (see e.g. Balsubramani [2014]), the deviation of MC approximation from the mean is almost surely bounded by

$$\sqrt{\frac{2\log\log M}{M}}\sqrt{Var((-\mathbf{v}^T\mathbf{x}_k-\gamma)_+)}$$

Let $\Sigma = Var_F(\mathbf{X})$ and λ_{max} be the largest eigenvalue of Σ, then we have

$$Var((-\mathbf{v}^T\mathbf{X}-\gamma)_+) \leq Var(\mathbf{v}^T\mathbf{X}) = \mathbf{v}^T\Sigma\mathbf{v} \leq \lambda_{max}\mathbf{v}^T\mathbf{v} \leq \lambda_{max},$$

for any $\mathbf{v} \in \mathscr{S}_d$ because $\mathbf{v}^T\mathbf{v} \leq \mathbf{v}^T\mathbf{1} = 1$. Thus, for an accuracy of $\epsilon > 0$ for the MC approximation, it is sufficient to generate M MC samples such that

$$\sqrt{\frac{2\log\log M}{M}\lambda_{max}} \leq \epsilon \;\; i.e. \;\; \frac{M}{\log\log M} \geq \frac{2\lambda_{max}}{\epsilon^2}$$

□

CHAPTER

4

ECBC-BASED NONPARAMETRIC ESTIMATION OF CONDITIONAL COPULA

4.1 Introduction

Copulas have found many applications in the field of finance, insurance, system reliability, etc., owing to its utility in modeling the dependence among variables (see, e.g., Nelsen [2007], Jaworski et al. [2010] and Joe [2014] for details about copulas and their applications). In some situations, the dependence structure between variables can be influenced by a set of covariates and it is thereby of interest to understand how such dependence changes with the values of covariates. For instance, it is well known that the life expectancy at birth of males and females in a country are often highly dependent due to shared economic or

environmental factors and it is possible that the strength of dependence relies on these factors. When the covariate is binary or discrete-valued with few levels, one can estimate a copula for each given level of the discrete-valued covariate separately. In constrast, the influence of continuous-value covariate on the dependence structure should be formulated in a functional way, and this is where conditional copulas (Patton [2006b]; Patton [2006a]) along with the corresponding conditional versions of dependence measures come into play.

Suppose we are interested in the dependence among the components of a random vector $\mathbf{Y} = (Y_1, Y_2, ..., Y_d)$, given covariates $\mathbf{X} = (X_1, X_2, ..., X_p)$. The conditional joint and marginal distribution of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$ can be denoted as

$$F_{\mathbf{x}}(\mathbf{y}) \equiv F_{\mathbf{x}}(y_1, y_2, ...y_d) = P(Y_1 \leq y_1, Y_2 \leq y_2, ..., Y_d \leq y_d \mid \mathbf{X} = \mathbf{x}), \tag{4.1}$$

and

$$F_{j\mathbf{x}}(y_j) = P(Y_j \leq y_j \mid \mathbf{X} = \mathbf{x}) \;\; j = 1, ..., d. \tag{4.2}$$

If $F_{1\mathbf{x}}, F_{2\mathbf{x}}, ..., F_{d\mathbf{x}}$ are continuous, then by an extension of the well-known Sklar's theorem (Sklar [1959]) for conditional distributions (e.g. see Patton [2006a]), there exists a unique copula $C_{\mathbf{x}}$ such that

$$F_{\mathbf{x}}(\mathbf{y}) = C_{\mathbf{x}}(F_{1\mathbf{x}}(y_1), F_{2\mathbf{x}}(y_2), ..., F_{d\mathbf{x}}(y_d)) \;\; \forall y \in \mathbb{R}^d, \forall y \in \mathbb{R}^d, \tag{4.3}$$

and the function $C_{\mathbf{x}}$ is called a conditional copula, which captures the conditional dependence structure of $\mathbf{Y}$ given $\mathbf{X} = \mathbf{x}$. The focus of this chapter is modeling continuous-valued responses and covariates. Thus, in what follows, we assume that the conditional marginal CDFs $F_{1\mathbf{x}}, j = 1, \ldots, d$ and the CDFs of each response and covariate are absolutely continuous.

The literature contains a variety of parametric families for modeling copulas. Some

commonly used copula families are Archimedean copulas, elliptical copulas, etc.; see Žežula [2009] and Joe [2014], etc. Assuming that the conditional copula belongs to a parametric copula family where the copula parameter is a function of the covariate(s), there has been previous work addressing the estimation of conditional copula in a semiparametric setting. In regard to frequentist methods based on an assumed parametric class, Acar et al. [2011] propose to estimate the functional relationship between the copula parameter and the covariate nonparametrically by using local likelihood approach. But they assume known marginals and the maximization is conducted for a fixed value of the covariate, i.e., With the intention of identifying the entire function between the copula parameter and the covariate, it is necessary to solve the maximization problem for a sufficiently large grid of values within the range of the covariate. Abegaz et al. [2012] extend the work to a more general setting of unknown marginals and apply a two-stage technique that has been widely adopted in copula estimation: in the first stage, the nonparametric estimates of conditional marginals are obtained using kernel-based method and by plugging in theses estimates, the functional link is estimated by maximizing the pseudo log-likelihood in the second stage. As alternative estimation methods for the function relationship, Vatter & Chavez-Demoulin [2015] develop generalized additive models for conditional dependence structures, and Fermanian & Lopez [2018] introduce so-called single-index copulas, etc. In the Bayesian framework, inference for bivariate conditional copula models have been constructed in Craiu & Sabeti [2012], Sabeti et al. [2014] and Levi & Craiu [2018], among others.

However, the misspecification of copula family could lead to severely biased estimation even though a sophisticated and flexible parametric model is employed (e.g. see Geerdens et al. [2018]), so it is required to select an appropriate copula model from a large number of candidate families. In order to do so, many copula selection techniques have been proposed either in either frequentist or Bayesian setting, e.g., Acar et al. [2011] select the copula family based on cross-validated prediction errors, while the deviance information criterion (DIC) is utilized for the choice of copula in Craiu & Sabeti [2012].

Acknowledging the limitations of parametric copula models as mentioned above, fully nonparametric approaches have also been proposed for conditional copula estimation. Gijbels et al. [2011] suggests the empirical estimators for conditional copulas where the weights are smoothed over the covariate space through kernel-based methods. They further derive nonparametric estimates for the conditional dependence measures including conditional Kendall's tau and conditional Spearman's rho. Since the bandwidth selection is very crucial for any of the smoothing methods, they also develop an algorithm for selecting the bandwidths. The asymptotic properties of the estimators together with conditional dependence measure estimates are established in Veraverbeke et al. [2011]. Gijbels et al. [2012] further consider more complex covariates like multivariate covariates and box-type conditioning events are studied in Derumigny & Fermanian [2020]. On the other hand, there has been recent work on Bayesian nonparametric estimation of conditional copula. Leisen et al. [2017] introduce the effect of a covariate to the Bayesian infinite mixture models proposed by [Wu et al., 2014]. However, large-sample asymptotic properties of the Bayesian models have been almost unexplored and still remains an area of open work.

In this chapter, we focus on the nonparametric estimation of conditional copulas and have realized that it can be done in a relatively easy way by employing the empirical checkerboard Bernstein copula (ECBC) estimator. When the covariates are continuous-valued, the main idea of extending the copula models to include covariates is to first estimate the full copula of responses along with covariates and then take partial derivatives to obtain the conditional distribution of responses given the covariates. As a fully nonparametric approach, it is not required to make any selection of the proper copula family, which is a key step in semiparametric methods to avoid the adverse consequence of model misspecification. Compared to the kernel-based empirical estimators, the selection of bandwidths is unnecessary either, making it easy to implement in practice. The proposed ECBC-based conditional copula estimator immediately leads to nonparametric estimates of the conditional dependence measures, which can be expressed in a very neat form under matrix

operations. The large-sample consistency of the proposed estimator is also provided in the chapter.

4.2 Models for Conditional Copula

In the following, we focus on the bivariate conditional copula of (Y_1, Y_2) with a single covariate X for simplicity. Notice that the extension to more than two dimensions and multiple covariates is straightforward.

Suppose we have i.i.d. samples $(Y_{i1}, Y_{i2}, X_i), i = 1, ..., n$, where $(Y_{i1}, Y_{i2}), i = 1, ..., n$ are i.i.d observations of the random vector (Y_1, Y_2) of which the conditional dependence structure is of our interest. $X_i, i = 1, ..., n$ are i.i.d. observations of the covariate X. We assume all components of (Y_1, Y_2, X) are continuous-valued random variables with absolutely continuous marginal distributions and the conditional marginal distributions of Y_1 and Y_2 given $X = x$ are also absolutely continuous. The goal is to estimate the conditional copula C_x from a random sample of i.i.d. observations $(Y_{i1}, Y_{i2}, X_i), i = 1, ..., n$.

