Summer 8-2020

D-Vine Pair-Copula Models for Longitudinal Binary Data

Huihui Lin
Old Dominion University, hlin005@odu.edu

Recommended Citation

Lin, Huihui. "D-Vine Pair-Copula Models for Longitudinal Binary Data" (2020). Doctor of Philosophy (PhD), Dissertation, Mathematics and Statistics, Old Dominion University, DOI: 10.25777/yajd-fq74
https://digitalcommons.odu.edu/mathstat_etds/114

D-VINE PAIR-COPULA MODELS FOR LONGITUDINAL BINARY DATA

by

Huihui Lin
B.S. May 2006, East China University Of Science and Technology, China
M.S. Sep. 2012, Michigan Technological University
M.S. Dec. 2014, Michigan Technological University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

COMPUTATIONAL AND APPLIED MATHEMATICS

OLD DOMINION UNIVERSITY
August 2020

Approved by:

N. Rao Chaganty (Director)

Lucia Tabacu (Member)

Sandipan Dutta (Member)

Hadiza Galadima (Member)

ABSTRACT

D-VINE PAIR-COPULA MODELS FOR LONGITUDINAL BINARY DATA

Huihui Lin
Old Dominion University, 2020
Director: Dr. N. Rao Chaganty

Dependent longitudinal binary data are prevalent in a wide range of scientific disciplines, including healthcare and medicine. A popular method for analyzing such data is the multivariate probit (MP) model. The motivation for this dissertation stems from the fact that the MP model fails even the binary correlations are within the feasible range. The reason being the underlying correlation matrix of the latent variables in the MP model may not be positive definite. In this dissertation, we study alternatives that are based on D-vine pair-copula models. We consider both the serial dependence modeled by the first order autoregressive (AR(1)) and the equicorrelated correlation structures. Simulation results show that our model is more effective than MP model. Some real life data analysis are presented to show usefulness of our models. We also consider a general situation where the marginal distributions are ordered multinomial. We extend the D-vine pair-copula model to handle multinomial longitudinal data, and compare the generated probability distributions with other methods that are available in R packages.

This dissertation is dedicated to my family.

ACKNOWLEDGEMENTS

I sincerely thank my advisor Dr. N. Rao Chaganty for suggesting such an interesting topic. It would be impossible to finish this research work without his guidance, support and mentorship throughout.

I am also extremely grateful for Dr. Lucia Tabacu, Dr. Sandipan Dutta and Dr. Hadiza Galadima for serving on my committee and providing kind comments to improve the quality of this work.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

I.1 BACKGROUND

Longitudinal dependent binary or multinomial data are prevalent in a wide range of scientific disciplines, such as biology, economics, medicine, public health, and social sciences. These data normally consists of measurements taken at several sequential time points on each respondent, which could be individuals, households, or other experimental units. The observed data could be grouped or ungrouped. An example of grouped longitudinal data from Stock et al. (1983) is shown in Table 1. These data are the traffic violations of high school students for a period of 4 years. Here Y_j is a binary indicator and represents whether the high school students had traffic violation(s) at year j, for $j = 1, 2, 3, 4$. There were 731 students who never had any traffic violation(s) during the monitoring period, and there were 310 subjects who only had traffic violation(s) at the forth year, and so on.

An example of ungrouped data is reported in Woolson and Clarke (1984). This is a longitudinal study of obesity status in school children. Besides binary obesity indicator the data also consists of covariate information such as age, gender etc. A subset of the data is presented in Table 2. The response variables are obesity indicators (1 for obese kid, 0 for non-obese kid) in years 1977, 1979 and 1981. The covariates are gender(0=Male, 1=Female), baseline age, time at observation taking values $1, 2, 3$ for the three years 1977, 1979 and 1981, respectively.

Although there are many well-developed tools for analyzing continuous longitudinal data, the methods for the analysis of categorical longitudinal data are still in development. The most widely used methodology for analysis of longitudinal categorical data is the generalized estimating equations (GEE) approach (see Liang and Zeger (1986)), which is based on estimating equations, moment estimates and a non-likelihood approach. However, the GEE method when applied to correlated binary data could lead to misleading conclusions see Chaganty and Joe (2004) and Sabo and Chaganty (2010). Alternative methods include the Gaussian method of estimation, which maximizes the Gaussian log likelihood function, (Crowder, 2001) or the quasi-least squares method of estimation for the correlation (Chaganty, 1997), the first-order Markov chains model and the multivariate probit model (Chaganty and Joe, 2004), (Yang and Chaganty, 2014).

Table 1: Traffic violation data

$(Y_1,\ Y_2,\ Y_3,\ Y_4)$	**Frequency**
(No, No, No, No)	731
(No, No, No, Yes)	310
(No, No, Yes, No)	256
(No, No, Yes, Yes)	196
(No, Yes, No, No)	156
(No, Yes, No, Yes)	121
(No, Yes, Yes, No)	114
(No, Yes, Yes, Yes)	152
(Yes, No, No, No)	61
(Yes, No, No, Yes)	40
(Yes, No, Yes, No)	45
(Yes, No, Yes, Yes)	39
(Yes, Yes, No, No)	47
(Yes, Yes, No, Yes)	42
(Yes, Yes, Yes, No)	46
(Yes, Yes, Yes, Yes)	53

In the last decade, the use of copulas has grown exponentially in various disciplines, including the analysis of discrete longitudinal data. A copula is simply a multivariate cumulative distribution function with uniform univariate marginals. A major advantage of copulas is that they can separate and capture the dependence in longitudinal categorical data. An important special case are the pair-copulas based on vines (Bedford and Cooke, 2002), (Czado, 2019). Instead of dealing directly, the pair-couplas factor the multivariate copula into pieces involving only bivariate copulas.

Table 2: Obesity data

ID	**Gender**	**Baseline age**	**Time point**	**Obesity**
1	0	6	1	1
1	0	6	2	1
1	0	6	3	1
2	0	6	1	1
2	0	6	2	1
2	0	6	3	1
3	0	8	1	1
3	0	8	2	1
3	0	8	3	1
...				

In the literature numerous authors have developed copula-based models for longitudinal categorical data. Noteworthy to mention are multivariate copula based models for binary, ordinal categorical and count data (Xu, 1996), multivariate Gaussian copula based model for quantile regression; elliptical copula based model for both discrete and continuous longitudinal data in the actuarial literature (Frees and Wang, 2006); D-vine pair-copula model for the high-dimensional discrete data in a Bayesian framework (Smith and Khaled, 2012). There are several R packages, (e.g., CDVine from Brechmann and Schepsmeier (2013), and VineCopula from Schepsmeier et al. (2015)) providing functions and tools for statistical inference of canonical vine (C-vine) and D-vine copulas. There is also a modified package, "rvinecopulib" from Nagler and Vatter (2018), that can handle discrete margins. In this dissertation, we propose a D-vine pair-copula model for longitudinal discrete binary data, and extend it later for longitudinal multinomial data.

I.2 OVERVIEW OF THE DISSERTATION

This dissertation is organized as follows. In Chapter II, we propose D-vine pair-copula model for analyzing longitudinal binary data. Since the pair-copula model uses bivariate copulas, we discuss important bivariate copulas including Gaussian, Clayton, Frank, and Gumbel. We also study the relationship between the copula parameters and the correlation between the binary variables. For the Gaussian copula, we present Hermite polynomials approximation of Gaussian copula parameter for a specified correlation between the binary variables. We show how to construct the probability mass function (PMF) for the D-vine pair-copula model for a given marginals that will result in AR(1) or equicorrelated correlation structures. We give numerical examples in three and four dimensions. We compare our PMF with the PMF obtained by the multivariate probit (MP) model (see Yang and Chaganty (2014)), which is a popular method for analyzing logitudinal binary data. We show with numerical examples our D-vine pair-copula model successfully generates a PMF in cases where the MP model fails, demonstrating the supreriority of our approach. We end the chapter fitting the D-vine pair-copula to real life data and compare the analysis with the MP model.

In Chapter III, we extend our work in Chapter II to the regression setting for longitudinal data that includes covariates besides the binary responses. We derive the necessary formulas for the score functions, which we use to obtain the maximum likelihood estimates of the regression and the correlation parameters. We present simulation studies to compare the asymptotic and small sample efficiencies of our D-vine pair-copula model that uses bivariate Gaussian copula and the MP model. We conclude the chapter fitting our models to a couple of real life longitudinal binary data.

In Chapter IV, we extend the models that we developed in Chapter II for longitudinal multinomial categorical data. We present step by step procedure for calculating the PMF using the D-vine pair-copula with bivariate Gaussian. As in Chapter II, we obtain the relationship between the copula parameter and the correlation between the multinomial variable's. Then, we compare the PMFs created by our model with the PMF's generated by two other R packages. We conduct a small simulation study to compare the performance of estimating the marginal distributions and the correlation beytween the multinomial variables. We end the dissertation with summary of this study and related research problems that we plan to pursue in future in Chapter V.

CHAPTER II

D-VINE PAIR-COPULA MODEL

II.1 INTRODUCTION

In clinical trials and research studies in medicine, and health care, the endpoint of the observed data most often consists of longitudinal binary observations. A popular statistical tool for analyzing such data has been the method of generalized estimating equations (GEEs), introduced by Liang and Zeger (1986). However, this method has several drawbacks. It uses an ambiguously defined working correlation to model the dependence in the longitudinal binary observations. Also it is a non-likelihood approach, in the sense it does not have an underlying probability model for the dependent binary observations. Alternatives to GEEs for the analysis of longitudinal binary data are Markov chains (MCs) and multivariate probit (MP) models. A contrasting study of the first order MC model and the MP model was presented by Yang and Chaganty (2014). They showed that both models are asymptotically efficient, and discussed situations where one is preferable over the other.

In recent years, due to their popularity, the copulas have been used as another alternative to the GEEs. Some researchers have combined copulas with MCs models. Escarela et al. (2009), have used Gaussian copula to construct conditional probabilities in MC models in the context of longitudinal binary data. The copula based bivariate probit models, were generalized by Winkelmann (2012), with non-normal dependence between binary responses using Frank and Clayton copulas. A nonlinear regression models were introduced by Radice et al. (2016), where the non-Gaussian copulas were used to deal with the dependence between binary responses. Smith et al. (2010) showed that longitudinal continuous data can be modeled by D-vine pair-copula, and later extended the work to the discrete case in a Bayesian framework in Smith and Khaled (2012). A Gaussian copula model for integer-valued ARMA structured time series data with or without covariates was developed by Lennon (2016). Panagiotelis et al. (2017) introduced two algorithms for optimizing vine structure and pair-copula selection for discrete regular vine copula, the first one was using a modified Akaike information criterion (mAIC) and the second was using predictive scores with cross-validation out of sample approach.

There remains still a need for efficiently modeling dependence among longitudinal binary variables using pair-copula. In this chapter we introduce D-vine pair-copula models for longitudinal binary data. These models are relatively easy to implement since they use only bivariate copulas, and flexible because they allow different types of bivariate copulas to model different type of dependence in conditional distributions. We will see that the D-vine pair-copula model enjoys several advantages over the MP model. The organization of this chapter is as follows. We first discuss the bivariate Gaussian, Clayton, Frank, Gumbel copulas and find the relation between the correlation of the binary variables and the copula parameters in Section II.2. Use of Hermite polynomials to numerically compute the Gaussian copula parameter for a given correlation is discussed in Section II.3. Comparisons between D-vine pair-copula models and the multivariate model (MP) model are discussed in Section II.4, together with some numerical examples where the vine model overcomes the difficulties associated with MP model. In Section II.5, we discuss parameter estimation by maximum likelihood for grouped data. Section II.6 contains analysis of two real life data. We end the chapter with some conclusions in Section II.7.

II.2 VINE PAIR-COPULA FOR BINARY DATA

Copula is extremely popularly used to bond the marginal distributions to their joint high-dimensional distribution, according to Sklar's Theorem (Sklar, 1959). In other words, copula is very useful to separate dependency relationship from multivariate distribution. In this paper, bivariate copula is sufficient because a multivariate copula, which is basically a joint cumulative distribution function (CDF), can be decomposed into pair vine copula with corresponding margins. Therefore, we start with presenting some widely used bivariate copulas and expanding to the specific situation of Bernoulli marginal distributions.

II.2.1 BIVARIATE COPULA FAMILIES

We will use and compare the well known bivariate copulas, as Gaussian copula from Elliptical copula families, Clayton/MTCJ copula, Gumbel copula and Frank copula from Archimediean copula families.

The bivariate Gaussian copula is given by

$$C(u_1, u_2; \gamma) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \gamma),$$

where Φ_2 is the standard bivariate normal CDF with correlation coefficient γ, and Φ^{-1} is the inverse of the standard normal CDF. Gaussian copula is the most popular copula, because of its well developed characteristics. The relationship between γ and the correlation coefficient ρ of binary variables (Y_1, Y_2) with marginal mean (p_1, p_2) is given in Equation (2.2.1) from Emrich and Piedmonte (1991),

$$\text{Corr}(Y_1, Y_2) = \rho = \frac{\Phi_2(\Phi^{-1}(p_1), \Phi^{-1}(p_2); \gamma) - p_1p_2}{\sqrt{p_1q_1p_2q_2}} \tag{2.2.1}$$

More details of this relationship can be found in Section II.2.4.

The Clayton copula has the CDF,

$$C(u_1, u_2; \alpha) = \max([\, u_1^{-\alpha} + u_2^{-\alpha} - 1]^{\,-\frac{1}{\alpha}}, 0),$$

where the parameter $\alpha \in [-1, \infty)\backslash\{0\}$. The relationship between α and the correlation coefficient of binary variables ρ is given in Equation (2.2.2).

$$\text{Corr}(Y_1, Y_2) = \rho = \frac{(q_1^{-\alpha} + q_2^{-\alpha} - 1)^{-\frac{1}{\alpha}} - q_1q_2}{\sqrt{p_1q_1p_2q_2}}, \tag{2.2.2}$$

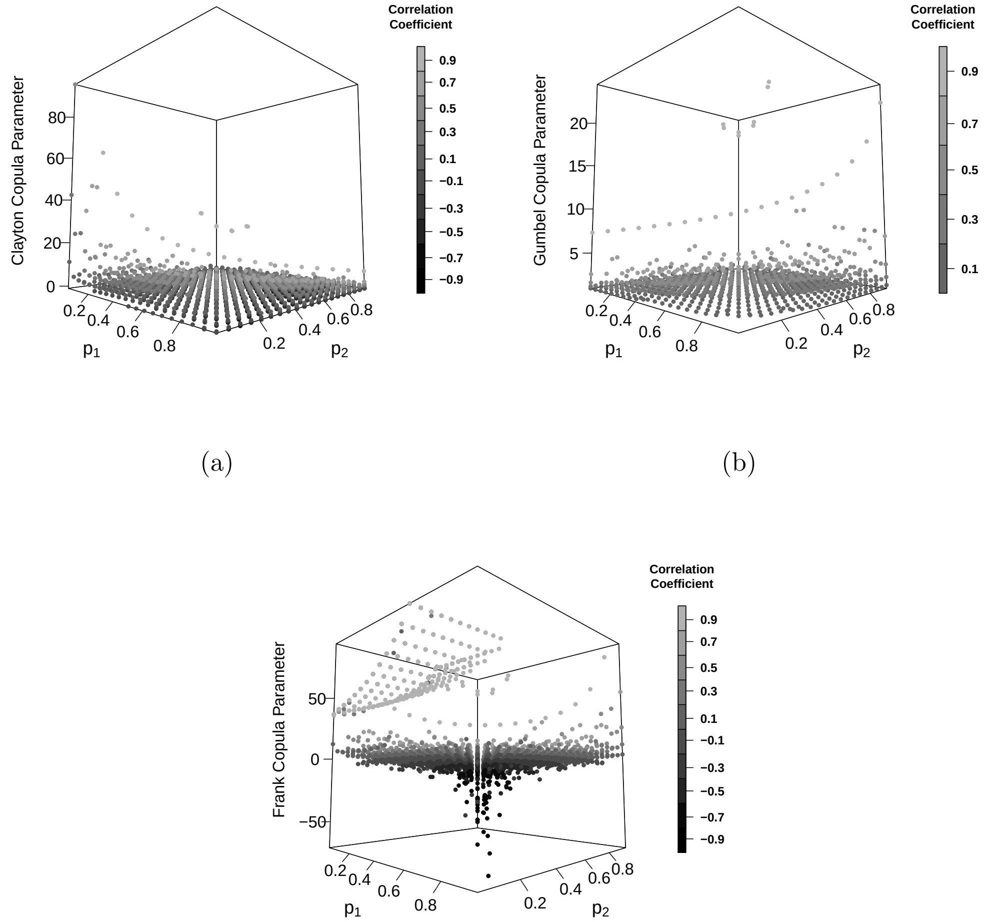
where $q_i = 1 - p_i$. More details are in Section II.2.4.

The Frank copula is given by

$$C(u_1, u_2; \alpha) = -\frac{1}{\alpha}\log(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1}),$$

where the parameter α can be any value except 0. The relationship between α and the correlation coefficient of binary variables ρ is given in Equation (2.2.3).

$$\text{Corr}(Y_1, Y_2) = \rho = \frac{-\frac{1}{\alpha}\log(1 + \frac{(e^{-\alpha q_1} - 1)(e^{-\alpha q_2} - 1)}{e^{-\alpha} - 1}) - q_1q_2}{\sqrt{p_1q_1p_2q_2}} \tag{2.2.3}$$



(a)

(b)

(c)

Figure 1: Plots of relationship between copula parameter α and marginal proportions with binary correlation. (a) Clayton copula; (b) Gumbel copula; (c) Frank copula

The Gumbel copula is given by

$$C(u_1, u_2; \alpha) = e^{-[(-\log u_1)^{\alpha} + (-\log u_2)^{\alpha}]^{\frac{1}{\alpha}}},$$

where the parameter α is greater or equal to 1. Gumbel copula works great for positively correlated data, but not for negatively correlated data. The relationship between α and the correlation coefficient of binary variables ρ is given by

$$\mathrm{Corr}(Y_1, Y_2) = \rho = \frac{e^{-[\,(-\log q_1)^{\alpha} + (-\log q_2)^{\alpha}\,]^{\frac{1}{\alpha}}} - q_1 q_2}{\sqrt{p_1 q_1 p_2 q_2}} \tag{2.2.4}$$

The independent copula is given by

$$C(u_1, u_2) = u_1 * u_2,$$

which can be used when pairs of variables are suspicious to be independent. That is, the bivariate copulas we covered as above become independent copula when $\gamma = 0$ for Gaussian copula, $\alpha = 0$ for Clayton copula and Frank copula, or $\alpha = 1$ for Gumbel copula.

Since in the longitudinal binary data case, the correlation structure is usually assumed as autoregressive or equicorrelated structure, which consists of correlation coefficient ρ of the binary variable. Tables in Appendix A shows numerical examples of the relationship between binary variables correlation coefficient ρ and the copula parameter α. Figure 1 shows that for all three copula while correlation coefficient of binary variables ρ is further away from 0, it has fewer available marginal proportions combinations to get the copula parameter α, this is because of the feasible range of correlation of the binary variable which can be found in Section II.2.6. The copula parameters ranges are consistent with the parameter requirement as well: Clayton copula has parameter α greater than -1, Frank copula has α not zero, Gumbel copula has α greater than 1.