As suggested by Gijbels et al. [2011], it is often favorable to remove the effect of the covariate on the marginal distributions before estimating C_x. In order to do that, we can transform the original observations (Y_{i1}, Y_{i2}) to marginally uniformly distributed (unobserved) samples

$$(U_{i1}, U_{i2}) \equiv (F_{1X_i}(Y_{i1}), F_{2X_i}(Y_{i2})), \quad i = 1, ..., n, \tag{4.4}$$

which can be estimated by pseudo-observations

$$(\hat{U}_{i1}, \hat{U}_{i2}) \equiv (\hat{F}_{1X_i}(Y_{i1}), \hat{F}_{2X_i}(Y_{i2})), \quad i = 1, ..., n, \tag{4.5}$$

where $\hat{F}_{1x}$ and $\hat{F}_{2x}$ are the estimated conditional marginal distributions.

Motivated by Janssen et al. [2016] who apply the empirical Bernstein estimator of bivariate copula derivative to conditional distribution estimation with a single covari-

ate, we are able to use the proposed multivariate copula estimator ECBC to estimate the conditional marginal distributions of Y_1 and Y_2 given $X = x$, respectively. Specifically, for $j \in \{1,2\}$, we have i.i.d samples $(Y_{ij}, X_i), i = 1,...,n$ and the corresponding pseudo-observations $(\hat{W}_{ij}, \hat{V}_i) = (F_{nY_j}(Y_{ij}), F_{nX}(X_i)), i = 1,...,n$, where F_{nY_j} and F_{nX} are the modified empirical estimation of the (unconditional) marginal distributions F_{Y_j} and F_X, respectively, e.g., $F_{nX}(x) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$. These pseudo-observations can be then treated as samples from a 2-dimensional copula C_j, which can be estimated by the ECBC copula estimator as follows

$$C_j^{\#}(w_j, v) = \sum_{h=0}^{g_j}\sum_{k=0}^{m_j} \tilde{\theta}_{h,k} \binom{g_j}{h} w_j^h (1-w_j)^{g_j-h} \binom{m_j}{k} v^k (1-v)^{m_j-k}, \tag{4.6}$$

where

$$\tilde{\theta}_{h,k} = C_{jn}^{\#}(\frac{h}{g_j}, \frac{k}{m_j}), \tag{4.7}$$

and $C_{jn}^{\#}$ is the empirical checkerboard copula. Then the partial derivative $C_j^{(1)}$ of C_j with respect to v can be estimated by using

$$\begin{aligned} C_j^{\#(1)}(w_j, v) &\equiv \frac{\partial C_j^{\#}(w_j, v)}{\partial v} \\ &= \sum_{h=0}^{g_j}\sum_{k=0}^{m_j-1} \tilde{\lambda}_{h,k} \binom{g_j}{h} w_j^{h_j} (1-w_j)^{g_j-h} m_j \binom{m_j-1}{k} v^k (1-v)^{m_j-k-1}, \end{aligned} \tag{4.8}$$

where

$$\tilde{\lambda}_{h,k} = \tilde{\theta}_{h,k+1} - \tilde{\theta}_{h,k}. \tag{4.9}$$

Notice that the following relationship holds between the conditional marginal distribution function of Y_j given $X = x$ and the partial derivative $C_j^{(1)}(w_j, v)$

$$F_{jx}(y_j) = P(Y_j \leq y_j \mid X = x) = C_j^{(1)}(F_{Y_j}(y_j), F_X(x)). \tag{4.10}$$

Thus, we can estimate the conditional marginal distributions using

$$\hat{F}_{jx}(y_j) = C_j^{\#(1)}(F_{nY_j}(y_j), F_{nX}(x)) \tag{4.11}$$

for $j = 1,2$, and then the corresponding pseudo-observations $(\hat{U}_{i1}, \hat{U}_{i2}), i = 1,...,n$ of the conditional copula C_x adjusted for the effect of the covariate on the marginal distributions can be estimated as given in (4.5).

Now we can use the covariate-adjusted pseudo-observations $(\hat{U}_{i1}, \hat{U}_{i2}), i = 1,...,n$ along with the pseudo-observations of the covariate $\hat{V}_i, i = 1,...,n$ to estimate a 3-dimensional copula $C(u_1, u_2, v)$ again using ECBC and denote it as $C^{\#}(u_1, u_2, v)$. Similar to (4.8), it is easy to obtain the partial derivative $C^{\#(1)}$ of $C^{\#}$ with respect to v, which is denoted as

$$\begin{aligned} C^{\#(1)}(u_1, u_2|v) &\equiv \frac{\partial C^{\#}(u_1, u_2, v)}{\partial v} \\ &= \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m-1} \tilde{\gamma}_{h_1,h_2,k} m \binom{m-1}{k} v^k (1-v)^{m-k-1} \prod_{s=1}^{2} \binom{l_s}{h_s} u_s^{h_s} (1-u_s)^{l_s-h_s}, \end{aligned} \tag{4.12}$$

where

$$\tilde{\gamma}_{h_1,h_2,k} = \tilde{\theta}_{h_1,h_2,k+1} - \tilde{\theta}_{h_1,h_2,k}. \tag{4.13}$$

Notice that we can use $C^{\#(1)}(u_1, u_2|F_{nX}(x))$ as an estimate of the conditional copula C_x, however, $C^{\#(1)}(u_1, u_2|v)$ is itself a valid bivariate copula for any value of $v \in [0,1]$ only asymptotically. This is because the conditional marginal distributions of $C^{\#(1)}(u_1, u_2|v)$ are not necessarily uniform distributions for finite samples. Aiming to obtain a more accurate estimate of the conditional copula for small samples, we consider the conditional marginal distributions of $C^{\#(1)}(u_1, u_2|v)$ given as

$$\begin{aligned} F_1(u_1|v) &\equiv C^{\#(1)}(u_1, 1|v) \\ &= \sum_{h_1=0}^{l_1} \sum_{k=0}^{m-1} \tilde{\gamma}_{h_1,l_2,k} m \binom{m-1}{k} v^k (1-v)^{m-k-1} \binom{l_1}{h_1} u_1^{h_1} (1-u_1)^{l_1-h_1}, \end{aligned} \tag{4.14}$$

and

$$F_2(u_2|v) \equiv C^{\#(1)}(1, u_2|v) = \sum_{h_2=0}^{l_2} \sum_{k=0}^{m-1} \tilde{\gamma}_{l_1,h_2,k} m \binom{m-1}{k} v^k (1-v)^{m-k-1} \binom{l_2}{h_2} u_2^{h_2} (1-u_2)^{l_2-h_2}, \tag{4.15}$$

By using Sklar's theorem, we are able to obtain a conditional copula estimator which is a genuine copula itself denoted as

$$C^{\#}(u_1, u_2|v) = C^{\#(1)}(F_1^{-1}(u_1|v), F_2^{-1}(u_2|v)|v), \tag{4.16}$$

where $F_1^{-1}(u_1|v)$ and $F_2^{-1}(u_2|v)$ are the inverse functions of F_1 and F_2, respectively. It is to be noted that $C^{\#}(u_1, u_2|v)$ is a valid copula for any value of $v \in [0,1]$, and as a result, the conditional copula C_x can be estimated by

$$C_x^{\#}(u_1, u_2) = P(F_{1x}(y_1) \leq u_1, F_{2x}(y_2) \leq u_2 \mid X = x) = C^{\#}(u_1, u_2|F_{nX}(x)). \tag{4.17}$$

Let $||g||(v) = \sup_{(u_1,u_2)\in[0,1]^2} |g(u_1, u_2|v)|$ denote the conditional supremum norm of a conditional function $g(u_1, u_2|v)$ defined on the unit square $[0,1]^2$ for a fixed v. We denote the common supremum norm as $||\cdot||$. The following theorem provides the large-sample consistency of the estimator $C^{\#}(u_1, u_2|v)$ for fixed value of $0 < v < 1$ using the conditional supremum norm.

Theorem 4.1. *Assume that the underlying trivariate copula $C(u_1, u_2, v)$ is absolutely continuous and conditional copula $C_v(u_1, u_2|v) = \frac{\partial C(u_1,u_2,v)}{\partial v}$ is Lipschitz continuous on $[0,1]^3$. Then for any fixed $0 < v < 1$, we have*

$$E(||C^{\#} - C_v||(v)) \xrightarrow{a.s.} 0 \ \ as \ \ n \to \infty.$$

where the expectation is taken with respect to the empirical prior distribution of l_1, l_2, and m

as given for ECBC.