II.2.2 BIVARIATE BINARY DISTRIBUTIONS

Consider first the case of two binary variables. Let $Y = (Y_1,\, Y_2)$, where 1 and 2 possibly may indicate two sequential time points. The joint CDF of Y using a copula C function would be, according to Sklar (1959) theorem, is $F(y_1, y_2) = C(F_1(y_1), F_2(y_2))$, where F_1 and F_2 are CDFs of the univariate binary distributions with means p_1 and p_2 respectively. Following Panagiotelis et al. (2012), we can recover the joint probability mass function of Y from the CDF as

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2) &= C(F_1(y_1), F_2(y_2);\ \theta) - C(F_1(y_1 - 1), F_2(y_2);\ \theta) \\ &\quad - C(F_1(y_1), F_2(y_2 - 1);\ \theta) + C(F_1(y_1 - 1), F_2(y_2 - 1);\ \theta) \end{aligned} \tag{2.2.5}$$

The $C(\cdot,\ \cdot;\ \theta)$ could be any copula presented in Section II.2.1. The copula parameter θ is the correlation coefficient γ for the Gaussian copula, and it is α for the Clayton, Frank, and Gumbel copulas. For binary variables $C(F_1(y_1), F_2(y_2);\ \theta) = 0$, if any y's are less than zero; otherwise if $y_1 = 1$, then $C(F_1(y_1), F_2(y_2);\ \theta) = F_2(y_2)$, and if $y_2 = 1$, then $C(F_1(y_1), F_2(y_2);\ \theta) = F_1(y_1)$; if both y_1 and y_2 are zero, $C(F_1(y_1),\ F_2(y_2);\ \theta) = C(q_1,\ q_2;\ \theta)$, since $F_i(0) = P(Y_i = 0) = q_i$ for $i = 1, 2$. Therefore the joint probability mass function in the bivariate case can be represented as in Table 3.

Table 3: PMF of bivariate binary variables

$(Y_1,\ Y_2)$	**Probability**
(0, 0)	$C(q_1, q_2;\ \theta)$
(0, 1)	$q_1 - C(q_1, q_2;\ \theta)$
(1, 0)	$q_2 - C(q_1, q_2;\ \theta)$
(1, 1)	$1 - q_1 - q_2 + C(q_1, q_2;\ \theta)$

NOTE: there are 3 parameters needed for the distribution of bivariate binary variables: marginal means p_1, p_2 and copula parameter θ.

II.2.3 TRIVARIATE BINARY DISTRIBUTIONS

In this section we will extend the pair-copula method to construct three dimensional binary distributions. Let Y_1, Y_2 and Y_3 be three dependent binary random variables. Note that

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = P(Y_2 = y_2) * P(Y_1 = y_1, Y_3 = y_3 | Y_2 = y_2) \tag{2.2.6}$$

The above equation shows the three dimension distribution can be obtained by constructing the bivariate conditional distribution of $(Y_1,\ Y_3)$ given $Y_2 = y_2$. To this end we introduce some notation. We first construct bivariate distributions for $(Y_1,\ Y_2)$ and $(Y_2,\ Y_3)$ using bivariate copulas $C_{12}(\cdot,\ \cdot)$ and $C_{23}(\cdot,\ \cdot)$ as in Table 3.

Let $q_{1|0} = P(Y_1 = 0|Y_2 = 0) = \frac{C_{12}(q_1,q_2)}{q_2}$ and $p_{1|0} = 1 - q_{1|0} = P(Y_1 = 1|Y_2 = 0) = 1 - \frac{C_{12}(q_1,q_2)}{q_2}$. Thus $Y_1|Y_2 = 0$ is distributed as Bernoulli with mean $p_{1|0}$. Similarly, $Y_3|Y_2 = 0$ is distributed as Bernoulli with mean $p_{3|0}$, where $p_{3|0} = 1 - \frac{C_{23}(q_2,q_3)}{q_2}$. We also have $Y_1|Y_2 = 1$ is Bernoulli with mean $p_{1|1} = 1 - \frac{q_1 - C_{12}(q_1,q_2)}{p_2}$, and $Y_3|Y_2 = 1$ is Bernoulli$(p_{3|1})$, where $p_{3|1} = 1 - \frac{q_3 - C_{23}(q_2,q_3)}{p_2}$.

Table 4: PMF of bivariate binary variables given y_2

	Probability
$(Y_1, Y_3)\|Y_2 = 0$	
(0, 0)	$C_{13\|0}(q_{1\|0}, q_{3\|0}; \theta_{13\|Y_2=0})$
(0, 1)	$q_{1\|0} - C_{13\|0}(q_{1\|0}, q_{3\|0}; \theta_{13\|Y_2=0})$
(1, 0)	$q_{3\|0} - C_{13\|0}(q_{1\|0}, q_{3\|0}; \theta_{13\|Y_2=0})$
(1, 1)	$1 - q_{1\|0} - q_{3\|0} + C_{13\|0}(q_{1\|0}, q_{3\|0}; \theta_{13\|Y_2=0})$
$(Y_1, Y_3)\|Y_2 = 1$	
(0, 0)	$C_{13\|1}(q_{1\|1}, q_{3\|1}; \theta_{13\|Y_2=1})$
(0, 1)	$q_{1\|1} - C_{13\|1}(q_{1\|1}, q_{3\|1}; \theta_{13\|Y_2=1})$
(1, 0)	$q_{3\|1} - C_{13\|1}(q_{1\|1}, q_{3\|1}; \theta_{13\|Y_2=1})$
(1, 1)	$1 - q_{1\|1} - q_{3\|1} + C_{13\|1}(q_{1\|1}, q_{3\|1}; \theta_{13\|Y_2=1})$

NOTE: $P(Y_1, Y_3|Y_2 = 0)$ needs 6 parameters: marginal means p_1, p_2, p_3 and copula parameter θ_{12}, θ_{23}, $\theta_{13|Y_2=0}$, $P(Y_1, Y_3|Y_2 = 1)$ needs the same parameters except the conditional correlation $\theta_{13|Y_2=1}$.

Table 4 show the conditional distributions of $(Y_1,\ Y_3)|Y_2 = 0$ and $(Y_1,\ Y_3)|Y_2 = 1$, respectively. Finally from Equation (2.2.6) and using the conditonal distributions we can get the joint trivariate PMF as given in table as below. For notational convenience we omit the copula parameter in some formulas.

Table 5: PMF of trivariate binary variables

(Y_1, Y_2, Y_3)	**Probability**
(0, 0, 0)	$q_2 * C_{13\|0}(q_{1\|0}, q_{3\|0})$
(0, 0, 1)	$C_{12}(q_1, q_2) - q_2 * C_{13\|0}(q_{1\|0}, q_{3\|0})$
(0, 1, 0)	$p_2 * C_{13\|1}(q_{1\|1}, q_{3\|1})$
(0, 1, 1)	$q_1 - C_{12}(q_1, q_2) - p_2 * C_{13\|1}(q_{1\|1}, q_{3\|1})$
(1, 0, 0)	$C_{23}(q_2, q_3) - q_2 * C_{13\|0}(q_{1\|0}, q_{3\|0})$
(1, 0, 1)	$q_2 - C_{23}(q_2, q_3) - C_{12}(q_1, q_2) + q_2 * C_{13\|0}(q_{1\|0}, q_{3\|0})$
(1, 1, 0)	$q_3 - C_{23}(q_2, q_3) - p_2 * C_{13\|1}(q_{1\|1}, q_{3\|1})$
(1, 1, 1)	$1 - q_1 - q_2 - q_3 + C_{12}(q_1, q_2) + C_{23}(q_2, q_3) + p_2 * C_{13\|1}(q_{1\|1}, q_{3\|1})$

NOTE: there are 7 parameters needed here: marginal means p_1, p_2, p_3 and copula parameter θ_{12}, θ_{23}, $\theta_{13|Y_2=0}$, $\theta_{13|Y_2=1}$.

Table 6: Summary of parameter values for PMF of the trivariate binary variables

Case	**Pair-copulas**	**Dependence**
1	All Gaussian	$\gamma_{12} = 0.752$, $\gamma_{23} = 0.607$
		$\gamma_{13\|Y_2=0} = 0.480$, $\gamma_{13\|Y_2=1} = 0.233$
2	All Clayton	$\alpha_{12} = 2$, $\alpha_{23} = 1.5$
		$\alpha_{13\|Y_2=0} = \alpha_{13\|Y_2=1} = 0.4$
3	All Frank	$\alpha_{12} = \alpha_{23} = 1.85$
		$\alpha_{13\|Y_2=0} = 0.95$, $\alpha_{13\|Y_2=1} = 0.85$
4	All Gumbel	$\alpha_{12} = \alpha_{23} = 10$
		$\alpha_{13\|Y_2=0} = \alpha_{13\|Y_2=1} = 4$
5	Gaussian for tree 1	$\gamma_{12} = 0.752$, $\gamma_{23} = 0.607$
	Frank for tree 2	$\alpha_{13\|Y_2=0} = 0.95$, $\alpha_{13\|Y_2=1} = 0.85$
6	All Independent	

Example is shown considering the following cases: if marginal means are $p = (0.8,\ 0.7,\ 0.6)$, and the dependence information is summarized in the Table 6 below.

Take the case 5 for example, we start with Table 3 using Gaussian copula with $\gamma_{12} = 0.752$, $\gamma_{23} = 0.607$, and $p = (0.8,\ 0.7,\ 0.6)$. The PMF of bivariate binary variables are in Table 7.

Table 7: PMF of bivariate binary variables using Gaussian copula

	Probability
$(Y_1,\ Y_2)$	
$(0,\ 0)$	$C(q_1, q_2;\ \gamma_{12}) = \Phi_2(\Phi^{-1}(0.2), \Phi^{-1}(0.3); 0.752) = 0.1517$
$(0,\ 1)$	$q_1 - C(q_1, q_2;\ \gamma_{12}) = 0.2 - 0.1517 = 0.0483$
$(1,\ 0)$	$q_2 - C(q_1, q_2;\ \gamma_{12}) = 0.3 - 0.1517 = 0.1483$
$(1,\ 1)$	$1 - q_1 - q_2 + C(q_1, q_2;\ \gamma_{12}) = 1 - 0.2 - 0.3 + 0.1517 = 0.6517$
$(y_2,\ y_3)$	
$(0,\ 0)$	$C(q_2, q_3;\ \gamma_{23}) = \Phi_2(\Phi^{-1}(0.3), \Phi^{-1}(0.4); 0.607) = 0.2097$
$(0,\ 1)$	$q_2 - C(q_2, q_3;\ \gamma_{23}) = 0.3 - 0.2097 = 0.0903$
$(1,\ 0)$	$q_3 - C(q_2, q_3;\ \gamma_{23}) = 0.4 - 0.2097 = 0.1903$
$(1,\ 1)$	$1 - q_2 - q_3 + C(q_2, q_3;\ \gamma_{23}) = 1 - 0.3 - 0.4 + 0.2097 = 0.5097$

Now, in order to have the conditional PMF of bivariate variables, we need to get the marginal mean of new Bernoulli variables $Y_1|Y_2 = 0$ and $Y_1|Y_2 = 1$.

$$\begin{aligned}
q_{1|0} &= \frac{P(Y_1 = 0, Y_2 = 0)}{q_2} = \frac{0.1517}{0.3} = 0.5057 \\
q_{3|0} &= \frac{P(Y_2 = 0, Y_3 = 0)}{q_2} = \frac{0.2097}{0.3} = 0.6990 \\
q_{1|1} &= \frac{P(Y_1 = 0, Y_2 = 1)}{p_2} = \frac{0.0483}{0.7} = 0.0690 \\
q_{3|1} &= \frac{P(Y_2 = 1, Y_3 = 0)}{p_2} = \frac{0.1903}{0.7} = 0.2719
\end{aligned}$$

Then, $p_{1|0} = 1 - 0.5057 = 0.4943$, $p_{3|0} = 1 - 0.6990 = 0.3010$, $p_{1|1} = 1 - 0.0690 = 0.931$ and $p_{3|1} = 1 - 0.2719 = 0.7281$. Also, Frank copula is used for tree 2 with

$\alpha_{13|Y_2=0} = 0.95$, $\alpha_{13|Y_2=1} = 0.85$. The PMF of conditional bivariate binary variables are caculated according to Table 4.

Table 8: Conditional PMF of bivariate binary variables using Frank copula

	Probability
$(Y_1,\ Y_3 \mid Y_2 = 0)$	
$(0,\ 0)$	$C_{13\mid 0}(q_{1\mid 0}, q_{3\mid 0}; \theta_{13\mid Y_2=0})$
	$= -\frac{1}{0.95}\log(1 + \frac{(e^{-0.95*0.5057}-1)(e^{-0.95*0.699}-1)}{e^{-0.95}-1}) = 0.3782$
$(0,\ 1)$	$q_{1\mid 0} - C_{13\mid 0}(q_{1\mid 0}, q_{3\mid 0}; \theta_{13\mid Y_2=0}) = 0.5057 - 0.3782 = 0.1275$
$(1,\ 0)$	$q_{3\mid 0} - C_{13\mid 0}(q_{1\mid 0}, q_{3\mid 0}; \theta_{13\mid Y_2=0}) = 0.699 - 0.3782 = 0.3208$
$(1,\ 1)$	$1 - q_{1\mid 0} - q_{3\mid 0} + C_{13\mid 0}(q_{1\mid 0}, q_{3\mid 0}; \theta_{13\mid Y_2=0}) = 0.1735$
$(Y_1,\ Y_3 \mid Y_2 = 1)$	
$(0,\ 0)$	$C_{13\mid 1}(q_{1\mid 1}, q_{3\mid 1}; \theta_{13\mid Y_2=1})$
	$= -\frac{1}{0.85}\log(1 + \frac{(e^{-0.85*0.069}-1)(e^{-0.85*0.2719}-1)}{e^{-0.85}-1}) = 0.0244$
$(0,\ 1)$	$q_{1\mid 1} - C_{13\mid 1}(q_{1\mid 1}, q_{3\mid 1}; \theta_{13\mid Y_2=1}) = 0.069 - 0.0244 = 0.0446$
$(1,\ 0)$	$q_{3\mid 1} - C_{13\mid 1}(q_{1\mid 1}, q_{3\mid 1}; \theta_{13\mid Y_2=1}) = 0.2719 - 0.0244 = 0.2475$
$(1,\ 1)$	$1 - q_{1\mid 1} - q_{3\mid 1} + C_{13\mid 1}(q_{1\mid 1}, q_{3\mid 1}; \theta_{13\mid Y_2=1}) = 0.6835$

Since $P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = P(Y_2 = y_2) * P(Y_1 = y_1, Y_3 = y_3|Y_2 = y_2)$, the last step is to multiply $P(Y_2 = y_2)$ to the table above to get the results as Table 5, for example, $P(Y_1 = 0, Y_2 = 0, Y_3 = 0) = 0.3782 * 0.3 = 0.1135$ or $P(Y_1 = 1,\ Y_2 = 1,\ Y_3 = 0) = 0.2475 * 0.7 = 0.1732$. PMF using other copulas can be obtained by the similar steps as above, and the results are presented in Table 9.

This example shows that different copula leads to different result. But for the AR(1) or equicorrelated structure binary variables as explained in Section II.2.4, we will choose the parameters to make correlation structure as required, thus, no matter which copula we use, we end up with the same joint distribution. Therefore, We will just use Gaussian copula in the rest of the thesis.

Four or higher dimensional multivariate binary distributions can be constructed in a similar fashion. We present the PMF's for four and five dimensions in Appendix B. These distributions are constructed using D-vines, since the second tree is conditioning on Y_2. We could also use a C-vine which involves conditioning on Y_1.

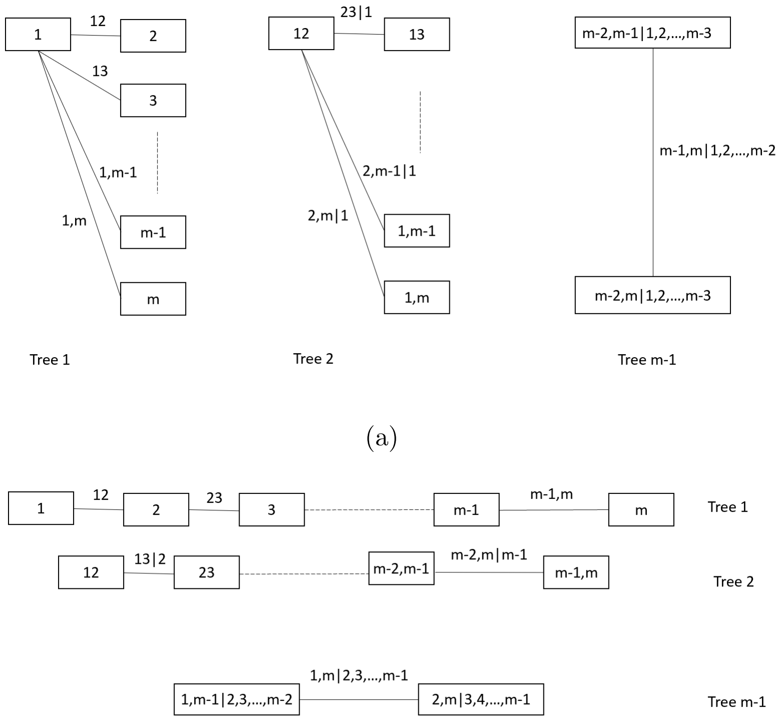
The tree structures below show the differences in general between D and C-vines. The vine decomposition is given in nested trees, for the ith tree there are $m-i$ nodes, represented by rectangular boxes in Figure 2.

Table 9: PMF of trivariate binary variables

(Y_1, Y_2, Y_3)	PMF_1	PMF_2	PMF_3	PMF_4	PMF_5	PMF_6
(0, 0, 0)	0.1267	0.1366	0.0582	0.1983	0.1135	0.024
(0, 0, 1)	0.0250	0.0322	0.0337	0.0000	0.0382	0.036
(0, 1, 0)	0.0210	0.0190	0.0450	0.0017	0.0171	0.056
(0, 1, 1)	0.0273	0.0122	0.0631	0.0000	0.0312	0.084
(1, 0, 0)	0.0830	0.0939	0.1077	0.0994	0.0963	0.096
(1, 0, 1)	0.0653	0.0373	0.1004	0.0023	0.0520	0.144
(1, 1, 0)	0.1693	0.1505	0.1891	0.1006	0.1732	0.224
(1, 1, 1)	0.4824	0.5182	0.4028	0.5977	0.4785	0.336

The main difference between C-vine and D-vine is that, for each tree there is one and only one node in C-vine that is connected to all the other nodes. Whereas there is no node in D-vine connected to more than two nodes. The C-vine structure is appropriate if one variable plays a central role, for example familial data where there is a head of the household. The D-vine is appropriate for longitudinal data where there is a natural sequence for the variables. Using D-vine, the joint distribution of the binary variables $(y_1, ..., y_m)$ with marginal mean $(p_1, ..., p_n)$ is as below,

$$\begin{aligned} P(y_1, ..., y_m) &= P(y_1, y_m|y_2, .., y_{m-1})P(y_2, y_{m-1}|y_3, .., y_{m-2})...P(y_{\frac{m+1}{2}}), \quad (2.2.7) \\ & \qquad \text{if } m \text{ is odd;} \\ &= P(y_1, y_m|y_2, .., y_{m-1})P(y_2, y_{m-1}|y_3, .., y_{m-2})...P(y_{\frac{m}{2}}, y_{\frac{m}{2}+1}), \\ & \qquad \text{if } m \text{ is even.} \end{aligned}$$



(a)

(b)

Figure 2: m-dimensional vine structures: (a) C-vine, (b) D-vine.

II.2.4 CORRELATION STRUCTURE FOR DEPENDENT BINARY VARIABLES

Longitudinal data consists of a series of observations and tend to be serially correlated. A common correlation model for the serial correlation is the first order autoregressive (AR(1)) structure. However, the equicorrelated structure (EQC) is also used when there is a small sequence of observations. The AR(1) structure arises when given the past and the present, the future depends on the present and not the past. Correlations needed for tree 1, either AR(1) or equicorrelated structure, can be solved using Equations (2.2.1)–(2.2.4).

The pair-copula model involves conditional ditributions which involve partial correlations. Therefore the above formulas need to adjusted for the partial correlations $\mathrm{Corr}(Y_1, Y_3|Y_2)$, $\mathrm{Corr}(Y_2, Y_4|Y_3)$, ..., $\mathrm{Corr}(Y_{m-2}, Y_m|Y_{m-1})$ for tree 2; $\mathrm{Corr}(Y_1, Y_4|Y_2, Y_3)$, ..., $\mathrm{Corr}(Y_{m-3}, Y_m|Y_{m-1}, Y_{m-2})$ for tree 3; ...., $\mathrm{Corr}(Y_1, Y_m|Y_2, ...Y_{m-1})$ for tree $m-1$.

For AR(1) structure, Poddar (2016) has shown that $\mathrm{Corr}(Y_i, Y_{i+k}|Y_{i+1}, ...Y_{i+k-1})$ is 0 for multivariate normal variables, but this conditional correlation depends on the values of Y_{i+1}, ..., Y_{i+k-1} for multivariate binary variables. For pair-copula model we make the assumption that $\mathrm{Corr}(Y_i, Y_{i+k}|Y_{i+1}, ...Y_{i+k-1}) = 0$ for the binary variables. We will see that numerically this leads to an approximate AR(1) structure when we use the Gaussian copula. For equicorrelated structure, $\mathrm{Corr}(Y_i, Y_{i+k}|Y_{i+1}, ...Y_{i+k-1}) = \rho/(1 + (k-1)\rho)$ for multivariate normal variables as proved in Appendix D. We make this assumption for constructing multivariate binary distribution with an approximate equicorrelation structure. We will show that numerically with these assumptions, we can generate PMF of approximately AR(1) or equicorrelated structured multivariate binary variables in Section II.4.3.

Since we will choose the copula parameters to make correlation structure as required, for example, binary variable correlation is ρ for tree 1, conditional correlations are zero for AR(1), or fixed as $\rho/(1 + (k-1)\rho)$ for equicorrelated structure, thus, no matter which copula we use, we end up with the same joint distribution. Therefore, we will use only Gaussian copula in the rest of this dissertation.

II.2.5 PMF FOR AR(1) STRUCTURED DEPENDENT BINARY DATA

For AR(1) structured binary data, we assume that $\mathrm{Corr}(Y_i, Y_{i+k}|Y_{i+1}, ...Y_{i+k-1}) = 0$, and for the Gaussian copula zero correlation implies independence. Hence we could use independence copula to generate the conditional distribution of Y_i and Y_{i+k} given $Y_{i+1}, ...Y_{i+k-1})$. Thus

$$\begin{aligned} f(y_i = 0, y_{i+k} = 0|i+1, ...i+k-1) &= f(y_{i|i+1,...i+k-1} = 0)f(y_{i+k|i+1,...i+k-1} = 0) \\ &= q_{i|i+1,...i+k-1}\; q_{i+k|i+1,...i+k-1} \end{aligned}$$

Similarly,

$$\begin{aligned} f(y_i = 0, y_{i+k} = 1|i+1, ...i+k-1) &= f(y_{i|i+1,...i+k-1} = 0)f(y_{i+k|i+1,...i+k-1} = 1) \\ &= q_{i|i+1,...i+k-1}\; p_{i+k|i+1,...i+k-1} \end{aligned}$$

and

$$
\begin{aligned}
f(y_i = 1, y_{i+k} = 0|i+1, ...i+k-1) &= p_{i|i+1,...i+k-1}\ q_{i+k|i+1,...i+k-1} \\
f(y_i = 1, y_{i+k} = 1|i+1, ...i+k-1) &= p_{i|i+1,...i+k-1}\ p_{i+k|i+1,...i+k-1}
\end{aligned}
$$

Therefore, we can replace the generalization of the conditional PMF in Section II.2.3 with the one shown in Table 10.

Table 10: Conditional PMF for AR(1) structure

$(y_i, y_{i+k})\|y_{i+1}, ...y_{i+k-1}$	**Probability**
(0,0)	$q_{i\|i+1,...i+k-1}q_{i+k\|i+1,...i+k-1}$
(0,1)	$q_{i\|i+1,...i+k-1}p_{i+k\|i+1,...i+k-1}$
(1,0)	$p_{i\|i+1,...i+k-1}q_{i+k\|i+1,...i+k-1}$
(1,1)	$p_{i\|i+1,...i+k-1}p_{i+k\|i+1,...i+k-1}$

II.2.6 BOUNDARIES OF THE BINARY VARIABLE CORRELATION

For binary variables with given marginal means, the correlation parameter has bounds which are known as Fréchet bounds, see Chaganty and Joe (2004). These bounds that depend on the correlation structure define the feasible range for the correlation. For the AR(1) structure the bounds for ρ are

$$\max_{2\leq t\leq m} L(p_{(t-1)}, p_t) \leq \rho \leq \min_{2\leq t\leq m} U(p_{(t-1)}, p_t) \tag{2.2.8}$$

where m is the dimension and

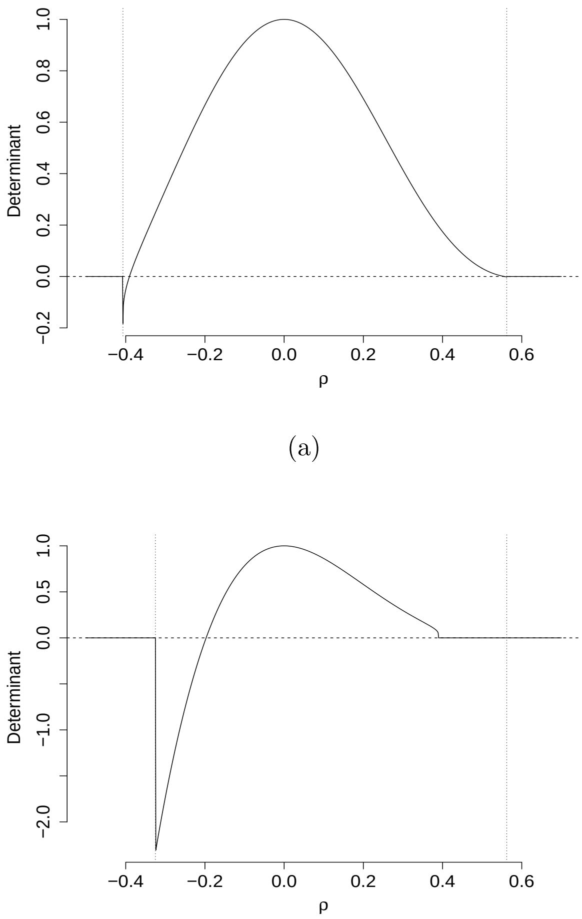
$$
\begin{aligned}
L(p_i, p_j) &= \max\left(-\sqrt{\frac{p_i p_j}{(1-p_i)(1-p_j)}}, -\sqrt{\frac{(1-p_i)(1-p_j)}{p_i p_j}}\right), \quad \text{and} \\
U(p_i, p_j) &= \min\left(\sqrt{\frac{p_i(1-p_j)}{(1-p_i)p_j}}, \sqrt{\frac{(1-p_i)p_j}{p_i(1-p_j)}}\right), \quad \text{for } 0 < p_i < 1,\ 0 < p_j < 1.
\end{aligned}
$$

For the equicorrelated structure, the bounds on ρ for any dimension m is an unsolved problem. However, when $m = 3$ Chaganty and Joe (2006) showed that

$$\max_{1\leq i<j\leq 3}(L_1(p_1, p_2, p_3), L(p_i, p_j)) \leq \rho \leq \min_{1\leq i<j\leq 3} U(p_i, p_j), \tag{2.2.9}$$

where the function L_1 is,

$$L_1(p_1, p_2, p_3) = \frac{-(p_1p_2p_3 + q_1q_2q_3)}{\sqrt{p_1q_1p_2q_2} + \sqrt{p_1q_1p_3q_3} + \sqrt{p_2q_2p_3q_3}}, \quad 0 < p_i < 1.$$



(a)

(b)

Figure 3: Comparison plotting of correlation feasible range vs. determinant of latent correlation matrix for (a) AR(1) and (b)Equicorrelated structure.

For instance, suppose $y = (y_1, y_2, y_3, y_4)$ is a four-dimensional binary vector with marginal mean $p = (0.43, 0.22, 0.37, 0.65)$. The feasible range for the correlation parameter ρ assuming a AR(1) structure is $(-0.4070, 0.5624)$. For the equicorrelated structure the range is a subset of $(-0.3248, 0.5624)$, where the lower bound is calculated as

$$\max_{1\le i<j\le 3}(L_1(p_1,p_2,p_3),L_1(p_1,p_2,p_4),L_1(p_1,p_3,p_4),L_1(p_2,p_3,p_4),L(p_i,p_j)).$$

A binary distribution does not exist for values of ρ outside these ranges. And even for values of ρ withing this feasible range, the multivaraite probit (MP) model may fail since the correlation matrix obtained solving Equation 2.2.1 may not lead to a positive definite matrix. Figure 3 has plot of the feasible range and the determinant of the correlation matrix associated with the Gaussian copula. We can see from the figure for values of ρ closer to the boundary, the determinant is negative, and so the MP model fails to generate a legitimate probability distribution of the binary variables.

However, the pair-copula model proposed in this chapter could generate a multivariate binary distribution for a feasible values of the correlation coefficient in cases where the MP model fails. We present some numerical examples in Section II.4.4 to illustrate this more concretely. And also present the PMF of the binary distributions in Section II.4.3, for values of ρ within the feasible range both for AR(1) and equicorrelated structures.

II.3 USE OF HERMITE POLYNOMIALS

A common method of solving Equation (2.2.1) to get γ for a given ρ is the root finding function “uniroot” in R to search the domain interval $(-1,1)$ for a solution. In a recent paper Xiao and Zhou (2019) suggested an alternative method employing Hermite polynomials which is more efficient and accurate. Following Xiao and Zhou (2019), we write Equation (2.2.1) as

$$\begin{aligned}\rho &= G(\gamma)\\ &= -\frac{p_1p_2}{\sqrt{p_1q_1p_2q_2}}+\frac{1}{\sqrt{p_1q_1p_2q_2}}\int_{\Phi^{-1}(q_1)}^{\infty}\int_{\Phi^{-1}(q_2)}^{\infty}\phi_2(z_1,z_2;\ \gamma)dz_1\,dz_2,\end{aligned}\tag{2.3.10}$$

where ϕ_2 is the density of standard bivariate normal with correlation γ. The double intergral on the right of Equation (2.3.10), is taken over the rectangular region $\{(z_1,z_2)|\Phi^{-1}(q_1)\le z_1<\infty,\Phi^{-1}(q_2)\le z_2<\infty\}$. Taking complements we can rewrite as

$$\rho \;=\; G(\gamma)=-\frac{p_1p_2}{\sqrt{p_1q_1p_2q_2}}+\frac{1}{\sqrt{p_1q_1p_2q_2}}(1-q_1-q_2+\Phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_2);\gamma))\tag{2.3.11}$$

The idea is to use a Taylor series expansion of $G(\gamma)$ around zero. This requires the nth order derivative of $\Phi_2(\Phi^{-1}(q_1), \Phi^{-1}(q_2); \gamma)$ at zero. Xiao and Zhou (2019) have shown that the nth order derivative is given by

$$\Phi_2^{(n)}(\Phi^{-1}(q_1), \Phi^{-1}(q_1); 0) = H_{n-1}(\Phi^{-1}(q_1))\ \phi(\Phi^{-1}(q_1))\ H_{n-1}(\Phi^{-1}(q_2))\ \phi(\Phi^{-1}(q_2)) \tag{2.3.12}$$

where H_j's are Hermite polynomials defined as

$$\begin{aligned} H_0(z) &= 1 \\ H_1(z) &= z \\ H_2(z) &= z^2 - 1 \\ H_3(z) &= z^3 - 3z \\ &\cdots \\ H_{k+1}(z) &= zH_k(z) - H_k'(z) \qquad \text{for } k \geq 1 \end{aligned}$$

Therefore,

$$\begin{aligned} G(0) &= 0 \\ G'(0) &= \frac{1}{\sqrt{p_1q_1p_2q_2}}\Phi_2'(\Phi^{-1}(q_1), \Phi^{-1}(q_1); 0) \\ &= \frac{1}{\sqrt{p_1q_1p_2q_2}}H_0(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_1))H_0(\Phi^{-1}(q_2))\phi(\Phi^{-1}(q_2)) \\ &= \frac{\phi(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_2))}{\sqrt{p_1q_1p_2q_2}} \\ G^{(2)}(0) &= \frac{1}{\sqrt{p_1q_1p_2q_2}}\Phi_2^{(2)}(\Phi^{-1}(q_1), \Phi^{-1}(q_1); 0) \\ &= \frac{1}{\sqrt{p_1q_1p_2q_2}}H_1(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_1))H_1(\Phi^{-1}(q_2))\phi(\Phi^{-1}(q_2)) \qquad (2.3.13) \\ &= \frac{\phi(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_2))}{\sqrt{p_1q_1p_2q_2}}\Phi^{-1}(q_1)\Phi^{-1}(q_2) \\ G^{(3)}(0) &= \frac{1}{\sqrt{p_1q_1p_2q_2}}\Phi_2^{(3)}(\Phi^{-1}(q_1), \Phi^{-1}(q_1); 0) \\ &= \frac{1}{\sqrt{p_1q_1p_2q_2}}H_2(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_1))H_2(\Phi^{-1}(q_2))\phi(\Phi^{-1}(q_2)) \\ &= \frac{\phi(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_2))}{\sqrt{p_1q_1p_2q_2}}(\Phi^{-1}(q_1)^2 - 1)(\Phi^{-1}(q_2)^2 - 1) \\ &\cdots \end{aligned}$$

$$
\begin{aligned}
G^{(k+2)}(0) &= \frac{1}{\sqrt{p_1q_1p_2q_2}}\Phi_2^{(k+2)}(\Phi^{-1}(q_1), \Phi^{-1}(q_1); 0)\\
&= \frac{1}{\sqrt{p_1q_1p_2q_2}}H_{k+1}(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_1))H_{k+1}(\Phi^{-1}(q_2))\phi(\Phi^{-1}(q_2))\\
&= \frac{\phi(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_2))}{\sqrt{p_1q_1p_2q_2}}(\Phi^{-1}(q_1)H_k(\Phi^{-1}(q_1)) - H_k'(\Phi^{-1}(q_1)))\\
&(\Phi^{-1}(q_2)H_k(\Phi^{-1}(q_2)) - H_k'(\Phi^{-1}(q_2)))\\
&\cdots
\end{aligned}
$$

The Taylor expansion of $\rho = G(\gamma)$ becomes,

$$
\begin{aligned}
\rho &= G(0) + \frac{G'(0)}{1!}\gamma + \frac{G^{(2)}(0)}{(2)!}\gamma^2 + \cdots + \frac{G^{(k+2)}(0)}{(k+2)!}\gamma^{k+2} + \cdots\\
&= \frac{\phi(\Phi^{-1}(q_1))\phi(\Phi^{-1}(q_2))}{\sqrt{p_1q_1p_2q_2}}\{\gamma + \frac{\Phi^{-1}(q_1)\Phi^{-1}(q_2)}{2}\gamma^2\\
&+ \cdots + \{(\Phi^{-1}(q_1)H_k(\Phi^{-1}(q_1)) - H_k'(\Phi^{-1}(q_1)))(\Phi^{-1}(q_2)H_k(\Phi^{-1}(q_2))\\
&- H_k'(\Phi^{-1}(q_2)))\}/(k+2)!\gamma^{k+2}\} + \cdots
\end{aligned}
\tag{2.3.14}
$$

Therefore, the coefficients of Equation (2.3.14) can be determined given p_i's, and γ can be solved for a given ρ.

A comparison plotting is shown in Figure 4, with binary variable coefficients on the x-axis and latent variable coefficients on y-axis. For the three cases considered, we can see from Figure 4, Hermite approximation is as good as solving Equation (2.2.1) directly when ρ is close to zero. But it starts to deviate more from the true solution of Equation (2.2.1), when ρ is at the boundary. It makes sense, because the Taylor expansion used for function $G()$ is around the point $\gamma = 0$, and the approximation more accurate around that point.

Now, in order to improve the estimation performance, we would like to do Taylor expansion around an arbitrary point $-1 < a < 1$. The first step of such adjustment is to find the nth-order derivative of bivariate Gaussian copula at $\gamma = a$ similar to Equation (2.3.12). We can let a equals the estimated value γ_0 from Equation (2.3.14) to get a more accurate estimation. According to Viskov (2008), Mehler's formula can be expressed using Hermite polynomials as below,

$$\begin{aligned}\frac{1}{\sqrt{1-(\gamma-a)^2}} & \quad \exp\{-\tfrac{(\gamma-a)^2 z_1^2-2(\gamma-a)z_1z_2+(\gamma-a)^2z_2^2}{2(1-(\gamma-a)^2)}\} \qquad (2.3.15)\\ & = \textstyle\sum_{k=0}^{\infty} H_k(z_1)H_k(z_2)\frac{(\gamma-a)^k}{k!}\\ \phi_2(z_1,z_2,\gamma) & = \tfrac{1}{2\pi}\sqrt{\tfrac{1-(\gamma-a)^2}{1-\gamma^2}}\exp\{-\tfrac{\gamma^2z_1^2-2\gamma z_1z_2+\gamma^2z_2^2}{2(1-\gamma^2)} \qquad (2.3.16)\\ & \quad +\tfrac{(\gamma-a)^2z_1^2-2(\gamma-a)z_1z_2+(\gamma-a)^2z_2^2}{2(1-(\gamma-a)^2)}\}\\ & \quad \textstyle\sum_{k=0}^{\infty} H_k(z_1)H_k(z_2)\frac{(\gamma-a)^k}{k!}\\ \phi_2(z_1,z_2,\gamma) & = \phi_2(z_1,z_2,a)\textstyle\sum_{k=0}^{\infty} H_k(z_1)H_k(z_2)\frac{(\gamma-a)^k}{k!}\end{aligned}$$

The Taylor expansion of $\phi_2(z_1, z_2, \gamma)$ about point $\gamma = a$ is as below,

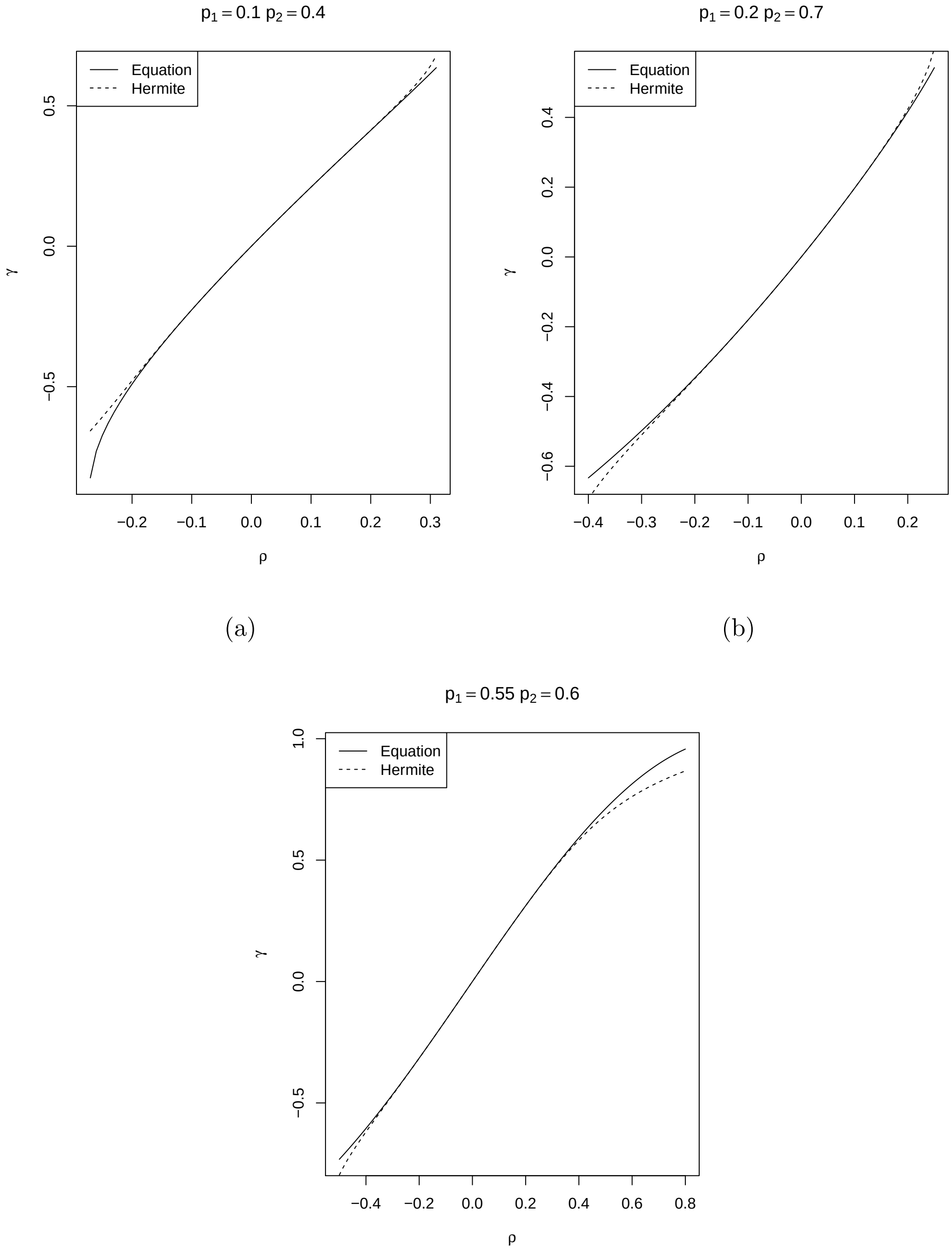
$$\phi_2(z_1,z_2,\gamma)=\sum_{k=0}^{\infty}\phi_2^{(k)}(z_1,z_2,a)\frac{(\gamma-a)^k}{k!} \qquad (2.3.17)$$