Remark: Following the hierarchical shifted Poisson distributions proposed for ECBC in Chapter 3, the empirical prior distribution of l_1, l_2, and m are given as

$$l_j|\alpha_j \sim Poisson(n^{\alpha_j})+1 \text{ and } \alpha_j \sim Unif\Big(\frac{1}{3},\frac{2}{3}\Big)\ j=1,2,$$

$$m|\alpha \sim Poisson(n^{\alpha})+2 \text{ and } \alpha \sim Unif\Big(\frac{1}{3},\frac{2}{3}\Big).$$

Next, by extending the dependence measures given in Schweizer & Wolff [1981] to conditional versions, we are able to estimate the conditional dependence measures (e.g. conditional Spearman's rho, conditional Kendall's tau, etc.) using the estimator $C^{\#(1)}(u_1,u_2|v)$. For instance, the estimate of conditional Kendall's tau takes the form

$$\hat{\tau}(v)=4\int_0^1\int_0^1 C^{\#(1)}(u_1,u_2|v)dC^{\#(1)}(u_1,u_2|v)-1, \tag{4.18}$$

and the estimate of conditional Spearman's rho is given as

$$\hat{\rho}(v)=12\int_0^1\int_0^1 \Big(C^{\#(1)}(u_1,u_2|v)-F_1(u_1|v)F_2(u_2|v)\Big)dF_1(u_1|v)dF_2(u_2|v). \tag{4.19}$$

Let us denote

$$\eta_{h_1,h_2|v}\equiv m\sum_{k=0}^{m-1}\tilde{\gamma}_{h_1,h_2,k}\binom{m-1}{k}v^k(1-v)^{m-k-1},\ \ h_1=0,...,l_1,h_2=0,...,l_2. \tag{4.20}$$

Then we can rewrite the estimator $C^{\#(1)}(u_1,u_2|v)$ and its conditional marginal distributions as

$$C^{\#(1)}(u_1,u_2|v)=\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\eta_{h_1,h_2|v}\prod_{s=1}^{2}\binom{l_s}{h_s}u_s^{h_s}(1-u_s)^{l_s-h_s}, \tag{4.21}$$

$$F_1(u_1|v)=\sum_{h_1=0}^{l_1}\eta_{h_1,l_2|v}\binom{l_1}{h_1}u_1^{h_1}(1-u_1)^{l_1-h_1}, \tag{4.22}$$

and

$$F_2(u_2|v)=\sum_{h_2=0}^{l_2}\eta_{l_1,h_2|v}\binom{l_2}{h_2}u_2^{h_2}(1-u_2)^{l_2-h_2}, \tag{4.23}$$

respectively. As a result, a closed-form estimate of conditional Kendall's tau takes the form

$$\begin{aligned}\hat{\tau}(v)=4\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{g_1=0}^{l_1-1}\sum_{g_2=0}^{l_2-1}\eta_{h_1,h_2|v}(\eta_{g_1+1,g_2+1|v}-\eta_{g_1+1,g_2|v}-\eta_{g_1,g_2+1|v}+\eta_{g_1,g_2|v})\\ \prod_{s=1}^{2}l_s\binom{l_s}{h_s}\binom{l_s-1}{g_s}B(h_s+g_s+1,2l_s-h_s-g_s)-1,\end{aligned} \tag{4.24}$$

where B is the beta function. Similarly, we are able to obtain a closed-form estimate of conditional Spearman's rho as

$$\begin{aligned}\hat{\rho}(v)=12\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{g_1=0}^{l_1-1}\sum_{g_2=0}^{l_2-1}(\eta_{h_1,h_2|v}-\eta_{h_1,l_2|v}\eta_{l_1,h_2|v})(\eta_{g_1+1,l_2|v}-\eta_{g_1,l_2|v})(\eta_{l_1,g_2+1|v}-\eta_{l_1,g_2|v})\\ \prod_{s=1}^{2}l_s\binom{l_s}{h_s}\binom{l_s-1}{g_s}B(h_s+g_s+1,2l_s-h_s-g_s).\end{aligned} \tag{4.25}$$

For the purpose of computing the estimates of conditional dependence measures more efficiently, we apply matrix operations to the tensor products in expressions (4.24) and (4.25). For given $(h_1,h_2), h_1=1,\ldots,l_1, h_2=1,\ldots,l_2$, let us denote

$$a_{h_1,g_1}=l_1\binom{l_1}{h_1}\binom{l_1-1}{g_1}B(h_1+g_1+1,2l_1-h_1-g_1),\ \ g_1=0,\ldots l_1-1. \tag{4.26}$$

and

$$b_{h_2,g_2}=l_2\binom{l_2}{h_2}\binom{l_2-1}{g_2}B(h_2+g_2+1,2l_2-h_2-g_2),\ \ g_2=0,\ldots l_2-1. \tag{4.27}$$

Then we have $\mathbf{a}_{h_1}=(a_{h_1,0},\ldots,a_{h_1,l_1-1})^T$ and $\mathbf{b}_{h_2}=(b_{h_2,0},\ldots,b_{h_2,l_2-1})^T$. We also denote a $l_1\times l_2$ matrix $\mathbf{D}_v=(d_{g_1,g_2|v})_{l_1\times l_2}$ where $d_{g_1,g_2|v}=\eta_{g_1+1,g_2+1|v}-\eta_{g_1+1,g_2|v}-\eta_{g_1,g_2+1|v}+\eta_{g_1,g_2|v}$. Thus, the estimate of conditional Kendall's tau given in (4.29) can be rewritten as

$$\hat{\tau}(v)=4\sum_{h_1=1}^{l_1}\sum_{h_2=1}^{l_2}\eta_{h_1,h_2|v}\mathbf{a}_{h_1}^T\mathbf{D}_v\mathbf{b}_{h_2}-1. \tag{4.28}$$

Furthermore, we can denote two $l_1 \times l_2$ matrices, $\mathbf{H}_v = (\eta_{h_1,h_2|v})_{l_1 \times l_2}$ and $\mathbf{G}_v = (\mathbf{a}_{h_1}^T \mathbf{D}_v \mathbf{b}_{h_2})_{l_1 \times l_2}$, an as a result, we have

$$\hat{\tau}(v) = 4\text{Tr}(\mathbf{H}_v^T \mathbf{G}_v) - 1. \tag{4.29}$$

Similarly, we are able to rewrite the estimate of conditional Spearman's rho given in (4.25). Let us first denote two vectors, $\mathbf{p}_v = (p_{g_1|v})_{l_1}^T$ where $p_{g_1|v} = \eta_{g_1+1,l_2|v} - \eta_{g_1,l_2|v}, g_1 = 0,\ldots,l_1-1$ and $\mathbf{q}_v = (q_{g_2|v})_{l_2}^T$ where $q_{g_2|v} = \eta_{l_1,g_2+1|v} - \eta_{l_1,g_2|v}, g_2 = 0,\ldots,l_2-1$. Then we have

$$\hat{\rho}(v) = 12 \sum_{h_1=1}^{l_1} \sum_{h_2=1}^{l_2} (\eta_{h_1,h_2|v} - \eta_{h_1,l_2|v}\eta_{l_1,h_2|v}) \mathbf{a}_{h_1}^T (\mathbf{p}_v \otimes \mathbf{q}_v) \mathbf{b}_{h_2}. \tag{4.30}$$

If we further denote two $l_1 \times l_2$ matrices, $\mathbf{R}_v = (r_{h_1,h_2|v})_{l_1 \times l_2}$ where $r_{h_1,h_2|v} = \eta_{h_1,h_2|v} - \eta_{h_1,l_2|v}\eta_{l_1,h_2|v}$ and $\mathbf{J}_v = (\mathbf{a}_{h_1}^T (\mathbf{p}_v \otimes \mathbf{q}_v) \mathbf{b}_{h_2})_{l_1 \times l_2}$, then we have

$$\hat{\rho}(v) = 12\text{Tr}(\mathbf{R}_v^T \mathbf{J}_v) \tag{4.31}$$

By applying the above matrix operations, we are able to obtain very neat expressions of the estimates of conditional dependence measures and the computational efficiency can be improved significantly.

4.3 Numerical Illustrations using Simulated Data

We now show the finite-sample performance of the conditional copula estimator $C_x^{\#}(u_1, u_2)$. Similar to the simulation setup in Acar et al. [2011], data $(U_{i1}, U_{i2}|X_i), i = 1, ..., n$ are generated from the Clayton copula using the package `copula` in `R` under the following models: $(U_{i1}, U_{i2})|X_i \sim C(u_1, u_2|\theta_i)$, where $\theta_i = \exp(0.8X_i - 2)$ and $X_i \sim Unif(0,3)$. The true copula parameter varies from 0.14 to 1.49 with Spearman's rho ranging from 0.10 to 0.60. The pseudo-observations of the covariate are defined as $V_i \equiv F_{nX}(X_i), i = 1,\ldots,n$, where $F_{nX}(x) = \frac{1}{n+1}\sum_{i=1}^n \mathbb{I}(X_i \le x)$. $N = 100$ replicates are drawn from the true copula with sample

size $n = 200$.