Comparing Equations (2.3.15) and (2.3.17), we can solve for the polynomial coefficients and obtain

$$\begin{aligned}G(a) &= \frac{-q_1q_2+\Phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_1);a)}{\sqrt{p_1q_1p_2q_2}}\\ G'(a) &= \frac{\phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_1);a)}{\sqrt{p_1q_1p_2q_2}}\\ G^{(2)}(a) &= \frac{\phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_1);a)}{\sqrt{p_1q_1p_2q_2}}\Phi^{-1}(q_1)\Phi^{-1}(q_2)\\ G^{(3)}(a) &= \frac{\phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_2);a)}{\sqrt{p_1q_1p_2q_2}}(\Phi^{-1}(q_1)^2-1)(\Phi^{-1}(q_2)^2-1)\\ \cdots\\ G^{(k+2)}(a) &= \frac{\phi_2(\Phi^{-1}(q_1),\Phi^{-1}(q_2);a)}{\sqrt{p_1q_1p_2q_2}}(\Phi^{-1}(q_1)H_k(\Phi^{-1}(q_1))-H_k'(\Phi^{-1}(q_1)))\\ & \quad (\Phi^{-1}(q_2)H_k(\Phi^{-1}(q_2))-H_k'(\Phi^{-1}(q_2)))\\ \cdots\end{aligned}$$

Therefore,

$$\begin{aligned}\rho = \quad & G(a)+\frac{G'(a)}{1!}(\gamma-a)+\frac{G^{(2)}(a)}{(2)!}(\gamma-a)^2+\cdots\\ & \qquad +\frac{G^{(k+2)}(a)}{(k+2)!}(\gamma-a)^{k+2}+\cdots\end{aligned}$$



(a)

(b)

(c)

Figure 4: Plots of relationship between binary variable correlation coefficient ρ and latent variable correlation coefficient γ. (a) p=(0.1, 0.4); (b) p=(0.2, 0.7); (c) p=(0.55, 0.6)

To get an improved solution, we first get an initial value γ_0 using Equation (2.3.14). Then we set $a = \gamma_0$, and obtain a more accurate value for γ from Equation (2.3.18). For example, when $p = (0.1, 0.4)$ and $\rho = -0.25$, $\gamma_0 = -0.608$ leads to an imrpoved approximation as $\gamma = -0.669$, which is very close to the real correlation as -0.674; when $p = (0.55, 0.6)$ and $\rho = 0.7$, $\gamma_0 = 0.8211$ leads to $\gamma = 0.9052$, which is very close to the real correlation as 0.8966.

II.4 COMPARISON OF PAIR-COPULA AND MP MODELS

The binary D-vine pair-copula model with Gaussian copulas and the multivariate probit model are not the same, since m-dimensional vine pair-copula model needs $m * (m-1)/2$ bivariate copulas and that many bivariate rectangular probabilities, while MP model requires m multidimensonal rectangular probabilities. Considering the flexibility of D-vine pair-copula with Gaussian copulas, it provides a good approximation to MP models as shown in Joe (2014), especially for higher dimensions.

II.4.1 INTRODUCTION TO MP MODEL

Let $Y = (Y_1, Y_2, ..., Y_m)$ a vector of binary random variables. Associated with the vector Y, there is a vector of latent variables $Z = (Z_1, Z_2, ...Z_m)$, which is distributed as multivariate normal (MVN), such that $Y_t = 1$ if $Z_t > 0$, and $Y_t = 0$ if $Z_t \leq 0$. Assume $Z_t = \mu_t + \epsilon_t$, where $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_m)$ is MVN$(0, R)$, and R is a correlation matrix. Thus, $p_t = P(y_t = 1) = P(Z_t > 0) = P(\mu_t + \epsilon_t > 0) = \Phi(\mu_t)$, and $q_t = (1 - p_t) = \Phi(-\mu_t)$. The PMF of $Y = (Y_1, Y_2, ..., Y_m)$ is given by

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_m = y_m) = \int_{D_m} ... \int_{D_1} \frac{1}{(2\pi)^{\frac{m}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{\boldsymbol{\epsilon}\ R^{-1}\ \boldsymbol{\epsilon}'}{2} \right)\ d\boldsymbol{\epsilon} \tag{2.4.18}$$

where, $D_t = (-\infty, \mu_t)$ if $y_t = 1$, and $D_t = (\mu_t, \infty)$ if $y_t = 0$. For example,

$$P(Y_1 = 0, Y_2 = 0, Y_3 = 0) \ = \ \int_{\mu_1}^{\infty} \int_{\mu_2}^{\infty} \int_{\mu_3}^{\infty} \phi_3(\epsilon_1, \epsilon_2, \epsilon_3\,;\ R)\ \ d\boldsymbol{\epsilon}$$

$$= \Phi_3(-\mu_1, -\mu_2, -\mu_3; R),$$

$$P(Y_1 = 1, Y_2 = 1, Y_3 = 1) = \int_{-\infty}^{\mu_1}\int_{-\infty}^{\mu_2}\int_{-\infty}^{\mu_3} \phi_3(\epsilon_1, \epsilon_2, \epsilon_3 \,;\, R) \;\; d\boldsymbol{\epsilon}$$

$$= \Phi_3(\mu_1, \mu_2, \mu_3 \,;\, R),$$

where $\Phi_3(\cdot; R)$ is the CDF of trivariate standard normal.

II.4.2 COMPARING PAIR GAUSSIAN COPULA WITH MP MODELS

To compare the pair-copula Gaussian model with MP models, we first look at the case of two dimensions. In this case, taking C_{12} as the bivariate Gaussian copula, the PMF as given in Table 3 is $P(Y_1 = 0, Y_2 = 0) = C_{12}(q_1, q_2; \gamma) = \Phi_2(\Phi^{-1}(q_1), \Phi^{-1}(q_2); \gamma) = \Phi_2(-\mu_1, -\mu_2; \gamma)$, which is identical to the probability under the MP model. Therefore, the probability distributions are the exactly the same for two dimensions. For three dimensions, for the MP model we have

$$P(Y_1 = 0, Y_2 = 0, Y_3 = 0) = \Phi_3(-\mu_1, -\mu_2, -\mu_3; R) \tag{2.4.19}$$

where Φ_3 is the three dimensional standard multivariate normal CDF and the correlation matrix R is given by

$$R = \begin{pmatrix} 1 & \gamma_{12} & \gamma_{13} \\ \gamma_{12} & 1 & \gamma_{23} \\ \gamma_{13} & \gamma_{23} & 1 \end{pmatrix}. \tag{2.4.20}$$

From Table 5, we see that for the pair-copula model we have

$$P(Y_1 = 0,\ Y_2 = 0,\ Y_3 = 0) = q_2\ C_{13|0}(q_{1|0},\ q_{3|0}). \tag{2.4.21}$$

With bivariate Gaussian copulas for the pairs, we have $q_i = \Phi(-\mu_i)$ for $i = 1, 2, 3$ and

$$\begin{aligned} q_{1|0} &= \frac{C_{12}(q_1, q_2)}{q_2} = \frac{\Phi_2(-\mu_1, -\mu_2;\ \gamma_{12})}{\Phi(-\mu_2)} = P(\epsilon_1 < -\mu_1 | \epsilon_2 < -\mu_2), \\ q_{3|0} &= \frac{C_{23}(q_2, q_3)}{q_2} = \frac{\Phi_2(-\mu_2, -\mu_3;\ \gamma_{23})}{\Phi(-\mu_2)} = P(\epsilon_3 < -\mu_3 | \epsilon_2 < -\mu_2), \end{aligned} \tag{2.4.22}$$

where $\boldsymbol{\epsilon} = (\epsilon_1,\, \epsilon_2,\, \epsilon_3)$ is trivariate standard multivariate normal with correlation matrix R in (2.4.20). The quantities $q_{1|0}$ and $q_{3|0}$ are same as the corresponding values for the MP model. Taking $C_{13|0}(\cdot,\cdot)$ as Gaussian copula with correlation $\gamma_{13|0}$ we have

$$P(Y_1 = 0,\ Y_2 = 0,\ Y_3 = 0) = \Phi(-\mu_2)\,\Phi_2\Big(\Phi^{-1}(q_{1|0}), \Phi^{-1}(q_{3|0}); \gamma_{13|0}\Big) \qquad (2.4.23)$$

The parameter $\gamma_{13|0}$ is unrelated to the elements of R given by (2.4.20). Clearly, (2.4.23) is not equal to (2.4.19) and the PMF of the pair-coupla model is different from the MP model. Numerical examples provided in Tables 11 and 12 reaffirm this observation.

II.4.3 NUMERICAL EXAMPLES OF BINARY DISTRIBUTIONS

In this section, we will present some numerical examples of probability mass functions generated by the coupla methods discussed in the previous sections. We generate the PMFs for both the MP and pair-copula models using Gaussian copula with given marginal means, and for both correlation structures AR(1) and equicorrelated. For the first example we took the marginal mean vector to be $p = (0.33, 0.26, 0.71)$. The feasible range of the correlation (2.2.8) for the AR(1) structure is $(-0.416, 0.379)$ and for the equicorrelated the range (2.2.9) is $(-0.331, 0.379)$. For our examples we took $\rho = 0.2$ which falls within the feasible range. The PMF's for the AR(1) structure are presented in Table 11 and for the equicorrelated structure in Table 12.

For the AR(1) structure, since coefficient is $\rho = 0.2$, marginal means are $p = (0.33, 0.26, 0.71)$. Then, the the latent correlation matrix is

$$R = \begin{bmatrix} 1 & 0.332 & 0.069 \\ 0.332 & 1 & 0.383 \\ 0.069 & 0.383 & 1 \end{bmatrix},$$

A close examination of Table 11 shows that for the AR(1) structure, the PMF values of the MP model are a little different from pair-copula models. But PMF of the pair-copula model with different copulas are the same, which is consistent with the result that we proved in the last section. Take $y = (0, 0, 0)$ for example, the probabilities are 0.1854 for MP model, 0.1846 for D-vine pair-copula using Gaussian. Also the two dimensional marginals are the same for MP model and pair Gaussian copula model, which is again consistent to the proof in Section II.4.2. For instance, $y = (0, 0) = 0.5370$ for both models.

Table 11: PMF of AR(1) trivariate binary variables generated by the MP model, D-vine pair-copula model

(Y_1, Y_2, Y_3)	PMF_{MP}	$PMF_{Gaussian}$
(0, 0, 0)	0.1854	0.1846
(0, 0, 1)	0.3516	0.3524
(0, 1, 0)	0.0174	0.0182
(0, 1, 1)	0.1154	0.1147
(1, 0, 0)	0.0689	0.0697
(1, 0, 1)	0.1339	0.1331
(1, 1, 0)	0.0181	0.0173
(1, 1, 1)	0.1088	0.1096

The correlation for binary variables are as below: $\text{Corr}(Y_1, Y_2) = 0.2003$, $\text{Corr}(Y_2, Y_3) = 0.2004$, $\text{Corr}(Y_1, Y_3) = 0.0408$ for the MP model; $\text{Corr}(Y_1, Y_2) = 0.2001$, $\text{Corr}(Y_2, Y_3) = 0.2004$, $\text{Corr}(Y_1, Y_3) = 0.0407$ for the D-vine pair-copula model using Gaussian copula. Since $\rho^2 = 0.04$, both model generates AR(1) structured binary variables.

For the equicorrelated structure, the correlation matrix is similar to the one of AR(1) structure except $\gamma_{13} = 0.357$. The PMF values of the MP model and pair-copula models are all very close but not identical. As an example, for $y = (0, 0, 1)$, the probability is 0.2123 for MP model, and it is 0.2108 for D-vine pair-copula using Gaussian copula.

The correlation for binary variables are check as well, and they're shown as below: $\text{Corr}(Y_1, Y_2) = 0.2004$, $\text{Corr}(Y_2, Y_3) = 0.2006$, $\text{Corr}(Y_1, Y_3) = 0.2008$ for the MP model; $\text{Corr}(Y_1, Y_2) = 0.2001$, $\text{Corr}(Y_2, Y_3) = 0.2003$, $\text{Corr}(Y_1, Y_3) = 0.1977$ for the D-vine pair-copula model using Gaussian copula. Both model generates equicorrelated structured binary variables.

Table 12: PMF of equicorrelated trivariate binary variables generated by the MP model, D-vine pair-copula model

(Y_1, Y_2, Y_3)	PMF_{MP}	$PMF_{Gaussian}$
(0, 0, 0)	0.2123	0.2108
(0, 0, 1)	0.3246	0.3263
(0, 1, 0)	0.0246	0.0256
(0, 1, 1)	0.1082	0.1073
(1, 0, 0)	0.0420	0.0436
(1, 0, 1)	0.1608	0.1593
(1, 1, 0)	0.0109	0.0099
(1, 1, 1)	0.1161	0.1171

II.4.4 EXAMPLES WHERE MP MODEL FAILS WHILE PAIR-COPULA MODEL WORKS

As mentioned in Section II.2.6, in some cases the MP model fails to generate a PMF even though it exists for certain marginal means and correlation parameter value in the feasible range. We show in this section the pair-copula model proposed in this dissertation is successful in generating a PMF in cases where the MP model fails. We consider the example given in Yang and Chaganty (2014). The marginal means for their example are given by the vector $p = (0.26,\ 0.36,\ 0.25,\ 0.24)$. For the AR(1) structure the feasible range of the correlation parameter ρ is $(-0.3244, 0.7698)$. Using the value $\rho = 0.72$, Yang and Chaganty (2014) have shown the latent correlation matrix is

$$R = \begin{bmatrix} 1 & 0.9378 & 0.7511 & 0.5869 \\ 0.9378 & 1 & 0.9460 & 0.7657 \\ 0.7511 & 0.9460 & 1 & 0.9157 \\ 0.5869 & 0.7657 & 0.9157 & 1 \end{bmatrix},$$

which turns out to be not positive definite and thus the MP method does not give a PMF for the binary variables. However, we now show that the pair-copula method is useful to generate a PMF for the binary variables. Note that the elements in R were obtained using Equation (2.2.1), for example $\gamma_{23} = 0.9460$ is obtained solving this equation

$$\begin{aligned}
\mathrm{Corr}(Y_2, Y_3) = 0.72 &= \frac{\Phi_2(\Phi^{-1}(p_2), \Phi^{-1}(p_3); \gamma) - p_2 p_3}{\sqrt{p_2 q_2 p_3 q_3}} \\
&= \frac{\Phi_2(\Phi^{-1}(0.36), \Phi^{-1}(0.25); \gamma) - 0.36 * 0.25}{\sqrt{0.36 * 0.64 * 0.25 * 0.75}}.
\end{aligned}$$

Selecting the bivariate Gaussian copula with parameter $\gamma_{23} = 0.9460$, and using $p_2 = 0.36\,(q_2 = 0.64)$, $p_3 = 0.25\,(q_3 = 0.75)$, we construct the joint distribution of $(Y_2,\, Y_3)$ as in Table 3 and get

$$\begin{aligned}
P(Y_2 = 0, Y_3 = 0) &= C(q_2, q_3; \gamma_{23}) = C(0.64, 0.75; 0.9460) = 0.6296 \\
P(Y_2 = 0, Y_3 = 1) &= q_2 - C(q_2, q_3; \gamma_{23}) = 0.64 - 0.6296 = 0.0104 \\
P(Y_2 = 1, Y_3 = 0) &= q_3 - C(q_2, q_3; \gamma_{23}) = 0.75 - 0.6296 = 0.1204 \\
P(Y_2 = 1, Y_3 = 1) &= 1 - q_2 - q_3 + C(q_2, q_3; \gamma_{23}) \\
&= 1 - 0.64 - 0.75 + 0.6296 = 0.2396
\end{aligned}$$

For the D-vine as shown in Figure 2, we also need bivariate PMFs of the pairs (Y_1, Y_2) and (Y_3, Y_4) for the first tree. These can be obtained similarly and the results are given in Table 13.

From the bivariate distributions we get the conditional probabilities

$$\begin{aligned}
p_{1|0} &= P(Y_1 = 1|Y_2 = 0) = \frac{P(Y_1 = 1, Y_2 = 0)}{P(Y_2 = 0)} = \frac{0.0148}{0.64} = 0.02313 \\
p_{1|1} &= P(Y_1 = 1|Y_2 = 1) = \frac{P(Y_1 = 1, Y_2 = 1)}{P(Y_2 = 1)} = \frac{0.2452}{0.36} = 0.68111 \\
p_{3|0} &= P(Y_3 = 1|Y_2 = 0) = \frac{P(Y_3 = 1, Y_2 = 0)}{P(Y_2 = 0)} = \frac{0.0104}{0.64} = 0.01625 \\
p_{3|1} &= P(Y_3 = 1|Y_2 = 1) = \frac{P(Y_3 = 1, Y_2 = 1)}{P(Y_2 = 1)} = \frac{0.2396}{0.36} = 0.66556
\end{aligned}$$

Note that $Y_1|Y_2 = 0$, $Y_1|Y_2 = 1$, $Y_3|Y_2 = 0$, and $Y_3|Y_2 = 1$ are binary variables with means $p_{1|0}$, $p_{1|1}$, $p_{3|0}$, and $p_{3|1}$, respectively. For standard normal random vector $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4)$ with AR(1) correlation structure ϵ_1, ϵ_3 given ϵ_2 are uncorrelated and therefore are independent. Therefore, we use independence copula to construct the joint PMF of $(Y_1,\, Y_3)$ given Y_2. Thus we get

$$\begin{aligned}
P(Y_1 = 0, Y_3 = 0|Y_2 = 0) &= q_{1|0} * q_{3|0} = 0.96100 \\
P(Y_1 = 0, Y_3 = 1|Y_2 = 0) &= q_{1|0} * p_{3|0} = 0.01587 \\
P(Y_1 = 1, Y_3 = 0|Y_2 = 0) &= p_{1|0} * q_{3|0} = 0.02275 \\
P(Y_1 = 1, Y_3 = 1|Y_2 = 0) &= p_{1|0} * p_{3|0} = 0.00038.
\end{aligned}$$

Table 13: Bivariate PMF from D-vine tree 1 using Gaussian copula

	Probability
$(Y_1,\ Y_2)$	
(0, 0)	$C(0.74, 0.64; 0.9378) = 0.6252$
(0, 1)	$0.74 - 0.6252 = 0.1148$
(1, 0)	$0.64 - 0.6252 = 0.0148$
(1, 1)	$1 - 0.74 - 0.64 + 0.6252 = 0.2452$
$(Y_2,\ Y_3)$	
(0, 0)	$C(0.64, 0.75; 0.9460) = 0.6296$
(0, 1)	$0.64 - 0.6296 = 0.0104$
(1, 0)	$0.75 - 0.6296 = 0.1204$
(1, 1)	$1 - 0.64 - 0.75 + 0.6296 = 0.2396$
$(Y_3,\ Y_4)$	
(0, 0)	$C(0.75, 0.76; 0.9157) = 0.7032$
(0, 1)	$0.75 - 0.7032 = 0.0468$
(1, 0)	$0.76 - 0.7032 = 0.0568$
(1, 1)	$1 - 0.75 - 0.76 + 0.7032 = 0.1932$

Similarly, we obtain the conditional bivariate pmf of $P(Y_1, Y_3|Y_2 = 1)$, $P(Y_2, Y_4|Y_3 = 0)$ and $P(Y_2, Y_4|Y_3 = 1)$ and these distributions are listed in Table 14. For tree 3 we need the conditional bivariate distributions $(P(Y_1,\ Y_4)|Y_2 =$

$y_2, Y_3 = y_3)$. is needed. These four conditional distributions depend on the following conditional probabilities

$$\begin{aligned}
p_{1|00} &= P(Y_1 = 1|Y_2 = 0, Y_3 = 0) = \frac{P(Y_1 = 1, Y_3 = 0|Y_2 = 0)}{P(Y_3 = 0|Y_2 = 0)} \\
&= \frac{P(Y_1 = 1, Y_3 = 0|Y_2 = 0)}{\frac{P(Y_2=0,Y_3=0)}{P(Y_2=0)}} = \frac{0.02275}{\frac{0.6296}{0.64}} = 0.023126 \\
p_{1|01} &= P(Y_1 = 1|Y_2 = 0, Y_3 = 1) = \frac{P(Y_1 = 1, Y_3 = 1|Y_2 = 0)}{P(Y_3 = 1|Y_2 = 0)} \\
&= \frac{P(Y_1 = 1, Y_3 = 1|Y_2 = 0)}{\frac{P(Y_2=0,Y_3=1)}{P(Y_2=0)}} = \frac{0.00038}{\frac{0.0104}{0.64}} = 0.02338 \\
p_{1|10} &= P(Y_1 = 1|Y_2 = 1, Y_3 = 0) = \frac{P(Y_1 = 1, Y_3 = 0|Y_2 = 1)}{P(Y_3 = 0|Y_2 = 1)} \\
&= \frac{P(Y_1 = 1, Y_3 = 0|Y_2 = 1)}{\frac{P(Y_2=1,Y_3=0)}{P(Y_2=1)}} = \frac{0.22779}{\frac{0.1204}{0.36}} = 0.68110 \\
p_{1|11} &= P(Y_1 = 1|Y_2 = 1, Y_3 = 1) = \frac{P(Y_1 = 1, Y_3 = 1|Y_2 = 1)}{P(Y_3 = 1|Y_2 = 1)} \\
&= \frac{P(Y_1 = 1, Y_3 = 1|Y_2 = 1)}{\frac{P(Y_2=1,Y_3=1)}{P(Y_2=1)}} = \frac{0.45332}{\frac{0.2396}{0.36}} = 0.68112
\end{aligned}$$

$$\begin{aligned}
p_{4|00} &= P(Y_4 = 1|Y_2 = 0, Y_3 = 0) = \frac{P(Y_2 = 0, Y_4 = 1|Y_3 = 0)}{P(Y_2 = 0|Y_3 = 0)} \\
&= \frac{P(Y_2 = 0, Y_4 = 1|Y_3 = 0)}{\frac{P(Y_2=0,Y_3=0)}{P(Y_3=0)}} = \frac{0.05244}{\frac{0.6296}{0.75}} = 0.06247 \\
p_{4|01} &= P(Y_4 = 1|Y_2 = 0, Y_3 = 1) = \frac{P(Y_2 = 0, Y_4 = 1|Y_3 = 1)}{P(Y_2 = 0|Y_3 = 1)} \\
&= \frac{P(Y_2 = 0, Y_4 = 1|Y_3 = 1)}{\frac{P(Y_2=0,Y_3=1)}{P(Y_3=1)}} = \frac{0.03199}{\frac{0.0104}{0.25}} = 0.76899 \\
p_{4|10} &= P(Y_4 = 1|Y_2 = 1, Y_3 = 0) = \frac{P(Y_2 = 1, Y_4 = 1|Y_3 = 0)}{P(Y_2 = 1|Y_3 = 0)} \\
&= \frac{P(Y_2 = 1, Y_4 = 1|Y_3 = 0)}{\frac{P(Y_2=1,Y_3=0)}{P(Y_3=0)}} = \frac{0.01002}{\frac{0.1204}{0.75}} = 0.06242 \\
p_{4|11} &= P(Y_4 = 1|Y_2 = 1, Y_3 = 1) = \frac{P(Y_2 = 1, Y_4 = 1|Y_3 = 1)}{P(Y_2 = 1|Y_3 = 1)} \\
&= \frac{P(Y_2 = 1, Y_4 = 1|Y_3 = 1)}{\frac{P(Y_2=1,Y_3=1)}{P(Y_3=1)}} = \frac{0.74062}{\frac{0.2396}{0.25}} = 0.77277
\end{aligned}$$

Assuming bivariate independent copulas we get

$$\begin{aligned}
P(Y_1 = 0, Y_4 = 0 | Y_2 = 0, Y_3 = 0) &= q_{1|00} * q_{4|00} \\
&= (1 - 0.023126) * (1 - 0.06247) = 0.91584 \\
P(Y_1 = 0, Y_4 = 1 | Y_2 = 0, Y_3 = 0) &= q_{1|00} * p_{4|00} \\
&= (1 - 0.023126) * 0.06247 = 0.06103 \\
P(Y_1 = 1, Y_4 = 0 | Y_2 = 0, Y_3 = 0) &= p_{1|00} * q_{4|00} \\
&= 0.023126 * (1 - 0.06247) = 0.02168 \\
P(Y_1 = 1, Y_4 = 1 | Y_2 = 0, Y_3 = 0) &= p_{1|00} * p_{4|00} \\
&= 0.023126 * 0.06247 = 0.00144
\end{aligned}$$

Similarly, the other conditional bivariate PMF for tree 3 can be calculated, and the distributions are listed in Table 15. Finally, the joint PMF can be obtained from Tables 13 and 15. For example,

$$\begin{aligned}
&P(Y_1 = 0, Y_2 = 0,\ Y_3 = 0,\ Y_4 = 0) \\
&\quad = P(Y_1 = 0,\ Y_4 = 0 | Y_2 = 0,\ Y_3 = 0) * P(Y_2 = 0,\ Y_3 = 0) \\
&\quad = 0.91584 * 0.6296 \\
&\quad = 0.5767
\end{aligned}$$

The four dimensional is given in Table 16. We can check that the correlation matrix of this distribution is

$$R_{binary} = \begin{bmatrix} 1.0 & 0.7200 & 0.5184 & 0.3733 \\ 0.7200 & 1.0 & 0.7200 & 0.5184 \\ 0.5184 & 0.7200 & 1.0 & 0.7200 \\ 0.3733 & 0.5184 & 0.7200 & 1.0 \end{bmatrix}$$

since $0.72^2 = 0.5184$, $0.72^3 = 0.3732$, this has an approximate AR(1) structure.

Table 14: Conditional bivariate PMF from D-vine tree 2 using Gaussian copula

	Probability
$(Y_1, Y_3 \mid Y_2 = 0)$	
(0, 0)	0.96100
(0, 1)	0.01587
(1, 0)	0.02275
(1, 1)	0.00038
$(Y_1, Y_3 \mid Y_2 = 1)$	
(0, 0)	0.10665
(0, 1)	0.21224
(1, 0)	0.22779
(1, 1)	0.45332
$(Y_2, Y_4 \mid Y_3 = 0)$	
(0, 0)	0.78710
(0, 1)	0.05244
(1, 0)	0.15044
(1, 1)	0.01002
$(Y_2, Y_4 \mid Y_3 = 1)$	
(0, 0)	0.00941
(0, 1)	0.03199
(1, 0)	0.21798
(1, 1)	0.74062

Table 15: Conditional bivariate PMF from D-vine tree 3 using Gaussian copula

	Probability
$(Y_1, Y_4 \vert Y_2 = 0, Y_3 = 0)$	
(0, 0)	0.91584
(0, 1)	0.02168
(1, 0)	0.06103
(1, 1)	0.00144
$(Y_1, Y_4 \vert Y_2 = 0, Y_3 = 1)$	
(0, 0)	0.22561
(0, 1)	0.75101
(1, 0)	0.00540
(1, 1)	0.01798
$(Y_1, Y_4 \vert Y_2 = 1, Y_3 = 0)$	
(0, 0)	0.29899
(0, 1)	0.01991
(1, 0)	0.63859
(1, 1)	0.04251
$(Y_1, Y_4 \vert Y_2 = 1, Y_3 = 1)$	
(0, 0)	0.07246
(0, 1)	0.24642
(1, 0)	0.15477
(1, 1)	0.52635

Table 16: Four dimensional distribution with specified marginals

(Y_1, Y_2, Y_3, Y_4)	**PMF**
(0, 0, 0, 0)	0.5767
(0, 0, 0, 1)	0.0384
(0, 0, 1, 0)	0.0023
(0, 0, 1, 1)	0.0078
(0, 1, 0, 0)	0.0360
(0, 1, 0, 1)	0.0024
(0, 1, 1, 0)	0.0174
(0, 1, 1, 1)	0.0590
(1, 0, 0, 0)	0.0137
(1, 0, 0, 1)	0.0009
(1, 0, 1, 0)	0.0001
(1, 0, 1, 1)	0.0002
(1, 1, 0, 0)	0.0768
(1, 1, 0, 1)	0.0051
(1, 1, 1, 0)	0.0371
(1, 1, 1, 1)	0.1261

II.5 PARAMETER ESTIMATION

In this section, we discuss the maximum likelihood estimation (mle) for the D-vine pair-copula model parameters. Assume there are n independent subjects, and there are m repeated binary observations on each subject. Thus we have a binary vector $\mathbf{y}_i = (y_{i1}, y_{i2}, \cdots, y_{im})$ of dimension m. Let p_j be the marginal probability of y_{ij} assumed to be the same for all i. There are 2^m possible combinations for $\mathbf{y}_i$. For instance, when $m = 4$, we have 16 combinations, that is, $\mathbf{y}_i = (0, 0, 0, 0)$, or $(0, 0, 0, 1)$, or $(0, 0, 1, 0)$, $\cdots$, or $(1, 1, 1, 1)$. The n observations can be grouped into 2^m counts. Assume the number of $(0, \cdots, 0)$ vectors is n_1, the number of $(0, \cdots, 1)$ is n_2, so on and so forth, the number of $(1, \cdots, 1)$ is n_{2^m}. Using these notations, the loglikelihood, $\ell(\boldsymbol{\theta})$, for D-vine pair-copula model for a sample of n independent observations is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}) \;&=\; n_1 \log P(Y_{i1} = 0, Y_{i2} = 0, \cdots, Y_{im} = 0) + n_2 \log P(Y_{i1} = 0, Y_{i2} = 0, \cdots, \\ &\quad Y_{im} = 1) + \cdots + n_{2^t} \log P(Y_{i1} = 1, Y_{i2} = 1, \cdots, Y_{im} = 1) \end{aligned} \tag{2.5.24}$$

where the parameter $\boldsymbol{\theta}$ consists of marginal probabilities and copula parameters that are functions of correlations between the binary variables. Take the two dimensional example shown in Table 3 for instance, the loglikelihood is shown in the Equation (2.5.25).

$$\begin{aligned} \ell(\gamma_{12}, p_1, p_2) \;&=\; n_1 \log(P(Y_{i1} = 0, Y_{i2} = 0)) + n_2 \log(P(Y_{i1} = 0, Y_{i2} = 1)) \\ &\quad + n_3 \log(P(Y_{i1} = 1, Y_{i2} = 0)) + n_4 \log(P(Y_{i1} = 1, Y_{i2} = 1)) \\ &=\; n_1 \log(C(q_1, q_2)) + n_2 \log(q_1 - C(q_1, q_2)) \\ &\quad + n_3 \log(q_2 - C(q_1, q_2)) + n_4 \log(1 - q_1 - q_2 + C(q_1, q_2)) \end{aligned} \tag{2.5.25}$$

The mle is obtained using the method "L-BFGS-B" by Byrd et al. (1995) which allows box constraints, the estimation of gradient function is approached using a finite-difference approximation, while the Hessian matrix of the parameters at optimized values is approximated using method "Richardson" of function "Hessian" in the R package "numDeriv" by Gilbert and Varadhan (2012).

II.6 DATA ANALYSIS

Three examples of longitudinal binary data are presented in this section. We illustrate the differences and similarities of the MP model and D-Vine pair-copula model with these examples. For example 1 that consists of traffic data, we fit the D-vine pair copula model with Gausian copula and AR(1) stucture. For example 2 that consists of church attendance data we stick with the Gaussian copula but use both AR(1) and equicorrelated structures. For example 3, that consists of repeated binary response to three drugs, we use the equicorrelated structure.

II.6.1 TRAFFIC VIOLATION DATA

This data arises from a randomized experiment conducted by Stock et al. (1983) (the "DeKalb study") to evaluate the impact of the driver education on the number of collisions and violations among teenage drivers. The study design consists of eligible students categorized into three groups based on the curriculum: safe performance curriculum, pre-driver licensing curriculum and a control group. The data was obtained using records from the state Department of Motor Vehicles for four consecutive years. We focus our attention on the control group and study the changes over the four years. Figure 5 displays pie chart of data from the control group consisting of 2409 males. Labels in the picture represents whether they had traffic violations, for example, 0110 indicates subjects in this category didn't have traffic violations in the first year and the forth year, but did have one or more violations in the second and third years.

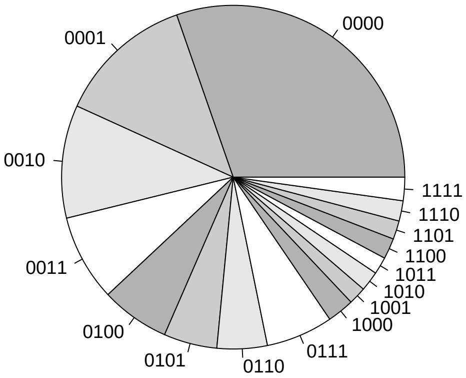


Figure 5: Traffic violation follow-up data of teenage drivers

Table 17 presents the mle estimates, standard errors and AIC for the models. The marginal parameter p_i represents the probability of having a traffic violation in the ith year for a teenager in this control group. The estimates of all marginal probabilities and AR(1) correlation parameter ρ are very close in both models. Since the estimates of p_i's are increasing we can concluse that the traffic violations rate is increasing with time in this group.

Table 17: Parameter estimation for the traffic data

	MP			**Gaussian D-Vine**		
Parameter	EST	SE	P-value	EST	SE	P-value
p_1	0.1673	0.0018	<0.0001	0.1539	0.0073	<0.0001
p_2	0.3288	0.0029	<0.0001	0.3030	0.0094	<0.0001
p_3	0.3777	0.0054	<0.0001	0.3743	0.0098	<0.0001
p_4	0.3824	0.0032	<0.0001	0.3959	0.0100	<0.0001
ρ	0.1429	0.0010	<0.0001	0.1672	0.0121	<0.0001
AIC	11277.85			11265.05		

II.6.2 CHURCH ATTENDANCE DATA

The Iowa 65+ rural health study by Mobily et al. (1994) studied the potential factors that effect low back pain of elderly persons. One variable of interest was the church attendance, surveyed three times over a six year period, as in year 0, year 3 and year 6. The response 0 means the subject isn't a regular church attender, while 1 means the subject attend church regularly. The complete data involves 1973 individuals. We compute the maximum likelihood estimates of the attendance rates and of the AR(1) parameter, assuming AR(1) structure.

Table 18 shows the mle estimates and standard errors for all the models. All marginal probabilities and AR(1) correlation parameter ρ are very close in each models. The p-values are not listed, because they all are very small and are significant.

Table 18: Estimates for the church attendance data with AR(1) structure

	MP			**Gaussian D-Vine**		
Parameter	EST	SE	P-value	EST	SE	P-value
p_1	0.7818	0.0083	<0.0001	0.7767	0.0100	<0.0001
p_2	0.7638	0.0091	<0.0001	0.7640	0.0100	<0.0001
p_3	0.7176	0.0121	<0.0001	0.7238	0.0107	<0.0001
ρ	0.6648	0.0150	<0.0001	0.6870	0.0146	<0.0001
AIC	4854.97			4856.16		

The church attendance rates of the three years are very close and large, thus y's are of high dependence, it is mainly due to the high frequency of $(1, 1, 1)$, which is 0.656 from the data. High weight brought by $(1, 1, 1)$ makes the marginal proportions large and close. It might be a good choice to assume $\mathbf{y}$'s equally correlated, and the mle is presented in Table 19.

Table 19: Estimates for the church attendance data using equicorrelated structure

	MP			**Gaussian D-Vine**		
Parameter	EST	SE	P-value	EST	SE	P-value
p_1	0.7902	0.0079	<0.0001	0.7768	0.0093	<0.0001
p_2	0.7525	0.0103	<0.0001	0.7657	0.0096	<0.0001
p_3	0.7394	0.0092	<0.0001	0.7262	0.0098	<0.0001
ρ	0.6477	0.0126	<0.0001	0.6586	0.0153	<0.0001
AIC	4794.55			4776.91		

Among both structures of all models, D-Vine pair-copula using Gaussian copula with equicorrelated correlation structure are the best, because they have the minimum AIC value.

II.6.3 DRUG RESPONSE DATA

This data was first reported by Grizzle et al. (1969). Here 46 subjects were treated with three drugs 1, 2 and 3, and recorded their response as 0 for unfavorable or 1 for favorable. The frequencies are given in Table 20. We assume the three binary responses are equicorrelated and fit our models. The maximum likelihood estimates of the marginal proportions and the correlation parameter of our models are presented in Table 21.

Table 20: Drug response data

(Y_1, Y_2, Y_3)	**Frequency**
(0, 0, 0)	6
(0, 0, 1)	16
(0, 1, 0)	2
(0, 1, 1)	4
(1, 0, 0)	2
(1, 0, 1)	4
(1, 1, 0)	6
(1, 1, 1)	6

Table 21: Parameter estimation for the drug response data

	MP			**Gaussian D-Vine**			**Indep. D-Vine**		
Parameter	EST	SE	P-value	EST	SE	P-value	EST	SE	P-value
p_1	0.3911	0.0871	<0.0001	0.3906	0.0725	<0.0001	0.3913	0.0719	<0.0001
p_2	0.4081	0.0994	<0.0001	0.3905	0.0725	<0.0001	0.3913	0.0719	<0.0001
p_3	0.6635	0.1042	<0.0001	0.6495	0.0700	<0.0001	0.6522	0.0702	<0.0001
ρ	0.0683	0.0216	0.0008	0.0395	0.0848	0.3207	NA	NA	NA
AIC	190.59			190.37			188.59		

Although the estimate of ρ is significant for MP model, it is close to zero for D-Vine pair-copula model and not significant. Take the MP model for example, $p = (0.3911, 0.4081, 0.6635)$ and correlation parameter $\rho = 0.0683$ which is within the feasible range $(-0.323, 0.591)$ for equicorrelated correlation structure. These values generate a correlation matrix of latent variable as below.

$$R = \begin{bmatrix} 1 & 0.109 & 0.113 \\ 0.109 & 1 & 0.112 \\ 0.113 & 0.112 & 1 \end{bmatrix}$$

All latent variable correlations are around 0.1, thus independent copula is a competing model. The estimate and standard error of p's using D-vine independent copulas are listed at the last two columns in the Table 20. The D-Vine independent copula model has the minimum AIC and seems to be a good choice for this data.