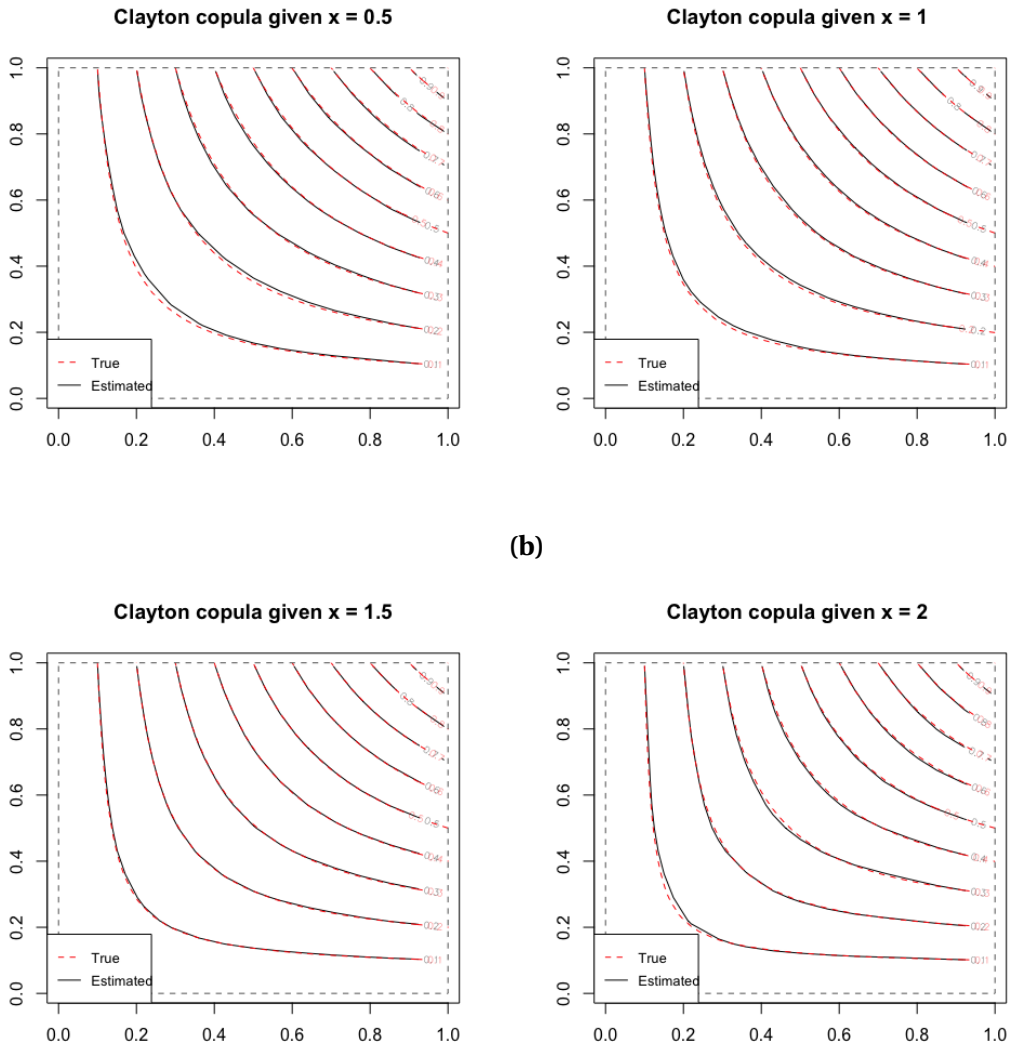
Figure 4.1 shows the contour plots of the Monte Carlo average of the estimated $C_x^{\#}(u_1, u_2)$ given $x = 0.5$, $x = 1$, $x = 1.5$, and $x = 2$, respectively, across 100 Monte Carlo replicates. The contour plots are drawn based on a 15×15 equally spaced grid of points in the unit square, meaning that for a given v we need to find $15 + 15 = 30$ roots. Since F_1 and F_2 are both non-decreasing functions, we can calculate the inverse functions $F_1^{-1}(u_1|v)$ and $F_2^{-1}(u_2|v)$ by applying the function `uniroot` in `R` to equations (4.22) and (4.23) for a given value of v. The true copula parameters are 0.20 (Spearman's rho equal to 0.09), 0.30 (Spearman's rho equal to 0.20). 0.45 (Spearman's rho equal to 0.27), and 0.67 (Spearman's rho equal to 0.37) for $x = 0.5$, $x = 1$, $x = 1.5$, and $x = 2$, respectively.

It can be observed from the plots that all the estimated contour lines overlap with the true lines at the boundaries, which is a evidence that the conditional copula estimator $C^{\#}(u_1, u_2|v)$ is a genuine copula with uniform conditional marginal distributions. Moreover, there is almost no bias between the estimated conditional copula averaged over 100 Monte Carlo samples and the true conditional copula across different values of the covariate, illustrating the proposed ECBC-based method works well in estimating conditional copula.

Then we can plot the conditional Kendall's tau and conditional Spearman's rho as given in (4.29) and (4.31) as a function of the covariate. The covariate x ranges from 0 to 3 so we compute the dependence measures at seven different values $(0.05, 0.5, 1, 1.5, 2, 2.5, 2.95)$. The following plots show the Monte Carlo average of estimates of dependence measures and the 90% Monte Carlo confidence bands ($5th$ and $95th$ percentiles of the dependence measure estimates) across 100 Monte Carlo replicates.

Overall, the estimates averaged over 100 Monte Carlo samples seem to be fairly close to the true conditional dependence measures across different values of the covariate. The variance tends to increase and the Monte Carlo average tends to underestimate a little bit when it gets closer to the boundaries of the covariate.

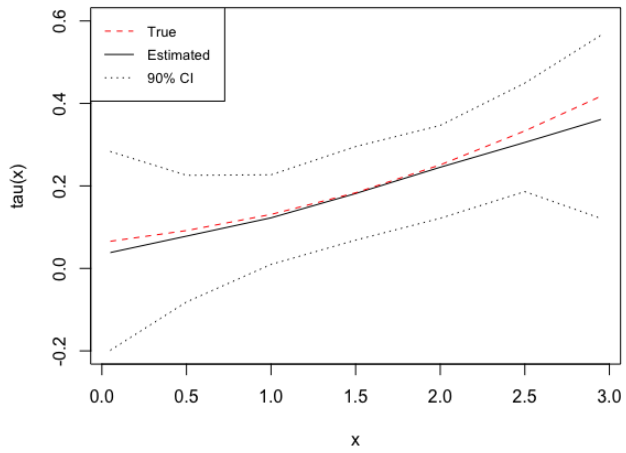
Next, we would like to compare the performance of our proposed nonparametric



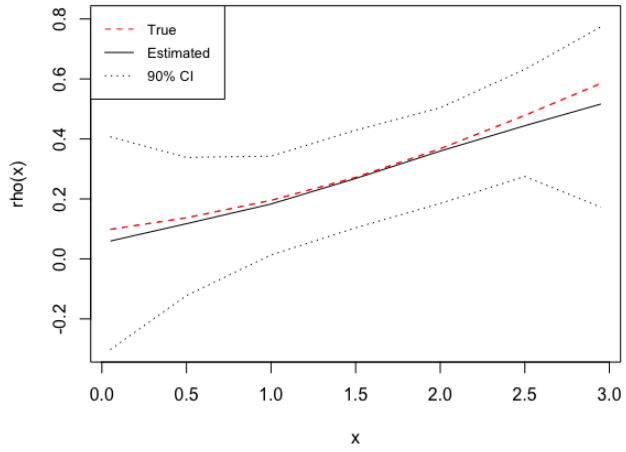
(a) **(b)**

(c) **(d)**

Figure 4.1 The contour plots of the Monte Carlo average of the estimated $C_x^{\#}(u_1, u_2)$ given $x = 0.5$, $x = 1$, $x = 1.5$, and $x = 2$, respectively.



(a)



(b)

Figure 4.2 The plots of the estimated conditional Kendall's tau and conditional Spearman's rho as a function of the covariate.

method to the semiparametric method in Acar et al. [2011] through simulation studies. They assume a conditional copula model where the copula function comes from a parametric copula family and the copula parameter is a function of the covariate. Different copula families, e.g., Clayton and Gumbel, were considered and the functional relationship between the copula parameter and the covariate was estimated using a nonparametric local likelihood approach. The severe consequence of misspecified copula model was investigated in Acar et al. [2011] and they proposed a copula selection method based on cross-validated prediction errors. In contrast, the proposed conditional copula estimator is fully nonparametric so there is no need to make any choice of the copula family.

Simulation setups follow Acar et al. [2011]. The data $(U_{i1}, U_{i2}|X_i), i = 1, ..., n$ are generated from the Clayton copula under the following models: $(U_{i1}, U_{i2})|X_i \sim C(u_1, u_2|\theta_i)$, where (i): $\theta_i = \exp(0.8X_i - 2)$ and $X_i \sim Unif(2,5)$; (ii): $\theta_i = \exp(2 - 0.3(X_i - 2)^2$ and $X_i \sim Unif(2,5)$. The sample size is $n = 200$.

The comparison can be done numerically by calculating the conditional Kendall's tau and some performance measures, including the integrated square Bias (IBIAS^2), integrated Variance (IVAR) and integrated mean square error (IMSE) as given in Acar et al. [2011]:

$$
\begin{aligned}
IBIAS^2 &= \int_{[2,5]} \left[E[\hat{\tau}_x(x)] - \tau_x(x)\right]^2 dx = 3\int_{[0,1]} \left[E[\hat{\tau}(v)] - \tau(v)\right]^2 dv, && (4.32)\\
IVAR &= \int_{[2,5]} E\left[[\hat{\tau}_x(x) - E[\hat{\tau}_x(x)]]^2\right] dx = 3\int_{[0,1]} E\left[[\hat{\tau}(v) - E[\hat{\tau}(v)]]^2\right] dv, && (4.33)\\
IMSE &= \int_{[2,5]} E\left[[\hat{\tau}_x(x) - \tau_x(x)]^2\right] dx = 3\int_{[0,1]} E\left[[\hat{\tau}(v) - \tau(v)]^2\right] dv, && (4.34)
\end{aligned}
$$

where the second equality holds because $\tau_x(X) = \tau(F_X(X)) = \tau(V)$ and $X \sim Unif(2,5)$. We compute Monte Carlo estimates of these performance measures by following the tricks in Segers et al. [2017] and compare our proposed method (referred to as "ECBC-based") to the local likelihood method (referred to as "Local") in Acar et al. [2011]. The results are shown in Table 4.1.

From the results we can see that when data are generated from the Clayton copula (the underlying true copula), our ECBC-based method outperforms the local likelihood method for the incorrect parametric case (Gumbel) in terms of bias and MSE, although the performance is not as good as the local likelihood method for the correct parametric case (Clayton). Nonetheless, the advantage of the proposed nonparametric method is that we can avoid the adverse impact of misspecified copula and obtain fairly good estimation of conditional copula and conditional dependence measures without having to select the 'best' copula model from numerous copula families.

Table 4.1 Comparison of the proposed method (referred to as "ECBC-based") to the local likelihood method (referred to as "Local") using Monte Carlo estimates of three performance measures, IBIAS^2, IVAR and IMSE. Data are generated from the Clayton copula under two different functional relationships between the copula parameter and the covariate.

Clayton copula: $\theta = \exp(0.8X - 2)$				
Estimation Method	Parametric Model	IBIAS^2 ($\times 10^{-2}$)	IVAR($\times 10^{-2}$)	IMSE ($\times 10^{-2}$)
Local	Clayton	0.017	0.553	0.570
Local	Gumbel	3.704	1.716	5.389
ECBC-based	N/A	0.323	2.569	2.892
Clayton copula: $\theta = \exp(2 - 0.3(X-4)^2)$				
Estimation Method	Parametric Model	IBIAS^2 ($\times 10^{-2}$)	IVAR($\times 10^{-2}$)	IMSE ($\times 10^{-2}$)
Local	Clayton	0.040	0.288	0.328
Local	Gumbel	4.808	1.301	6.109
ECBC-based	N/A	0.855	1.876	2.731

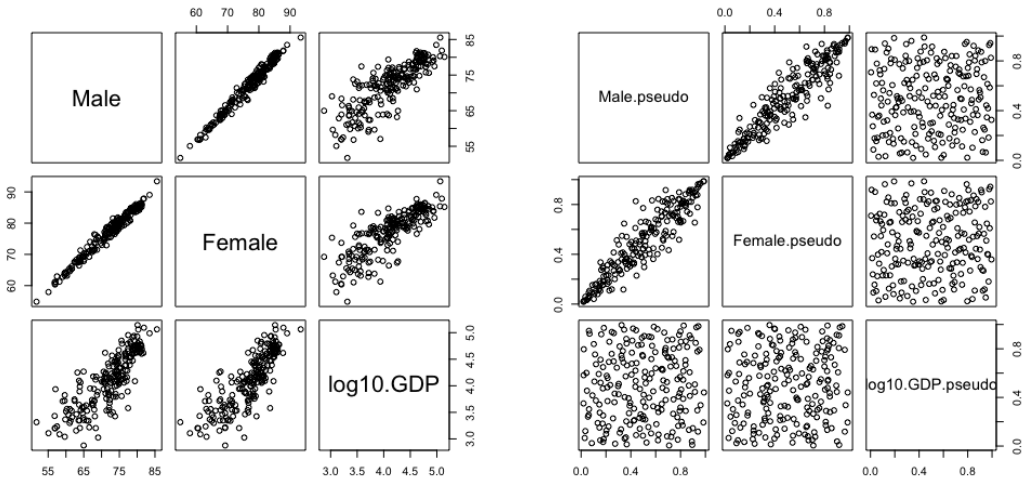
4.4 Real Case Study

We now apply the proposed methodology to a data set of life expectancy at birth of males and females with GDP (in USD) per capita as a covariate for 210 countries or regions. The data are available from the World Factbook 2020 of CIA. Similar data sets were analyzed

in Gijbels et al. [2011] and Abegaz et al. [2012]. Life expectancy at birth summarizes the average number of years to be lived in a country while GDP per capita is often considered as an indicator of a country's standard of living. We are interested in the dependence between the life expectancy at birth of males and females and would like to see if the strength of dependence is influenced by the GDP per capita. In other words, it is of interest to investigate the dependence between the life expectancy at birth of males (Y_1) and females (Y_2) conditioning on the covariate X, where $X = \log_{10}(\text{GDP})$ is $\log_{10}$ transformation of GDP per capita.

The pairwise scatterplots of the data are shown in Figure 4.3(a), from which we can see that there is strong positive correlation between the life expectancy of males (referred to as Male) and females (referred to as Female). Figure 4.3(a) also shows that the life expectancy tends to increase with the $\log_{10}$ transformation of GDP per capita (referred to as log10.GDP) for both males and females. Before estimating the conditional copula of (Y_1, Y_2) given X, we first remove the effect of the covariate X on the marginal distributions of Y_1 and Y_2. As a result, the covariate-adjusted pseudo-observations of Y_1 and Y_2 (referred to as Male.pseudo and Female.pseudo, respectively) and the pseudo-observations of X (referred to ad log10.GDP.pseudo) are given in 4.3(b).

We then estimate the conditional copula and the conditional dependence of life expectancy at birth of males and females given the covariate X. Figure 4.4 shows the estimated conditional Kendall's tau as a function of $\log_{10}(\text{GDP})$. It can be observed from the plot that the estimate of Kendall's tau decreases from around 0.8 to 0.6 as GDP per capita increases from $10^3 = 1000$ to $10^{4.6} \approx 40000$ USD and it picks up slightly as GDP per capita becomes greater than 40000 USD. Overall, the dependence between the life expectancy at birth of males and females is relatively larger for countries with lower GDP per capita (less than 10000 USD) and the dependence is relatively smaller for countries with higher GDP per capita (greater than 10000 USD).



(a) **(b)**

Figure 4.3 Life expectancy data.

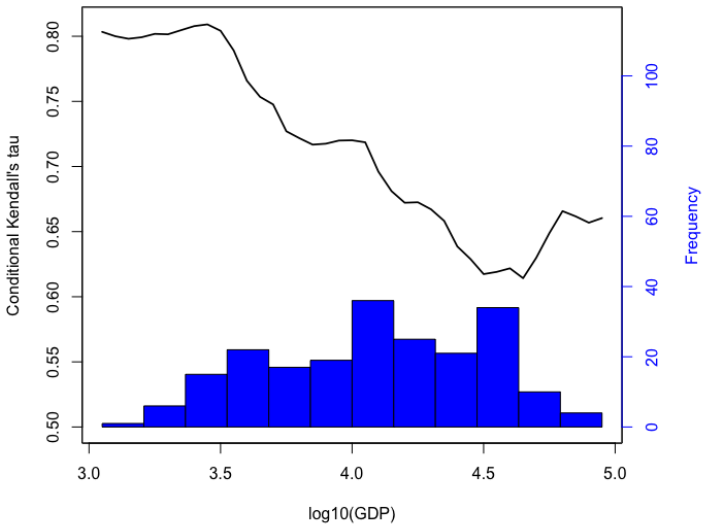


Figure 4.4 Estimated conditional Kendall's tau as a function of $\log_{10}(\text{GDP})$.

4.5 Concluding Remarks

This chapter provides a nonparametric approach for estimating conditional copulas based on the empirical checkerboard Bernstein copula (ECBC) estimator. The proposed nonparametric method has its own advantages compared to the semiparametric methods as it gets rid of model misspecification by not relying on any selection of copula family and demonstrates a good finite-sample performance. The large-sample consistency of the proposed ECBC-based conditional copula estimator is also presented. In addition, we derive closed-form nonparametric estimates of the conditional dependence measures from the proposed estimator.

Due to the complexity in modelling and inference caused by the dependence of conditional copula on the covariates, it is quite common in practice, particularly for vine copulas, to assume that the dependence structure is not influenced by the value of covariates, which is referred to as 'simplifying assumption' (e.g. see Haff et al. [2010], Acar et al. [2012], Stoeber et al. [2013], Nagler & Czado [2016] and Schellhase & Spanhel [2018]). In the literature, there have been some available tests of the simplifying assumption; see Acar et al. [2013], Gijbels et al. [2017b], Gijbels et al. [2017a], Derumigny & Fermanian [2017] and Kurz & Spanhel [2017], etc. Our proposed ECBC-based conditional copula estimator can be useful for constructing new tests of the simplifying assumption. We have shown the framework of obtaining a general estimate of the conditional copula that is allowed to vary with the value of covariates. It is also straightforward to obtain an estimate satisfying the simplifying assumption based on the covariate-adjusted pseudo-observations again using ECBC estimator. Therefore, it could be possible to build test statistics based on some discrepancy criteria like Kolmogorov-Smirnov type, Anderson-Darling type, etc., where the distributions of such test statistics could be approximated by bootstrap schemes.