II.7 CONCLUSIONS

The multivariate probit model is one of the most popular model to analyze the dependence relationships of longitudinal binary data. Some studies focused on the vine pair-copula of the discrete data. Our aim in this paper is to develop the D-vine pair-copula models for the specific situation as longitudinal binary data, assumed as first order autoregressive or equicorrelated structured, to estimates the marginal proportions and correlation parameter. We have also shown that MP model is different from vine Gaussian pair-copula starting three dimensions. The main advantage of our model is that it can produce the PMF around correlation feasible boundaries where MP model fails.

Relationship between the binary variable correlation and the copula parameters are presented. Our model is flexible due to the multiple options of copulas, including Gaussian, Clayton, Frank, Gumbel and even independent copulas. However, We fixed the conditional dependence of multivariate binary variables with equicorrelated structure, or assumed conditional independence with AR(1) structure, therefore, no matter which copula is used, the same PMF will be produced. Independent copula is applied when variables have very low correlated coefficients. Best model could be chosen using AIC. Work in the next chapter would be considering the covariates as well for binary longitudinal data, which is more common in the real life data.

CHAPTER III

D-VINE PAIR GAUSSIAN COPULA REGRESSION MODEL

III.1 INTRODUCTION

In Chapter II, we discussed the D-vine pair-copula method of constructing a joint distribution for dependent binary observations that arise in longitudinal studies. We have also discussed estimation of the parameter that include the marginal means and copula parameters which are related to the correlation betweeen the binary observations. In this chapter, we will extend the D-vine pair-copula model to the regression setting assuming the data consists of covariates associated with the binary responses. We give mathematical details for the maximum likelihood estimation of the regression and correlation parameters. We develop formulas for the score functions which will be used to develop a R-code for obtaining the estimates and standard errors for real life longitudinal data. We also compare our results with the MP model through efficiency calculations via simulations.

III.2 PARAMETER ESTIMATION FOR REGRESSION MODEL

Assume that we have n independent subjects. Let $y_i = (y_{i1}, y_{i2}, ..., y_{im})$ be a sequence of binary observations on the ith subject. Associated with y_i we have a matrix $X_i = (x_{i1}, x_{i2}, ..., x_{im})$, where the column vector $x_{it} = (x_{it1}, x_{it2}, ..., x_{itk})'$ consist of k covariates for subject i at time t. For the regression model we assume the mean $E(y_{it}) = p_{it}$ is linked to the covariates via a probit link $p_{it} = \Phi(x'_{it}\beta)$ or a logit link $\text{logit}(p_{it}) = x'_{it}\beta$, where β is the regression coefficient. For simplicity we assume the correlation parameter ρ is a constant and does not depend on the covariates. The goal is to find the maximum likelihood estimate of the parameter vector $\theta = (\beta,\, \rho)$ and associated standard errors. The loglikehood for the D-vine pair-copula model for a random sample of n subjects is given as $\ell(\theta|y, x) = \sum_{i=1}^{n} \log\, f(y_i; X_i, \theta)$, where the multivariate binary probability mass function $f(y_i; X_i, \theta)$ is determined by the D-vine pair-copula described in the Equation (2.2.7) in the previous chapter. In this chapter we only deal with AR(1) structure since we are dealing with longitudinal data. For notational convenience we drop the arguments X_i and θ and simply write $f(y_i)$ for $f(y_i; X_i, \theta)$. The log likelihood is

$$\begin{aligned}
\ell(\theta|y,x) &= \sum_{i=1}^{n} \log\, f(y_i) \qquad (3.2.26)\\
&= \sum_{i=1}^{n}\{\log\, f(y_{i1}, y_{i,m}|y_{i2},..,y_{i,m-1}) + \log\, f(y_{i2}, y_{i,m-1}|y_3,..,y_{i,m-2})...\\
&\qquad +\log\, f(y_{i,\frac{m+1}{2}})\} \qquad \text{if } m \text{ is an odd number;}\\
&= \sum_{i=1}^{n}\{\log\, f(y_{i1}, y_{i,m}|y_{i2},..,y_{i,m-1}) + \log\, f(y_{i2}, y_{i,m-1}|y_3,..,y_{i,m-2})...\\
&\qquad +\log\, f(y_{i,\frac{m}{2}}, y_{i,\frac{m}{2}+1})\} \qquad \text{if } m \text{ is an even number.}
\end{aligned}$$

To find the maximum likelihood estimates we need to find the derivatives of (3.2.26) with respect to $\theta = (\beta,\, \rho)$. This involves find the derivative of univariate marginal, bivariate distribution, and bivariate conditional distributions. Derivative of the marginal probability, $f(y_{ij})$, with respect to β is presented in Section III.3.2, partial derivatives of bivariate probabilities from D-vine tree 1, $f(y_{ij}, y_{i,j+1})$ with respect to β, and ρ are presented in Section III.3.3. Next, partial derivatives of conditional bivariate probability from D-vine tree 2, $f(y_{ij}, y_{i,j+2}|y_{i,j+1})$ with respect to β and ρ are presented in Section III.3.4. Finally, partial derivatives of conditional bivariate probability from D-vine tree 3 and higher order $f(y_{ij}, y_{i,j+k}|y_{i,j+1}, ..., y_{i,j+k-1})$ with respect to β and ρ are presented in Section III.3.5.

III.3 SCORE FUNCTION DETAILS

For notational convenience in the following subsections we drop the subscript i, for example we write y_j for y_{ij}, pmf $f(y_j)$ for $f(y_{ij})$, p_j for p_{ij}, and $\gamma_{j,j+1}$ for $\gamma_{ij,ij+1}$ etc. Recall that the subscript i stands for the ith subject, so that the expressions that we derive are valid for the variables on any subject. Note that f stands for a generic pmf and the arguments determine the distribution of the variables involved. For example, $f(y_j)$, $f(y_j,\, y_{j+1})$, and $f(y_j|y_{j+1})$ denote the mariginal pmf of y_j, the joint pmf of $(y_j,\, y_{j+1})$ and the conditional pmf of y_j given y_{j+1} etc.

III.3.1 LINK FUNCTION

We usually relate the mean p_j of the binary variable y_j to the covariates vector x_j either by the probit $p_j = \Phi(x_j'\beta)$ or the logit $p_j = 1/(1+\exp(-x_j'\beta))$ link functions. The partial derivative of p_j with respect to β_l is given by

$$\frac{\partial p_j}{\partial \beta_l} = \phi(x_j'\beta)\ x_{jl},$$

if the link function is probit and it is

$$\frac{\partial p_j}{\partial \beta_l} = \frac{x_{jl}\ \exp(-x_j'\beta)}{(1+\exp(-x_j'\beta))^2}.$$

for the logit function. The partial derivative of p_j with respect to ρ is zero.

III.3.2 DERIVATIVE OF THE MARGINAL DISTRIBUTION

Note that the mariginal distribution of y_j is

$$f(y_j) = (p_j)^{y_j}(1-p_j)^{(1-y_j)} = p_j I(y_j = 1) + (1-p_j)I(y_j = 0).$$

The partial derivative with respect to the regression coefficient β_l can be written as,

$$\begin{aligned}\frac{\partial f(y_j)}{\partial \beta_l} &= \tfrac{\partial p_j}{\partial \beta_l} I(y_j = 1) - \tfrac{\partial p_j}{\partial \beta_l} I(y_j = 0)\\ &= \tfrac{\partial p_j}{\partial \beta_l}(I(y_j = 1) - I(y_j = 0)).\end{aligned}$$

Clearly, the partial derivative with respect to ρ is zero.

III.3.3 DERIVATIVES OF THE DISTRIBUTION IN D-VINE TREE 1

III.3.3.1 Partial derivative with respect to ρ

For the D-vine pair-copula, the first tree consists of distributions of the pairs {12}, {23}, {34},...,$\{m-1, m\}$. Using the Gaussian copula $C(u_j, u_{j+1}) = \Phi_2(\Phi^{-1}(u_j), \Phi^{-1}(u_{j+1}); \gamma_{j,j+1})$, where $\gamma_{j,j+1}$ is the correlation coefficient of latent variables, the joint distribution of $(y_j\ y_{j+1})$, is presented in Table 22.

Table 22: PMF of bivariate distribution in tree 1

$(y_j,\ y_{j+1})$	**Probability**
$(0,\ 0)$	$\Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})$
$(0,\ 1)$	$q_j - \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})$
$(1,\ 0)$	$q_{j+1} - \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})$
$(1,\ 1)$	$1 - q_j - q_{j+1} + \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})$

NOTE: The parameters of this distribution are means p_j, p_{j+1} and correlation $\gamma_{j,j+1}$.

In order to take derivative of PMF with respect to the binary correlation coefficient ρ, we start working with the bivariate normal CDF $\Phi_2(a, b; \gamma)$ with respect to its arguments a, b and latent correlation coefficient γ. Plackett (1954) showed

$$\begin{aligned} \frac{\partial}{\partial\gamma}\Phi_2(a, b; \gamma) &= \frac{\partial^2}{\partial a \partial b}\Phi_2(a, b; \gamma) \\ &= \phi_2(a, b; \gamma). \end{aligned} \tag{3.3.27}$$

Also, the relationship between ρ and $\gamma_{j,j+1}$ is:

$$\mathrm{Corr}(y_j,\ y_{j+1}) = \rho \;=\; \frac{\Phi_2(\Phi^{-1}(p_j), \Phi^{-1}(p_{j+1}); \gamma_{j,j+1}) - p_j p_{j+1}}{\sigma_j\ \sigma_{j+1}} \tag{3.3.28}$$

where $\sigma_j = \sqrt{p_j(1-p_j)}$. Taking derivative with respect to ρ on both sides we get

$$\begin{aligned} \frac{\partial\rho}{\partial\rho} = 1 &= \frac{1}{\sigma_j\sigma_{j+1}}\frac{\partial}{\partial\rho}\{\Phi_2(\Phi^{-1}(p_j), \Phi^{-1}(p_{j+1}); \gamma_{j,j+1}) - p_j p_{j+1}\} \\ \sigma_j\sigma_{j+1} &= \frac{\partial\Phi_2(\Phi^{-1}(p_j), \Phi^{-1}(p_{j+1}); \gamma_{j,j+1})}{\partial\gamma_{j,j+1}}\frac{\partial\gamma_{j,j+1}}{\partial\rho} \\ \sigma_j\sigma_{j+1} &= \phi_2(\Phi^{-1}(p_j), \Phi^{-1}(p_{j+1}); \gamma_{j,j+1})\frac{\partial\gamma_{j,j+1}}{\partial\rho}. \end{aligned}$$

The last equality follows from Plackett's identity (3.3.27). Rearranging the terms we get

$$\begin{aligned}
\frac{\partial \gamma_{j,j+1}}{\partial \rho} &= \frac{\sigma_j \sigma_{j+1}}{\phi_2(\Phi^{-1}(p_j), \Phi^{-1}(p_{j+1}); \gamma_{j,j+1})} \\
&= \frac{\sigma_j \sigma_{j+1}}{\phi_2(\Phi^{-1}(1-q_j), \Phi^{-1}(1-q_{j+1}); \gamma_{j,j+1})} \\
&= \frac{\sigma_j \sigma_{j+1}}{\phi_2(-\Phi^{-1}(q_j), -\Phi^{-1}(q_{j+1}); \gamma_{j,j+1})} \\
&= \frac{\sigma_j \sigma_{j+1}}{\phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})}.
\end{aligned}$$

Now, the derivative of PMF with respect to the binary correlation coefficient ρ,

$$\begin{aligned}
\frac{\partial}{\partial \rho} f(y_j = 0, y_{j+1} = 0) &= \frac{\partial}{\partial \rho} \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1}) \\
&= \frac{\partial \gamma_{j,j+1}}{\partial \rho} \frac{\partial}{\partial \gamma_{j,j+1}} \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1}) \\
&= \frac{\partial \gamma_{j,j+1}}{\partial \rho} \phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1}) \\
&= \frac{\sigma_j \sigma_{j+1}}{\phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1})} \phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1}) \\
&= \sigma_j \sigma_{j+1}.
\end{aligned}$$

Next,

$$\begin{aligned}
\frac{\partial}{\partial \rho} f(y_j = 0, y_{j+1} = 1) &= \frac{\partial}{\partial \rho} q_j - \frac{\partial}{\partial \rho} \Phi_2(\Phi^{-1}(q_j), \Phi^{-1}(q_{j+1}); \gamma_{j,j+1}) \\
&= -\sigma_j \sigma_{j+1}.
\end{aligned}$$

Similarly, the derivative of the PMF at other values can be obtained, and thus we have

$$\frac{\partial f(y_j, y_{j+1})}{\partial \rho} = \begin{cases} \sigma_j \sigma_{j+1} \quad if \quad y_j = 0 \quad and \quad y_{j+1} = 0, \\ -\sigma_j \sigma_{j+1} \quad if \quad y_j = 0 \quad and \quad y_{j+1} = 1, \\ -\sigma_j \sigma_{j+1} \quad if \quad y_j = 1 \quad and \quad y_{j+1} = 0, \\ \sigma_j \sigma_{j+1} \quad if \quad y_j = 1 \quad and \quad y_{j+1} = 1. \end{cases} \tag{3.3.29}$$

III.3.3.2 Partial derivative with respect to β

In order to take derivative of the bivariate PMF from tree 1 with respect to regression coefficients β's, the partial derivative of the Gaussian copula with respect to the first argument is needed. It follows from the properties of the normal distribution (see Section 3.10.1 on page 129 in Joe (2014) or Table 1 in Smith et al. (2010)), we have

$$\frac{\partial C(u_j, u_{j+1})}{\partial u_j} = \Phi\left(\frac{\Phi^{-1}(u_{j+1}) - \gamma_{j,j+1}\Phi^{-1}(u_j)}{\sqrt{1-\gamma_{j,j+1}^2}}\right).$$

We are now ready to take the derivative with respect to the regression coefficient β_l. Note that $u_j = F(y_j) = I(y_j = 1) + (1 - p_j)I(y_j = 0)$, hence $\partial u_j/\partial\beta_l = -\partial p_j/\partial\beta_l I(y_j = 0)$. Then,

$$\begin{aligned}
\frac{\partial C(u_j, u_{j+1})}{\partial \beta_l} &= \frac{\partial C}{\partial \gamma_{j,j+1}}\frac{\partial \gamma_{j,j+1}}{\partial \beta_l} + \frac{\partial C}{\partial u_j}\frac{\partial u_j}{\partial \beta_l} + \frac{\partial C}{\partial u_{j+1}}\frac{\partial u_{j+1}}{\partial \beta_l} \\
&= 0 + \frac{\partial C}{\partial u_j}\frac{\partial u_j}{\partial \beta_l} + \frac{\partial C}{\partial u_{j+1}}\frac{\partial u_{j+1}}{\partial \beta_l} \\
&= \Phi\left(\frac{\Phi^{-1}(u_{j+1}) - \gamma_{j,j+1}\Phi^{-1}(u_j)}{\sqrt{1-\gamma_{j,j+1}^2}}\right)\left(-\frac{\partial p_j}{\partial \beta_l} I(y_j = 0)\right) \\
&\quad +\Phi\left(\frac{\Phi^{-1}(u_j) - \gamma_{j,j+1}\Phi^{-1}(u_{j+1})}{\sqrt{1-\gamma_{j,j+1}^2}}\right)\left(-\frac{\partial p_{j+1}}{\partial \beta_l} I(y_{j+1} = 0)\right)
\end{aligned}$$

Now, the derivative of PMF with respect to the regression coefficient β_l,

$$\begin{aligned}
\frac{\partial}{\partial \beta_l} f(y_j = 0,\ y_{j+1} = 0) &= \frac{\partial C(q_j, q_{j+1})}{\partial \beta_l} \\
&= \Phi\left(\frac{\Phi^{-1}(q_{j+1}) - \gamma_{j,j+1}\Phi^{-1}(q_j)}{\sqrt{1-\gamma_{j,j+1}^2}}\right)\left(-\frac{\partial p_j}{\partial \beta_l}\right) \\
&\quad +\Phi\left(\frac{\Phi^{-1}(q_j) - \gamma_{j,j+1}\Phi^{-1}(q_{j+1})}{\sqrt{1-\gamma_{j,j+1}^2}}\right)\left(-\frac{\partial p_{j+1}}{\partial \beta_l}\right) \\
&= -\frac{\partial p_j}{\partial \beta_l} B_1 - \frac{\partial p_{j+1}}{\partial \beta_l} B_2,
\end{aligned}$$

where

$$B_1 = \Phi\left(\frac{\Phi^{-1}(q_{j+1}) - \gamma_{j,j+1}\Phi^{-1}(q_j)}{\sqrt{1-\gamma_{j,j+1}^2}}\right) \text{ and } B_2 = \Phi\left(\frac{\Phi^{-1}(q_j) - \gamma_{j,j+1}\Phi^{-1}(q_{j+1})}{\sqrt{1-\gamma_{j,j+1}^2}}\right).$$

Using Table 22, the partial derivative of $f(y_j = 0,\ y_{j+1} = 1)$ with respect to the regression coefficient is

$$\begin{aligned}\frac{\partial}{\partial\beta_l}f(y_j=0,\ y_{j+1}=1) &= \frac{\partial q_j}{\partial\beta_l}-\frac{\partial C(q_j,q_{j+1})}{\partial\beta_l}\\ &= -\frac{\partial p_j}{\partial\beta_l}+\frac{\partial p_j}{\partial\beta_l}B_1+\frac{\partial p_{j+1}}{\partial\beta_l}B_2.\end{aligned}$$

Similarly, the derivative of the PMF with other values can be obtained. Thus we have

$$\frac{\partial f(y_j,\ y_{j+1})}{\partial\beta_l}=\begin{cases}-\frac{\partial p_j}{\partial\beta_l}B_1-\frac{\partial p_{j+1}}{\partial\beta_l}B_2 & \text{if } y_j=0 \text{ and } y_{j+1}=0,\\ -\frac{\partial p_j}{\partial\beta_l}+\frac{\partial p_j}{\partial\beta_l}B_1+\frac{\partial p_{j+1}}{\partial\beta_l}B_2 & \text{if } y_j=0 \text{ and } y_{j+1}=1,\\ -\frac{\partial p_{j+1}}{\partial\beta_l}+\frac{\partial p_j}{\partial\beta_l}B_1+\frac{\partial p_{j+1}}{\partial\beta_l}B_2 & \text{if } y_j=1 \text{ and } y_{j+1}=0,\\ \frac{\partial p_j}{\partial\beta_l}+\frac{\partial p_{j+1}}{\partial\beta_l}-\frac{\partial p_j}{\partial\beta_l}B_1-\frac{\partial p_{j+1}}{\partial\beta_l}B_2 & \text{if } y_j=1 \text{ and } y_{j+1}=1,\end{cases}$$

where $\partial p_j/\partial\beta_l$ is given in Section III.3.1.

III.3.4 DERIVATIVES OF THE DISTRIBUTION IN D-VINE TREE 2

III.3.4.1 Partial derivative with respect to ρ

It is slightly different to take derivate of conditional copula from unconditional copula, due to both the changes in parameters and arguments. For tree 2, the conditional PMF of $\{13|2\}$, $\{24|3\}$,...,$\{m-2,m|m-1\}$ are obtained assuming conditional independence. Thus, independent copula $C(u_{j|j+1},\ u_{j+2|j+1};\ \theta_{j,j+2|j+1})=u_{j|j+1}*u_{j+2|j+1}$ is used for tree 2.

The first argument $u_{j|j+1}=F(y_j|y_{j+1})$ is the CDF of a new binary variable $y_j|y_{j+1}$ with new marginal mean $p_{j|j+1}$, and the second argument $u_{j+2|j+1}=F(y_{j+2}|y_{j+1})$ is the CDF of a new binary variable $y_{j+2}|y_{j+1}$ with new marginal mean $p_{j+2|j+1}$. The probabilities $p_{j|j+1}$ and $p_{j+2|j+1}$ can be obtained as

$$\begin{aligned}p_{j|j+1}=1-q_{j|j+1} &= 1-\frac{f(y_j=0,y_{j+1})}{f(y_{j+1})};\\ p_{j+2|j+1}=1-q_{j+2|j+1} &= 1-\frac{f(y_{j+1},y_{j+2}=0)}{f(y_{j+1})}.\end{aligned}$$

Therefore, the conditional bivariate PMF $f(y_j,\ y_{j+2}|y_{j+1})$, according to Table 10, is given by

Table 23: Conditional bivariate distribution

$(y_j,\, y_{j+2})\vert y_{j+1}$	**Probability**
(0,0)	$q_{j\vert j+1}\; q_{j+2\vert j+1}$
(0,1)	$q_{j\vert j+1}\; p_{j+2\vert j+1}$
(1,0)	$p_{j\vert j+1}\; q_{j+2\vert j+1}$
(1,1)	$p_{j\vert j+1}\; p_{j+2\vert j+1}$

The partial derivative of $q_{j|j+1}$ with respect to ρ is

$$\begin{aligned}\frac{\partial q_{j|j+1}}{\partial \rho} &= \frac{\partial}{\partial \rho}\frac{f(y_j=0,y_{j+1})}{f(y_{j+1})}\\ &= \frac{1}{f(y_{j+1})}\frac{\partial}{\partial \rho}f(y_j=0,y_{j+1})\\ &= \begin{cases}\frac{\sigma_j\sigma_{j+1}}{q_{j+1}} \quad if \quad y_{j+1}=0;\\ -\frac{\sigma_j\sigma_{j+1}}{p_{j+1}} \quad if \quad y_{j+1}=1.\end{cases}\end{aligned}$$

where we have used the result given in (3.3.29). In a compact form this can be written as

$$\frac{\partial q_{j|j+1}}{\partial \rho} = \frac{\sigma_j\sigma_{j+1}}{q_{j+1}}I(y_{j+1}=0) - \frac{\sigma_j\sigma_{j+1}}{p_{j+1}}I(y_{j+1}=1) = (-1)^{y_{j+1}}\frac{\sigma_j\sigma_{j+1}}{f(y_{j+1})}.$$

Similarly,

$$\frac{\partial q_{j+2|j+1}}{\partial \rho} = (-1)^{y_{j+1}}\frac{\sigma_{j+1}\sigma_{j+2}}{f(y_{j+1})}.$$

Therefore,

$$\begin{aligned}\frac{\partial}{\partial \rho}f(y_j=0,y_{j+2}=0|y_{j+1}) &= q_{j+2|j+1}\frac{\partial}{\partial \rho}q_{j|j+1} + q_{j|j+1}\frac{\partial}{\partial \rho}q_{j+2|j+1}\\ &= q_{j+2|j+1}(-1)^{y_{j+1}}\frac{\sigma_j\sigma_{j+1}}{f(y_{j+1})} + q_{j|j+1}(-1)^{y_{j+1}}\frac{\sigma_{j+1}\sigma_{j+2}}{f(y_{j+1})}\\ &= (-1)^{y_{j+1}}\frac{\sigma_{j+1}}{f(y_{j+1})}(q_{j+2|j+1}\sigma_j + q_{j|j+1}\sigma_{j+2})\end{aligned}$$

We also have,

$$\frac{\partial}{\partial \rho}f(y_j=0,y_{j+2}=1|y_{j+1}) = (-1)^{y_{j+1}}\frac{\sigma_{j+1}}{f(y_{j+1})}(p_{j+2|j+1}\sigma_j - q_{j|j+1}\sigma_{j+2})$$

Thus,

$$\frac{\partial}{\partial \rho}f(y_j,y_{j+2}|y_{j+1}) = (-1)^{y_{j+1}}\frac{\sigma_{j+1}}{f(y_{j+1})}((-1)^{y_j}f(y_{j+2}|y_{j+1}) + (-1)^{y_{j+2}}f(y_j|y_{j+1}))$$

III.3.4.2 Partial derivative with respect to β

Then the derivative of $q_{j|j+1}$ with respect to the regression coefficient is

$$\begin{aligned}\frac{\partial q_{j|j+1}}{\partial \beta_l} &= \frac{\partial}{\partial \beta_l}\frac{f(y_j=0,y_{j+1})}{f(y_{j+1})}\\ &= \frac{f(y_{j+1})\frac{\partial}{\partial \beta_l}f(y_j=0,y_{j+1})-f(y_j=0,y_{j+1})\frac{\partial f(y_{j+1})}{\partial \beta_l}}{f(y_{j+1})^2}\\ &= \frac{\frac{\partial}{\partial \beta_l}f(y_j=0,y_{j+1})-q_{j|j+1}\frac{\partial f(y_{j+1})}{\partial \beta_l}}{f(y_{j+1})},\end{aligned}$$

where $\frac{\partial}{\partial \beta_l}f(y_j=0,y_{j+1})$ are the results from tree 1 in Section III.3.3, and $\frac{\partial f(y_{j+1})}{\partial \beta_l}$ are from the derivative results of Bernoulli PMF in Section III.3.2.

The derivative of $q_{j+2|j+1}$ with respect to the β_l is similar. Then,

$$\frac{\partial}{\partial \beta_l}f(0,0|y_{j+1}) = \frac{\partial q_{j|j+1}}{\partial \beta_l}q_{j+2|j+1}+q_{j|j+1}\frac{\partial q_{j+2|j+1}}{\partial \beta_l}.$$

Similarly, $\frac{\partial}{\partial \beta_l}f(0,1|y_{j+1})$ and others can be obtained as well.

III.3.5 THE THIRD OR FURTHER TREES

Now, considering Gaussian copula conditional on two variables. Similar to tree 2, $p_{j|j+1,j+2}=1-q_{j|j+1,j+2}=1-\frac{f(y_j=0,y_{j+2}|y_{j+1})}{f(y_{j+2}|y_{j+1})}$. The derivative to the AR(1) parameter ρ or the regression coefficients,

$$\begin{aligned}\frac{\partial q_{j|j+1,j+2}}{\partial \theta} &= \frac{\partial}{\partial \theta}\frac{f(y_j=0,y_{j+2}|y_{j+1})}{f(y_{j+2}|y_{j+1})}\\ &= \frac{f(y_{j+2}|y_{j+1})\frac{\partial}{\partial \theta}f(y_j=0,y_{j+2}|y_{j+1})-f(y_j=0,y_{j+2}|y_{j+1})\frac{\partial f(y_{j+2}|y_{j+1})}{\partial \theta}}{f(y_{j+2}|y_{j+1})^2}\\ &= \frac{f(y_{j+2}|y_{j+1})\frac{\partial}{\partial \theta}f(y_j=0,y_{j+2}|y_{j+1})-q_{j|j+1,j+2}f\frac{\partial f(y_{j+2}|y_{j+1})}{\partial \theta}}{f(y_{j+2}|y_{j+1})^2}\\ &= \frac{\frac{\partial}{\partial \theta}f(y_j=0,y_{j+2}|y_{j+1})-q_{j|j+1,j+2}\frac{\partial f(y_{j+2}|y_{j+1})}{\partial \theta}}{f(y_{j+2}|y_{j+1})}\end{aligned}$$

where $\frac{\partial}{\partial \theta}f(y_j=0,y_{j+2}|y_{j+1})$ is from tree 2 in Section III.3.4, and $\frac{\partial f(y_{j+2}|y_{j+1})}{\partial \theta}$ is from tree 1 in Section III.3.3. Thus, the derivative of conditional bivariate PMF would be as shown in the table as below.

Table 24: Score function of conditional bivariate distribution from tree 3

$(y_j, y_{j+3})\|y_{j+1}, y_{j+2}$	**Probability**	$\frac{\partial}{\partial\theta}f(y_j, y_{j+3}\|y_{j+1}, y_{j+2})$
(0,0)	$q_{j\|j+1,j+2}q_{j+3\|j+1,j+2}$	$\frac{\partial q_{j\|j+1,j+2}}{\partial\theta}q_{j+3\|j+1,j+2}$ $+\frac{\partial q_{j+3\|j+1,j+2}}{\partial\theta}q_{j\|j+1,j+2}$
(0,1)	$q_{j\|j+1,j+2} - q_{j\|j+1,j+2}q_{j+3\|j+1,j+2}$	$\frac{\partial q_{j\|j+1,j+2}}{\partial\theta} - \frac{\partial f(0,0\|y_{j+1},y_{j+2})}{\partial\theta}$
(1,0)	$q_{j+3\|j+1,j+2} - q_{j\|j+1,j+2}q_{j+3\|j+1,j+2}$	$\frac{\partial q_{j+3\|j+1,j+2}}{\partial\theta} - \frac{\partial f(0,0\|y_{j+1},y_{j+2})}{\partial\theta}$
(1,1)	$(1 - q_{j\|j+1,j+2})(1 - q_{j+3\|j+1,j+2})$	$-\frac{\partial q_{j\|j+1,j+2}}{\partial\theta} - \frac{\partial q_{j+3\|j+1,j+2}}{\partial\theta}$ $+\frac{\partial f(0,0\|y_{j+1},y_{j+2})}{\partial\theta}$

Table 25: Conditional bivariate distribution from tree k

$(y_j, y_{j+k+1})\|y_{j+1}, ..., y_{j+k}$	**Probability**
(0,0)	$q_{j\|j+1,...,j+k}q_{j+k+1\|j+1,...,j+k}$
(0,1)	$q_{j\|j+1,...,j+k} - q_{j\|j+1,...,j+k}q_{j+k+1\|j+1,...,j+k}$
(1,0)	$q_{j+k+1\|j+1,...,j+k} - q_{j\|j+1,...,j+k}q_{j+k+1\|j+1,...,j+k}$
(1,1)	$(1 - q_{j\|j+1,...,j+k})(1 - q_{j+k+1\|j+1,...,j+k})$

Table 26: Score function of conditional bivariate distribution from tree k

$(y_j, y_{j+k+1})\|y_{j+1}, ..., y_{j+k}$	$\frac{\partial}{\partial\theta}f(y_j, y_{j+k+1}\|y_{j+1}, ..., y_{j+k})$
(0,0)	$\frac{\partial q_{j\|j+1,...,j+k}}{\partial\theta}q_{j+k+1\|j+1,...,j+k}+$ $\frac{\partial q_{j+k+1\|j+1,...,j+k}}{\partial\theta}q_{j\|j+1,...,j+k}$
(0,1)	$\frac{\partial q_{j\|j+1,...,j+k}}{\partial\theta} - \frac{\partial f(y_j=0,y_{j+k+1}=0\|y_{j+1},...,y_{j+k})}{\partial\theta}$
(1,0)	$\frac{\partial q_{j+k+1\|j+1,...,j+k}}{\partial\theta} - \frac{\partial f(y_j=0,y_{j+k+1}=0\|y_{j+1},...,y_{j+k})}{\partial\theta}$
(1,1)	$-\frac{\partial q_{j\|j+1,...,j+k}}{\partial\theta} - \frac{\partial q_{j+k+1\|j+1,...,j+k}}{\partial\theta}+$ $\frac{\partial f(y_j=0,y_{j+k+1}=0\|y_{j+1},...,y_{j+k})}{\partial\theta}$

In fact, the further trees will go just like tree 3, that the previous results will be used to obtain both the value of conditional bivariate PMF and derivative of it. Table 25 and 26 show tree k for example, where derivative results from tree $k-2$, $k-1$ will be needed.

III.4 PARAMETER ESTIMATION

The loglikelihood is shown at the beginning of this chapter in Equation (3.2.26). The score function is the derivative of the loglikelihood with respect to $\theta = (\rho, \beta)$, and it can be evaluated using the partial derivatives given in Section III.3.

$$
\begin{aligned}
\ell'(\theta|y,x) &= \sum_{i=1}^{n} \frac{\partial f(y_i)/\partial\theta}{f(y_i)} \\
&= \sum_{i=1}^{n} \left\{ \frac{\partial f(y_{i1}, y_{i,m}|y_{i2}, .., y_{i,m-1})/\partial\theta}{f(y_{i1}, y_{i,m}|y_{i2}, .., y_{i,m-1})} + \frac{\partial f(y_{i2}, y_{i,m-1}|y_3, .., y_{i,m-2})/\partial\theta}{f(y_{i2}, y_{i,m-1}|y_3, .., y_{i,m-2})} \cdots \right. \\
&\qquad \left. + \frac{\partial f(y_{i,\frac{m+1}{2}})/\partial\theta}{f(y_{i,\frac{m+1}{2}})} \right\} \qquad \text{if } m \text{ is an odd number;} \\
&= \sum_{i=1}^{n} \left\{ \frac{\partial f(y_{i1}, y_{i,m}|y_{i2}, .., y_{i,m-1})/\partial\theta}{f(y_{i1}, y_{i,m}|y_{i2}, .., y_{i,m-1})} + \frac{\partial f(y_{i2}, y_{i,m-1}|y_3, .., y_{i,m-2})/\partial\theta}{f(y_{i2}, y_{i,m-1}|y_3, .., y_{i,m-2})} \cdots \right. \\
&\qquad \left. + \frac{\partial f(y_{i,\frac{m}{2}}, y_{i,\frac{m}{2}+1})/\partial\theta}{f(y_{i,\frac{m}{2}}, y_{i,\frac{m}{2}+1})} \right\} \qquad \text{if } m \text{ is an even number.}
\end{aligned}
$$

The maximum likelihood estimates are obtained solving $\ell'(\theta|y,x) = 0$ using non-linear routines in R. The standard errors of the ML estimates are obtained by inverting the observed Fisher information matrix computed numerically.

III.5 SIMULATION STUDY

III.5.1 COMPARISONS BASED ON ASYMPTOTIC EFFICIENCY

To compare the performance of the D-vine pair-copula model against the multivariate probit model, in this section we perform some simulation studies. We consider three dimensional balanced longitudinal binary data. We use two covariates, the first covariate x^c_{it1}, $t = 1, 2, 3$ and $i = 1, 2, \ldots, n$, is continuous and distributed as uniform on (0,1) and represents a time-varying factor. The second covariate $x^d_{it2} = t$, for

$t = 1, 2, 3$ and for all $i = 1, \ldots, n$, and it represents a fixed discrete covariate. The regression model is given by

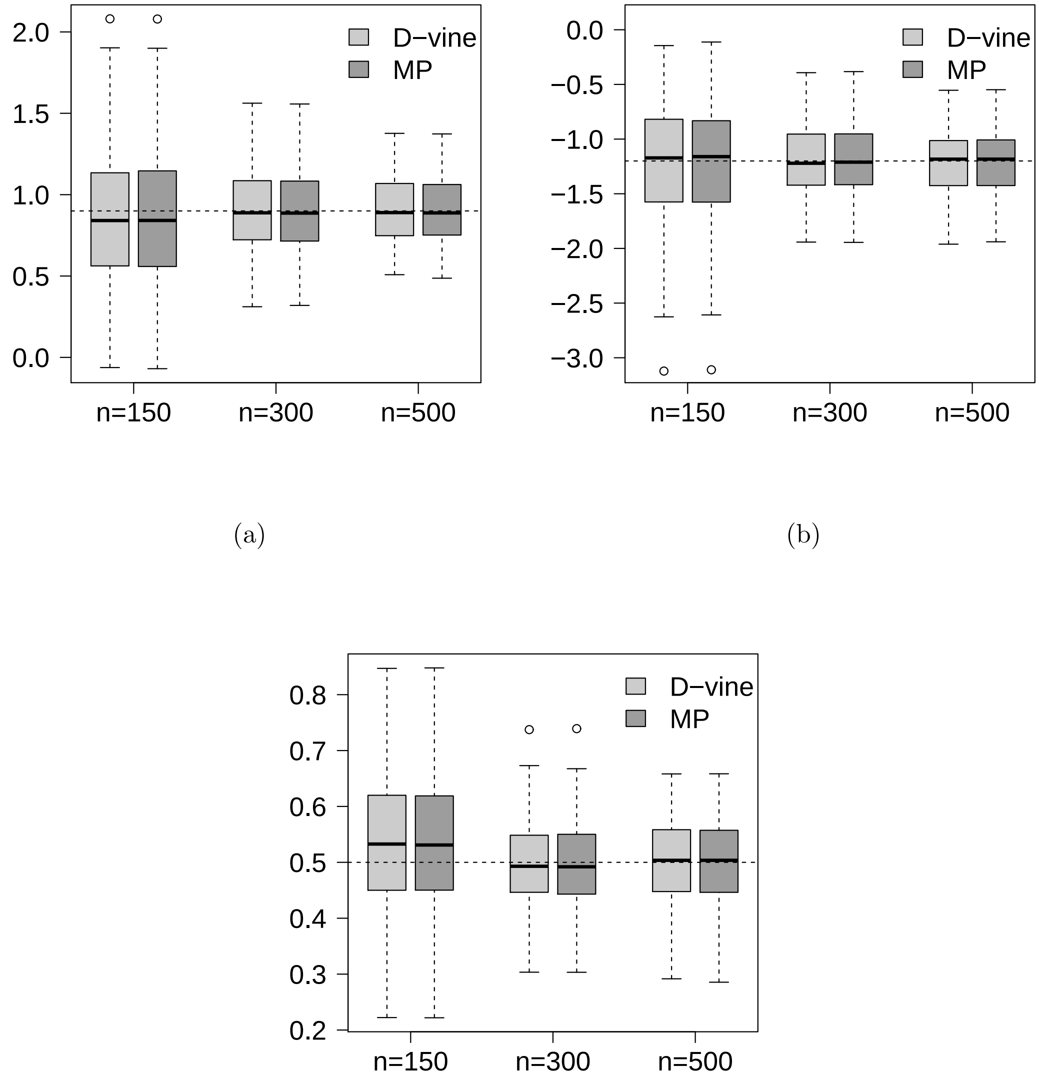
$$\text{logit}(p_{it}) = \beta_0 + \beta_1 \, x_{it1}^c + \beta_2 \, x_{it2}^d.$$

The true regression coefficients were fixed as $\beta_0 = 0.9$, $\beta_1 = -1.2$ and $\beta_2 = 0.5$. We assume AR(1) structure with true correlation value taken as $\rho = 0.4$. With these parameter values we simulated samples of sizes $n = 150, 300$, and 500 from the multivariate probit model of dimension three.

Table 27: The summary of estimates from both models

		Gaussian D-vine		**MP**		**Efficiency**
		Estimate	Standard Error	Estimate	Standard Error	
β_0	n=150	0.873	0.408	0.874	0.410	1.011
	n=300	0.907	0.252	0.903	0.254	1.012
	n=500	0.912	0.209	0.912	0.211	1.018
β_1	n=150	-1.210	0.556	-1.207	0.556	1.002
	n=300	-1.204	0.351	-1.198	0.352	1.007
	n=500	-1.210	0.283	-1.210	0.285	1.017
β_2	n=150	0.521	0.136	0.521	0.136	0.988
	n=300	0.498	0.087	0.498	0.087	0.988
	n=500	0.499	0.076	0.499	0.076	1.014

The results of fitting D-vine pair-copula and the MP model for the simulated samples are given in the Table 27 and the box-plots in Figure 6. The estimates of the regression coefficients from using the D-vine pair-copula model are all close to the true regression coefficients, and the standard errors are comparable to the MP model. Also the standard errors are smaller with increased sample size, for example, the standard error of $\widehat{\beta}_2$ is 0.136, 0.087 and 0.076 using D-vine pair-copula for sample sizes of 150, 300 and 500, respectively. This shows that the D-vine pair-copula model is consistently estimating the regression coefficients. The efficiency of the MP model with respect to the D-vine pair-copula model as measured by the ratio of the variances are in the range 0.988 to 1.018, demonstrating the D-vine pair-copula model is a good competitor to the true MP model.