Another interesting topic for future work would be extending the estimation framework to high-dimensional conditional copula. We can perhaps first use some dimension reduc-

tion methods like principal component analysis (PCA) and then develop copula models based on the lower dimensional principal components of the covariates.

4.6 Technical Details

4.6.1 Proof of Theorem 4.1

Proof. We denote

$$P_{m,k}(v) = \binom{m}{k} v^k (1-v)^{m-k}. \tag{4.35}$$

Then we can rewrite the Bernstein copula as

$$B(C; u_1, u_2, v) = \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m} C\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k}{m}\right) P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P_{m,k}(v), \tag{4.36}$$

and the ECBC copula estimator as

$$B(C_n^{\#}; u_1, u_2, v) = \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m} C_n^{\#}\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k}{m}\right) P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P_{m,k}(v), \tag{4.37}$$

where $C_n^{\#}$ is the empirical checkerboard copula. Thus the partial derivative of 3-dimensional ECBC $B(C_n^{\#}; u_1, u_2, v)$ with respect to v takes the form of

$$\begin{aligned} C^{\#(1)}(u_1, u_2|v) &\equiv \frac{\partial B(C_n^{\#}; u_1, u_2, v)}{\partial v} \\ &= \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m} C_n^{\#}\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k}{m}\right) P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P'_{m,k}(v), \end{aligned} \tag{4.38}$$

where $P'_{m,k}(v)$ is the derivative of $P_{m,k}(v)$ with respect to v.

Let us denote the partial derivative of the Bernstein copula $B(C; u_1, u_2, v)$ with respect

to v as

$$\begin{aligned} C^{(1)}(u_1,u_2|v) &\equiv \frac{\partial B(C;u_1,u_2,v)}{\partial v} \\ &= \sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m} C\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k}{m}\right) P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P^{'}_{m,k}(v), \end{aligned} \tag{4.39}$$

and the partial derivative of the empirical Bernstein copula $B(C_n;u_1,u_2,v)$ with respect to v as

$$\begin{aligned} C_n^{(1)}(u_1,u_2|v) &\equiv \frac{\partial B(C_n;u_1,u_2,v)}{\partial v} \\ &= \sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m} C_n\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k}{m}\right) P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P^{'}_{m,k}(v). \end{aligned} \tag{4.40}$$

Using the triangle inequality we have

$$\begin{aligned} ||C^{\#(1)}-C_v||(v) &\leq ||C^{\#(1)}-C^{(1)}||(v)+||C^{(1)}-C_v|(v) \\ &\leq ||C^{\#(1)}-C_n^{(1)}||(v)+||C_n^{(1)}-C^{(1)}|(v)+||C^{(1)}-C_v||(v). \end{aligned}$$

First, we can show that

$$\begin{aligned} &||C^{\#(1)}-C_n^{(1)}||(v) \\ &= ||\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m}\left(C_n^{\#}\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k}{m}\right)-C_n\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k}{m}\right)\right)P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P^{'}_{m,k}(v)||(v) \\ &\leq \max_{0\leq h_1\leq l_1,0\leq h_2\leq l_2,0\leq k\leq m-1}\left|C_n^{\#}\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k+1}{m}\right)-C_n\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k+1}{m}\right)\right| \\ &\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m}|P_{l_1,h_1}(u_1)||P_{l_2,h_2}(u_2)||P^{'}_{m,k}(v)| \\ &\leq \max_{0\leq h_1\leq l_1,0\leq h_2\leq l_2,0\leq k\leq m-1}\left|C_n^{\#}\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k+1}{m}\right)-C_n\left(\frac{h_1}{l_1},\frac{h_2}{l_2},\frac{k+1}{m}\right)\right|\sum_{k=0}^{m}|P^{'}_{m,k}(v)|. \end{aligned}$$

In above the second inequality follows from the fact that since $\binom{l_j}{h_j}u_j^{l_j}(1-u_j)^{l_j-h_j}, l_j = 0,1,\ldots,h_j,\ j=1,2$ are binomial probabilities, $\sum_{h_j=0}^{l_j}\binom{l_j}{h_j}u_j^{h_j}(1-u_j)^{l_j-h_j}=1$ for any $u_j\in[0,1]$, $j=1,2$. Under the assumption that the marginal CDFs are continuous, it follows from the

Remark 2 in Genest et al. [2017] that for d-dimensional copula

$$||C_n^{\#} - C_n|| \leq \frac{3}{n},$$

and from Lemma 1 in Janssen et al. [2014] it follows that for any fixed $0 < v < 1$,

$$\sum_{k=0}^{m} |P_{m,k}^{'}(v)| \sim \sqrt{\frac{2}{\pi}} \frac{m^{1/2}}{\sqrt{v(1-v)}} = O(m^{1/2}) \text{ as } m \to \infty.$$

Thus, for any fixed $0 < v < 1$ we have

$$\begin{aligned}
&||C^{\#(1)} - C_n^{(1)}||(v) \\
&\leq \max_{0 \leq h_1 \leq l_1, 0 \leq h_2 \leq l_2, 0 \leq k \leq m-1} \left| C_n^{\#}\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k+1}{m}\right) - C_n\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k+1}{m}\right) \right| \sum_{k=0}^{m} |P_{m,k}^{'}(v)| \\
&\leq ||C_n^{\#} - C_n|| \sum_{k=0}^{m} |P_{m,k}^{'}(v)| = O(m^{1/2} n^{-1}).
\end{aligned}$$

Next, we can use a similar technique to show that

$$\begin{aligned}
&||C_n^{(1)} - C^{(1)}|(v) \\
&\leq \max_{0 \leq h_1 \leq l_1, 0 \leq h_2 \leq l_2, 0 \leq k \leq m-1} \left| C_n\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k+1}{m}\right) - C\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k+1}{m}\right) \right| \sum_{k=0}^{m} |P_{m,k}^{'}(v)|
\end{aligned}$$

By using the Lemma 1 in Janssen et al. [2012] and equation (3) in Kiriliouk et al. [2019], for d-dimensional copula we obtain

$$\begin{aligned}
||C_n - C|| &\leq \frac{3}{n} + O(n^{-1/2}(\log\log n)^{1/2}) \ \ a.s. \\
&= O(n^{-1/2}(\log\log n)^{1/2}) \ \ a.s..
\end{aligned}$$

Thus it follows that for any fixed $0 < v < 1$,

$$||C_n^{(1)} - C^{(1)}||(v) = O(m^{1/2} n^{-1/2}(\log\log n)^{1/2}), \ \ a.s..$$

Hence, for any fixed $0 < v < 1$ we have

$$||C^{\#(1)} - C^{(1)}||(v) \le ||C^{\#(1)} - C_n^{(1)}||(v) + ||C_n^{(1)} - C^{(1)}|(v) = O(m^{1/2} n^{-1/2} (\log\log n)^{1/2}), \quad a.s..$$

Next, by mean value theorem there exists $\frac{k}{m} < \xi_k < \frac{k+1}{m}$ s.t.

$$\begin{aligned}
&C^{(1)}(u_1, u_2|v) \\
&= \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m-1} m \left(C\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k+1}{m}\right) - C\left(\frac{h_1}{l_1}, \frac{h_2}{l_2}, \frac{k}{m}\right)\right) P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P_{m-1,k}(v), \\
&= \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m-1} C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \xi_k\right) P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P_{m-1,k}(v), \\
&= \sum_{h_1=0}^{l_1} \sum_{h_2=0}^{l_2} \sum_{k=0}^{m-1} \left(C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \xi_k\right) - C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \frac{k}{m-1}\right) + C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \frac{k}{m-1}\right)\right) \\
&P_{l_1,h_1}(u_1) P_{l_2,h_2}(u_2) P_{m-1,k}(v).
\end{aligned}$$

Notice that

$$\frac{k}{m} - \frac{k+1}{m} < \frac{k}{m} - \frac{k}{m-1} < \xi_k - \frac{k}{m-1} < \frac{k+1}{m} - \frac{k}{m-1} < \frac{k+1}{m} - \frac{k}{m},$$

which means that

$$\left| \xi_k - \frac{k}{m-1} \right| < \frac{1}{m}.$$

If $C_v(u_1, u_2|v) = \frac{\partial C(u_1,u_2,v)}{\partial v}$ is Lipschitz continuous on $[0,1]^3$, then there exists a Lipschitz constant L s.t.

$$\left| C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \xi_k\right) - C_v\left(\frac{h_1}{l_1}, \frac{h_2}{l_2} \middle| \frac{k}{m-1}\right) \right| \le L \left| \xi_k - \frac{k}{m-1} \right| \le \frac{L}{m},$$

and based on Lemma 3.2 in Segers et al. [2017] we also have

$$\begin{aligned}&||\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m-1}C_v\left(\frac{h_1}{l_1},\frac{h_2}{l_2}\Big|\frac{k}{m-1}\right)P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P_{m-1,k}(v)-C_v||\\&\leq L\left(\frac{1}{2\sqrt{l_1}}+\frac{1}{2\sqrt{l_2}}+\frac{1}{2\sqrt{m-1}}\right).\end{aligned}$$

Thus

$$\begin{aligned}&||C^{(1)}-C_v||(v)\\&\leq||\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m-1}\left(C_v\left(\frac{h_1}{l_1},\frac{h_2}{l_2}\Big|\xi_k\right)-C_v\left(\frac{h_1}{l_1},\frac{h_2}{l_2}\Big|\frac{k}{m-1}\right)\right)P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P_{m-1,k}(v)||\\&+||\sum_{h_1=0}^{l_1}\sum_{h_2=0}^{l_2}\sum_{k=0}^{m-1}C_v\left(\frac{h_1}{l_1},\frac{h_2}{l_2}\Big|\frac{k}{m-1}\right)P_{l_1,h_1}(u_1)P_{l_2,h_2}(u_2)P_{m-1,k}(v)-C_v||\\&\leq\frac{L}{m}+L\left(\frac{1}{2\sqrt{l_1}}+\frac{1}{2\sqrt{l_2}}+\frac{1}{2\sqrt{m-1}}\right).\end{aligned}$$

Finally, for any fixed $0<v<1$ we obtain

$$\begin{aligned}||C^{\#(1)}-C_v||(v)&\leq||C^{\#(1)}-C^{(1)}||(v)+||C^{(1)}-C_v||(v)\\&\leq\frac{L}{m}+L\left(\frac{1}{2\sqrt{l_1}}+\frac{1}{2\sqrt{l_2}}+\frac{1}{2\sqrt{m-1}}\right)+O(m^{1/2}n^{-1/2}(\log\log n)^{1/2}),\quad a.s..\end{aligned}$$

The empirical prior of the degrees m, l_1, and l_2 are given as

$$m|\alpha\sim Poisson(n^{\alpha})+2 \text{ and } \alpha\sim Unif\Big(\frac{1}{3},\frac{2}{3}\Big),$$

$$l_j|\alpha_j\sim Poisson(n^{\alpha_j})+1 \text{ and } \alpha_j\sim Unif\Big(\frac{1}{3},\frac{2}{3}\Big)\ j=1,2.$$

Notice that $\Pr(\frac{1}{3}\leq\alpha\leq\frac{2}{3})=1$ and $\Pr(\frac{1}{3}\leq\alpha_j\leq\frac{2}{3})=1, j=1,2$, then $E(m^{1/2}n^{-1/2}(\log\log n)^{1/2})$ $\leq n^{1/3}n^{-1/2}(\log\log n)^{1/2}\to 0$ as $n\to\infty$. In the proof of Theorem 3.1 it has been shown $E\left(\sqrt{\frac{1}{l_j}}\Big|\alpha_j\right)\leq\sqrt{\frac{1-e^{-n^{\alpha_j}}}{n^{\alpha_j}}}\to 0, j=1,2$, $E\left(\sqrt{\frac{1}{m-1}}\Big|\alpha\right)\leq\sqrt{\frac{1-e^{-n^{\alpha}}}{n^{\alpha}}}\to 0$ as $n\to\infty$ and $E\left(\frac{1}{m}\Big|\alpha\right)\leq$ $\frac{1-e^{-n^{\alpha}}}{n^{\alpha}}\to 0$ as $n\to\infty$. Thus, taking expectation with respect to the prior distributions of l_1,

l_2 and m as given for ECBC, it follows that

$$
\begin{aligned}
&E(||C^{\#(1)} - C_v||(v)) \\
&\leq E\left(\frac{L}{m}\right) + E\left(L\left(\frac{1}{2\sqrt{l_1}} + \frac{1}{2\sqrt{l_2}} + \frac{1}{2\sqrt{m-1}}\right)\right) + E\left(O(m^{1/2}n^{-1/2}(\log\log n)^{1/2})\right) \\
&\to 0 \; a\,s \; n \to \infty \;\; a.s.
\end{aligned}
$$

Since $C(u_1, 1|v) = u_1$ and $C(1, u_2|v) = u_2$ and $C^{\#(1)}(u_1, u_2|v)$ converges to $C_v(u_1, u_2|v)$ uniformly on $[0,1]^2$ as $n \to \infty$ for any fixed $0 < v < 1$, then we have

$$
\begin{aligned}
&E\left(||F_1(u_1|v) - u_1||(v)\right) \equiv E\left(\sup_{u_1 \in [0,1]} |F_1(u_1|v) - u_1|\right) \\
&= E\left(||C^{\#(1)}(u_1, 1|v) - C(u_1, 1|v)||(v)\right) \leq E\left(||C^{\#(1)} - C^{(1)}||(v)\right) \xrightarrow{\text{a.s.}} 0,
\end{aligned}
$$

and

$$
\begin{aligned}
&E\left(||F_2(u_2|v) - u_2||(v)\right) \equiv E\left(\sup_{u_2 \in [0,1]} |F_2(u_2|v) - u_2|\right) \\
&= E\left(||C^{\#(1)}(1, u_2|v) - C(1, u_2|v)||(v)\right) \leq E\left(||C^{\#(1)} - C^{(1)}||(v)\right) \xrightarrow{\text{a.s.}} 0.
\end{aligned}
$$

For any fixed $0 < v < 1$, F_1 and F_2 are non-decreasing functions, so $F_1^{-1}(u_1|v) \xrightarrow{\text{a.s.}} u_1$ and $F_2^{-1}(u_2|v) \xrightarrow{\text{a.s.}} u_2$, Thus we obtain the uniform convergence of

$$
E(||C^{\#} - C_v||(v)) = E(||C^{\#(1)}(F_1^{-1}(u_1|v), F_2^{-1}(u_2|v)|v) - C_v||(v)) \xrightarrow{\text{a.s.}} 0
$$

as $n \to \infty$ for any fixed $0 < v < 1$.

□

BIBLIOGRAPHY

Abegaz, F. et al. (2012). "Semiparametric estimation of conditional copulas". *Journal of Multivariate Analysis* **110**, pp. 43–73.

Acar, E. F. et al. (2011). "Dependence calibration in conditional copulas: A nonparametric approach". *Biometrics* **67**.2, pp. 445–453.

Acar, E. F. et al. (2012). "Beyond simplified pair-copula constructions". *Journal of Multivariate Analysis* **110**, pp. 74–90.

Acar, E. F. et al. (2013). "Statistical testing of covariate effects in conditional copula models". *Electronic Journal of Statistics* **7**, pp. 2822–2850.

Amblard, C. & Girard, S. (2002). "Symmetry and dependence properties within a semiparametric family of bivariate copulas". *Journal of Nonparametric Statistics* **14**.6, pp. 715–727.

Amblard, C. & Girard, S. (2005). "Estimation procedures for a semiparametric family of bivariate copulas". *Journal of Computational and Graphical Statistics* **14**.2, pp. 363–377.

Balsubramani, A. (2014). "Sharp finite-time iterated-logarithm martingale concentration". *arXiv preprint arXiv:1405.2639*.

Belalia, M et al. (2017). "Testing independence based on Bernstein empirical copula and copula density". *Journal of Nonparametric Statistics* **29**.2, pp. 346–380.

Berger, J. O. et al. (1999). "Integrated likelihood methods for eliminating nuisance parameters". *Statistical science* **14**.1, pp. 1–28.

Bouezmarni, T. et al. (2010). "Asymptotic properties of the Bernstein density copula estimator for α-mixing data". *Journal of Multivariate Analysis* **101**.1, pp. 1–10.

Burda, M. & Prokhorov, A. (2014). "Copula based factorization in Bayesian multivariate infinite mixture models". *Journal of Multivariate Analysis* **127**, pp. 200–213.

Carley, H & Taylor, M. (2002). "A new proof of Sklar's theorem". *Distributions with given marginals and statistical modelling*. Springer, pp. 29–34.

Chen, S. X. & Huang, T.-M. (2007). "Nonparametric estimation of copula functions for dependence modelling". *Canadian Journal of Statistics* **35**.2, pp. 265–282.

Craiu, V. R. & Sabeti, A. (2012). "In mixed company: Bayesian inference for bivariate conditional copula models with discrete and continuous outcomes". *Journal of Multivariate Analysis* **110**, pp. 106–120.

Deheuvels, P. (1979). “La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d’indépendance”. *Bulletins de l’Académie Royale de Belgique* **65**.1, pp. 274–292.

Denuit, M. & Scaillet, O. (2004). “Nonparametric tests for positive quadrant dependence”. *Journal of Financial Econometrics* **2**.3, pp. 422–450.

Derumigny, A. & Fermanian, J.-D. (2020). “Conditional empirical copula processes and generalized dependence measures”. *arXiv preprint arXiv:2008.09480*.

Derumigny, A. & Fermanian, J.-D. (2017). “About tests of the “simplifying" assumption for conditional copulas”. *Dependence Modeling* **5**.1, pp. 154–197.

Diers, D. et al. (2012). “Dependence modeling in non-life insurance using the Bernstein copula”. *Insurance: Mathematics and Economics* **50**.3, pp. 430–436.

Durante, F. & Sempi, C. (2015). *Principles of copula theory*. CRC press.

Embrechts, P. et al. (2001). “Modelling dependence with copulas”. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich* **14**.

Fermanian, J.-D. & Lopez, O. (2018). “Single-index copulas”. *Journal of Multivariate Analysis* **165**, pp. 27–55.

Fermanian, J.-D. et al. (2004). “Weak convergence of empirical copula processes”. *Bernoulli* **10**.5, pp. 847–860.

Geerdens, C. et al. (2018). “Conditional copula models for right-censored clustered event time data”. *Biostatistics* **19**.2, pp. 247–262.

Genest, C. et al. (2017). “Asymptotic behavior of the empirical multilinear copula process under broad conditions”. *Journal of Multivariate Analysis* **159**, pp. 82–110.

Gijbels, I. & Mielniczuk, J. (1990). “Estimating the density of a copula function”. *Communications in Statistics-Theory and Methods* **19**.2, pp. 445–464.

Gijbels, I. & Sznajder, D. (2013). “Positive quadrant dependence testing and constrained copula estimation”. *Canadian Journal of Statistics* **41**.1, pp. 36–64.

Gijbels, I. et al. (2010). “Positive quadrant dependence tests for copulas”. *Canadian Journal of Statistics* **38**.4, pp. 555–581.

Gijbels, I. et al. (2011). “Conditional copulas, association measures and their applications”. *Computational Statistics & Data Analysis* **55**.5, pp. 1919–1932.

Gijbels, I. et al. (2012). “Multivariate and functional covariates and conditional copulas”. *Electronic Journal of Statistics* **6**, pp. 1273–1306.

Gijbels, I. et al. (2017a). “Nonparametric testing for no covariate effects in conditional copulas”. *Statistics* **51**.3, pp. 475–509.

Gijbels, I. et al. (2017b). “Score tests for covariate effects in conditional copulas”. *Journal of Multivariate Analysis* **159**, pp. 111–133.

Haff, I. H. et al. (2010). “On the simplified pair-copula construction-simply useful or too simplistic?” *Journal of Multivariate Analysis* **101**.5, pp. 1296–1310.

Hofert, M. et al. (2014). “Package ‘copula’”. *URL http://ie. archive. ubuntu. com/disk1/disk1 /cran. r-project. org/web/packages/copula/copula. pdf*.

Janssen, P. et al. (2012). “Large sample behavior of the Bernstein copula estimator”. *Journal of Statistical Planning and Inference* **142**.5, pp. 1189–1197.

Janssen, P. et al. (2014). “A note on the asymptotic behavior of the Bernstein estimator of the copula density”. *Journal of Multivariate Analysis* **124**, pp. 480–487.

Janssen, P. et al. (2016). “Bernstein estimation for a copula derivative with application to conditional distribution and regression functionals”. *Test* **25**.2, pp. 351–374.

Jaworski, P. et al. (2010). *Copula theory and its applications*. Vol. 198. Springer.

Joe, H. (2014). *Dependence modeling with copulas*. CRC press.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. Chapman and Hall/CRC.

— (2005). “Asymptotic efficiency of the two-stage estimation method for copula-based models”. *Journal of multivariate Analysis* **94**.2, pp. 401–419.

Jorion, P. (2007). *Value at risk: the new benchmark for managing financial risk*. The McGraw-Hill Companies, Inc.

Kiriliouk, A. et al. (2019). “On some resampling procedures with the empirical beta copula”. *Manuscript*.

Kosorok, M. R. (2007). *Introduction to empirical processes and semiparametric inference*. Springer Science & Business Media.

Kurz, M. S. & Spanhel, F. (2017). “Testing the simplifying assumption in high-dimensional vine copulas”. *arXiv preprint arXiv:1706.02338*.

Leisen, F. et al. (2017). “Bayesian Nonparametric Conditional Copula Estimation of Twin Data”. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*.

Levi, E. & Craiu, R. V. (2018). “Bayesian inference for conditional copulas using Gaussian Process single index models”. *Computational Statistics & Data Analysis* **122**, pp. 115–134.

Li, X et al. (1997). “On approximation of copulas”. *Distributions with given marginals and moment problems*. Springer, pp. 107–116.

Liew, C. K. (1976). “Inequality constrained least-squares estimation”. *Journal of the American Statistical Association* **71**.355, pp. 746–751.

Lu, L. & Ghosh, S. K. (2020). “Nonparametric Estimation and Testing for Positive Quadrant Dependent Bivariate Copula”. *Journal of Business & Economic Statistics*, pp. 1–36.

McNeil, A. J., Nešlehová, J., et al. (2009). “Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions”. *The Annals of Statistics* **37**.5B, pp. 3059–3097.

Nagler, T. & Czado, C. (2016). “Evading the curse of dimensionality in nonparametric density estimation with simplified vine copulas”. *Journal of Multivariate Analysis* **151**, pp. 69–89.

Nelsen, R. B. (1996). “Nonparametric measures of multivariate association”. *Lecture Notes-Monograph Series*, pp. 223–232.

— (2007). *An introduction to copulas*. Springer Science & Business Media.

Omelka, M. et al. (2009). “Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing”. *The Annals of Statistics* **37**.5B, pp. 3023–3058.

Patton, A. J. (2006a). “Modelling asymmetric exchange rate dependence”. *International economic review* **47**.2, pp. 527–556.

Patton, A. J. (2006b). “Estimation of multivariate models for time series of possibly different lengths”. *Journal of applied econometrics* **21**.2, pp. 147–173.

Pérez, A. & Prieto-Alaiz, M. (2016). “A note on nonparametric estimation of copula-based multivariate extensions of Spearman’s rho”. *Statistics & Probability Letters* **112**, pp. 41–50.

Quiroz, M. et al. (2018). “Speeding up MCMC by efficient data subsampling”. *Journal of the American Statistical Association*.

Rémillard, B. & Scaillet, O. (2009). “Testing for equality between two copulas”. *Journal of Multivariate Analysis* **100**.3, pp. 377–386.

Robert, C. P. et al. (2018). “Accelerating MCMC algorithms”. *Wiley Interdisciplinary Reviews: Computational Statistics* **10**.5, e1435.

Rockafellar, R. T., Uryasev, S., et al. (2000). "Optimization of conditional value-at-risk". *Journal of risk* **2**, pp. 21–42.

Sabeti, A. et al. (2014). "Additive models for conditional copulas". *Stat* **3**.1, pp. 300–312.

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

Scaillet, O. (2005). "A Kolmogorov-Smirnov type test for positive quadrant dependence". *Canadian Journal of Statistics* **33**.3, pp. 415–427.

Scaillet, O. & Fermanian, J.-D. (2002). "Nonparametric estimation of copulas for time series". *FAME Research paper* 57.

Schellhase, C. & Spanhel, F. (2018). "Estimating non-simplified vine copulas using penalized splines". *Statistics and Computing* **28**.2, pp. 387–409.

Schmid, F. & Schmidt, R. (2007). "Multivariate extensions of Spearman's rho and related statistics". *Statistics & Probability Letters* **77**.4, pp. 407–416.

Schweizer, B., Wolff, E. F., et al. (1981). "On nonparametric measures of dependence for random variables". *Annals of Statistics* **9**.4, pp. 879–885.

Segers, J. et al. (2017). "The empirical beta copula". *Journal of Multivariate Analysis* **155**, pp. 35–51.

Semenov, M. & Smagulov, D. (2017). "Portfolio risk assessment using copula models". *arXiv preprint arXiv:1707.03516*.

Sklar, M (1959). "Fonctions de repartition an dimensions et leurs marges". *Publ. inst. statist. univ. Paris* **8**, pp. 229–231.

Smith, M. S. (2011). "Bayesian approaches to copula modelling". *arXiv preprint arXiv:1112.4204*.

Stoeber, J. et al. (2013). "Simplified pair copula constructions-limitations and extensions". *Journal of Multivariate Analysis* **119**, pp. 101–118.

Uryasev, S. (2000). "Conditional value-at-risk: Optimization algorithms and applications". *Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr)(Cat. No. 00TH8520)*. IEEE, pp. 49–57.

Vatter, T. & Chavez-Demoulin, V. (2015). "Generalized additive models for conditional dependence structures". *Journal of Multivariate Analysis* **141**, pp. 147–167.

Veraverbeke, N. et al. (2011). "Estimation of a conditional copula and association measures". *Scandinavian Journal of Statistics* **38**.4, pp. 766–780.

Wu, J. et al. (2014). “Bayesian nonparametric inference for a multivariate copula function”. *Methodology and Computing in Applied Probability* **16**.3, pp. 747–763.

Žežula, I. (2009). “On multivariate Gaussian copulas”. *Journal of Statistical Planning and Inference* **139**.11, pp. 3942–3946.