(a)

(b)

(c)

Figure 6: Boxplot of regression coefficients estimates from simulation with $\beta = (0.9, -1.2, 0.5)$ and $\rho = 0.4$: (a) Intercept β_0, (b) regression coefficient of time-varying covariate, β_1, (c) regression coefficient of fixed covariate, β_2

Table 28: The mean, variance and bias of mle of regression coefficients from large sample simulation with $\beta = (0.9, -1.2, 0.5)$

		Gaussian D-vine			**MP**			**ARE**
		Mean	Var.	Bias	Mean	Var.	Bias	
β_0	$\rho = -0.1$	0.9041	0.0277	0.0041	0.9033	0.0277	0.0033	0.9988
	$\rho = 0.0$	0.9098	0.0387	0.0098	0.9098	0.0387	0.0098	1.0000
	$\rho = 0.1$	0.9127	0.0355	0.0127	0.9122	0.0354	0.0122	0.9969
	$\rho = 0.2$	0.9016	0.0388	0.0016	0.8994	0.0389	-0.0006	1.0012
	$\rho = 0.3$	0.8768	0.0456	-0.0232	0.8761	0.0460	-0.0239	1.0088
	$\rho = 0.4$	0.9127	0.0439	0.0127	0.9116	0.0469	0.0116	1.0179
	$\rho = 0.5$	0.8813	0.0361	-0.0187	0.8793	0.0364	0.0207	1.0094
	$\rho = 0.6$	0.8902	0.0467	-0.0098	0.8869	0.0474	-0.0131	1.0147
	$\rho = 0.7$	0.9102	0.0434	0.0102	0.9003	0.0451	0.0003	1.0165
β_1	$\rho = -0.1$	-1.1462	0.0265	0.0538	-1.1442	0.0269	0.0558	1.0018
	$\rho = 0.0$	-1.2003	0.0591	-0.0003	-1.2003	0.0591	-0.0003	0.9999
	$\rho = 0.1$	-1.1816	0.0582	0.0184	-1.1814	0.0581	0.0186	0.9986
	$\rho = 0.2$	-1.2154	0.0687	-0.0154	-1.2176	0.0696	-0.0176	1.0127
	$\rho = 0.3$	-1.1685	0.0729	0.0315	-1.1681	0.0728	0.0319	0.9977
	$\rho = 0.4$	-1.2105	0.0801	-0.0105	-1.2101	0.0815	-0.0101	1.0172
	$\rho = 0.5$	-1.1773	0.0936	0.0227	-1.1734	0.0941	0.0266	1.0056
	$\rho = 0.6$	-1.1978	0.0930	0.0022	-1.1933	0.0932	0.0067	1.0024
	$\rho = 0.7$	-1.1855	0.1068	0.0145	-1.1700	0.0191	0.0300	1.0219
β_2	$\rho = -0.1$	0.4816	0.0060	-0.0184	0.4812	0.0059	-0.0188	0.9857
	$\rho = 0.0$	0.4949	0.0053	-0.0051	0.4949	0.0053	-0.0051	0.9999
	$\rho = 0.1$	0.4909	0.0052	-0.0091	0.4910	0.0052	-0.0090	1.0043
	$\rho = 0.2$	0.5001	0.0062	0.0001	0.4999	0.0062	-0.0001	0.9991
	$\rho = 0.3$	0.4970	0.0047	-0.0030	0.4972	0.0047	-0.0028	1.0019
	$\rho = 0.4$	0.4992	0.0057	-0.0008	0.4992	0.0058	-0.0008	1.0143
	$\rho = 0.5$	0.5077	0.0044	0.0077	0.5071	0.0047	0.0071	1.0064
	$\rho = 0.6$	0.4987	0.0043	-0.0013	0.4985	0.0043	-0.0015	0.9995
	$\rho = 0.7$	0.4965	0.0013	-0.0035	0.4966	0.0013	-0.0034	1.0034

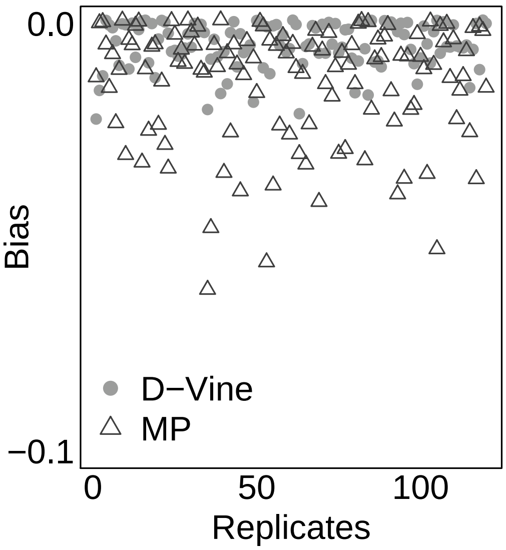


Figure 7: Bias of correlation coefficients ρ of 3 dimension simulation with $\beta = (0.9, -1.2, 0.5)$ and $\rho = 0.778$, which is the upper boundary of feasible range

Next to study the behavior of the bias and the asymptotic relative efficiency (ARE) as a function of ρ, we looked at 1000 replications of simulated samples of size $n = 500$ for different values of ρ. For each value of ρ using the 1000 estimates of the regression coefficients we calculated the mean, variance, and bias for each of the two models. We computed ARE taking the ratio of the mean square errors of the MP model over the D-vine pair-copula model. A value of the ARE more than one indicates D-vine pair-copula model is better than the MP model. The results are summarized in Table 28. The D-vine Gaussian copula model has larger bias for the intercept β_0, smaller bias for β_1, and almost the same value for β_2 when compared to the MP model. Also, AREs are in the range 0.9857 to 1.0219, as ρ traverses in the feasible range. This demonstrates that the D-vine Gaussian pair-copula model is performing well on a data where the true model is the MP model.

The bias of the correlation estimate for the replicates is given in Figure 7 for $\rho = 0.78$ which is close to the upper boundary in the feasible region. The figure shows the bias is smaller for the D-vine model compared to the MP model for estimating ρ.

Thus our simulations show that the D-vine model is superior not only in estimating the regression parameters but also in estimating the correlation parameter.

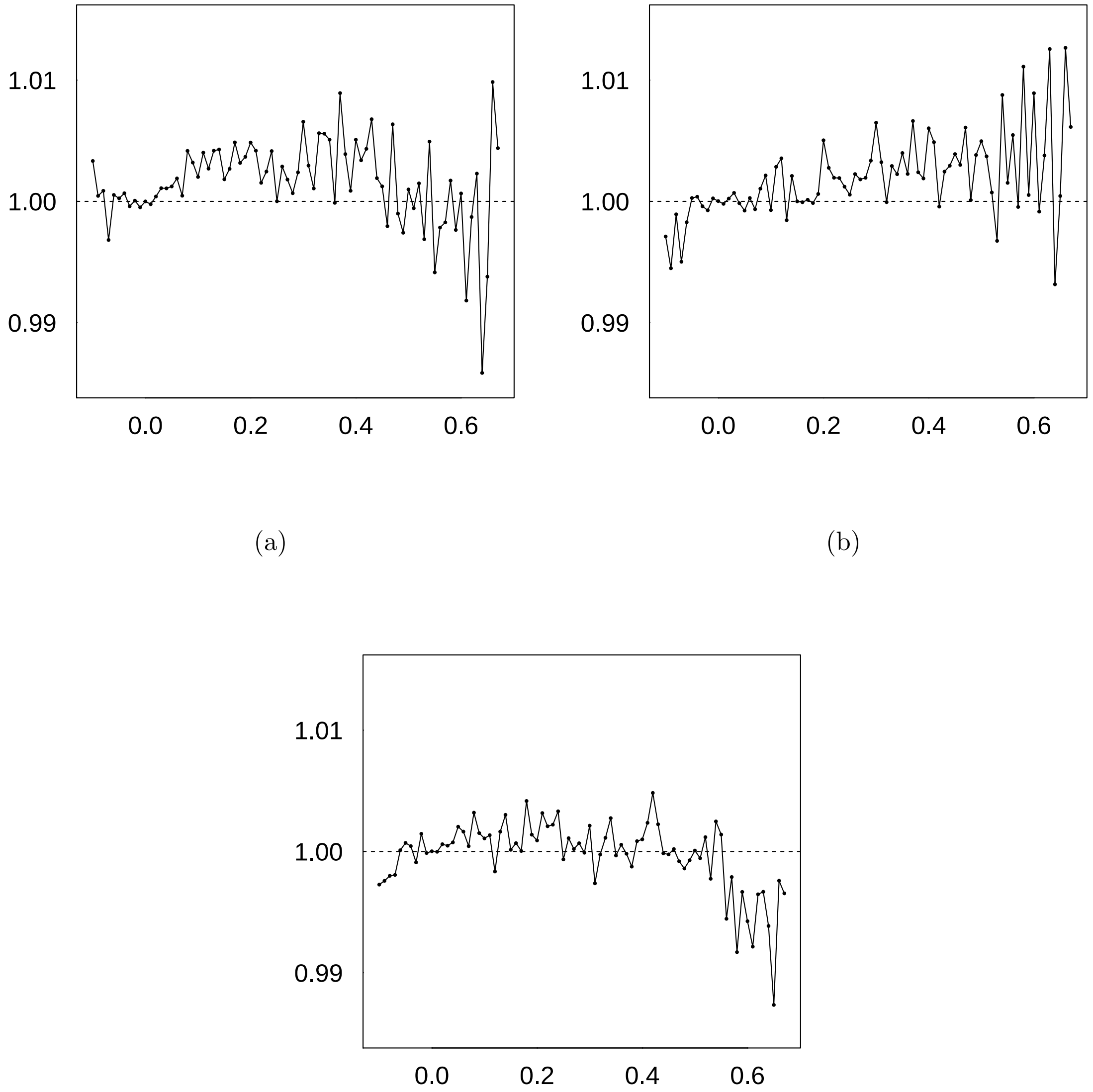
III.5.2 COMPARISONS BASED ON SMALL SAMPLE EFFICIENCY

To compare the small sample efficiencies, we used the same covariates x^c and x^d, regression coefficients β's with the values as in the previous section. However, we took a small sample size $n = 30$ for each replication. The number of repeated measurements were taken to be 3 and then 5. We did 1000 replications in order to calculate the mean square error of the regression and correlation parameters. Comparisons were made using the ratio of mean square errors (MSE) of the multivariate probit model and the D-vine Gaussian copula model.

For the small sample with three dimensions, the MP works better for the coefficients around the boundaries (for example, $-0.13 \leq \rho \leq 0$ or $0.5 \leq \rho \leq 0.67$), while the D-vine Gaussian copula model works better for the coefficients in the middle (for example, $0 \leq \rho \leq 0.5$). With five dimensions, the D-vine Gaussian copula model works better for the coefficients β_0 and β_1, except around the upper boundary, while the MP model is more efficient in estimating β_2.

We also ran 1000 replicates to check estimation of the binary variable correlation, ρ, at the upper boundary 0.67, as in our example. As shown in Figure 10, for the D-Vine Gaussian copula model, most of the bias of estimate of ρ is around 0, while for the MP model there is significant bias.

In fact, the estimate of ρ from the D-Vine model has mean 0.667 and standard deviation of 0.008, while MP model has a mean 0.614 and standard deviation 0.052. Thus D-Vine Gaussian copula model performs better than the MP model to estimate the correlation coefficient around the boundary of feasible range.

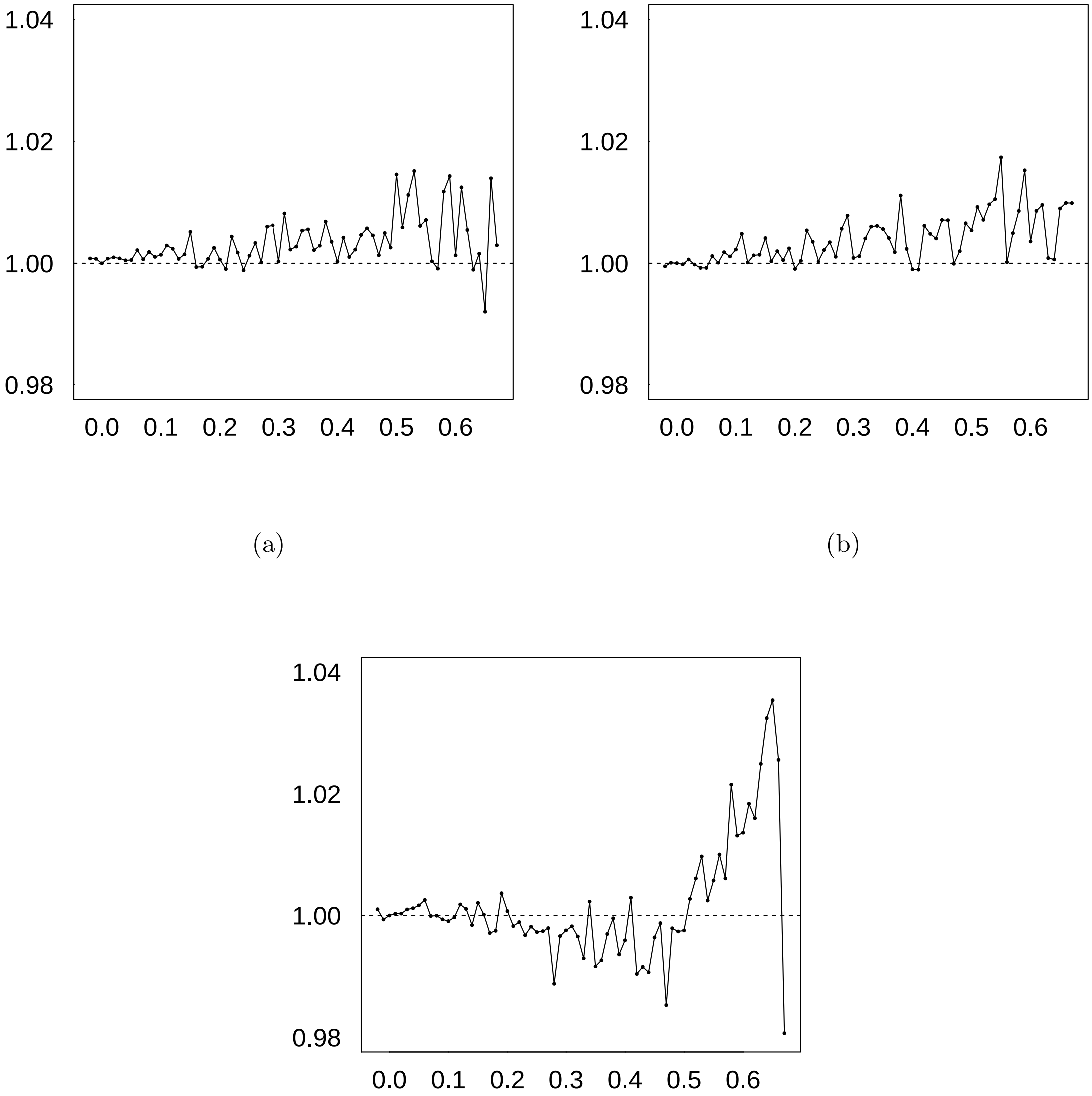


(a)

(b)

(c)

Figure 8: Plot of RMSE for regression coefficients of 3 dimension simulation with $\beta = (-1, 1, 0.8)$ and ρ within feasible range $(-0.1358, 0.6703)$, (a) intercept β_0, (b) regression coefficient of time-varying covariate, (c) regression coefficient of fixed covariate



(a)

(b)

(c)

Figure 9: Plot of RMSE for regression coefficients of 5 dimension simulation with $\beta = (-1, 1, 0.8)$ and ρ within feasible range $(-0.0277, 0.6703)$, (a) intercept β_0, (b) regression coefficient of time-varying covariate, (c) regression coefficient of fixed covariate

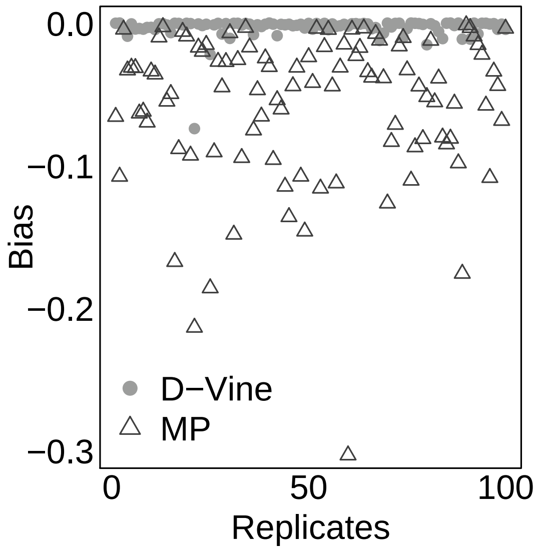


Figure 10: Bias of correlation coefficients ρ of 3 dimension simulation with $\beta = (-1, 1, 0.8)$ and $\rho = 0.67$, which is the upper boundary of feasible range

III.6 DATA ANALYSIS

III.6.1 OBESITY DATA

We analyze a subset of the Obesity data from the Muscatine coronary risk factor study conducted during 1977, 1979 and 1981, and reported by Woolson and Clarke (1984). In this study, 4856 school-aged kids were classified as obese if their weight was 210% or more than the median weight given their gender, age and height. We took a subset of 1700 children who hadn't missed any survey. The response variables are obesity indicators (1 for obese kid, 0 for non-obese kid) in the three years 1977, 1979 and 1981, the independent variables are baseline age (age at year 1977) and gender. We ruled out the effect of gender, because insignificant relationship was found between gender and obesity. Therefore, the covariates are baseline age (x_i^{age}), time at observation (x_{it}^{time}) and interaction between baseline age and time. The covariate

x_{it}^{time} takes values $1, 2, 3$ for the three years 1977, 1979 and 1981, respectively.

Table 29: Parameter estimates for the obesity data

	MP			**D-vine**		
Parameter	EST	SE	p-Value	EST	SE	p-Value
Intercept	-2.671	0.0.064	< 0.001	-2.683	0.061	< 0.001
Baseline Age	0.203	0.047	< 0.001	0.202	0.009	< 0.001
Time	0.679	0.028	< 0.001	0.678	0.022	< 0.001
Time×Baseline Age	0.010	0.005	0.023	0.012	0.005	0.016
ρ	0.638	0.006	< 0.001	0.638	0.008	< 0.001
AIC	17.050			17.040		

NOTE: Range of ρ is (0, 0.7659).

We fit the binary regression model

$$\text{logit}(p_{it}) = \beta_0 + \beta_1 x_i^{age} + \beta_2 x_{it}^{time} + \beta_3 x_i^{age} x_{it}^{time},$$

where $i = 1, 2, ..., 1765$; and $t = 1, 2, 3$. The point estimates, standard errors and p-values for both the MP model and D-vine pair-copula model are presented in Table 29. The parameter estimates seem to be in agreement for both the models. The p-values indicate baseline age, time and the interaction are significant factors . The estimated regression coefficients of baseline age and time are positive, which means that the older subject is more likely to have obesity issue, and the subject is more likely to have obesity issue as time goes.

III.6.2 RESPIRATORY ILLNESS DATA

As a second example we consider the clinical trials data from Stokes et al. (1995) of the SAS Institute, Inc. This clinical study compares two approaches for treating respiratory disease. In the trial there were 111 patients from two different clinics (center 1 is denoted as 1, and center 2 is denoted as 2). The patients were randomly assigned to receive placebo (denoted as 1) or aggressive treatment (denoted as 0) for their respiratory illness, as summarized in Table 30.

Table 30: Respiratory illness data

Treatment	Center 1	Center 2
Placebo	29	28
Active	27	27

The patients were examined for respiratory illness at baseline and at four follow up visits, recording breathing condition as 1 for good response and 0 for poor. Then, we ruled out the effect from clinic, because insignificant relationship was found between the different center and response. Therefore, the covariates are treatment (x_i^{treat}), baseline response (x_i^{base}), and time at observation (x_{it}^{time}). The covariate x_i^{treat} takes values 0, 1 for aggressive treatement and placebo, respectively; x_i^{base} takes values 0, 1 for poor and good response, respectively; x_{it}^{time} takes values 1, 2, 3, 4 for the four visit time points $t = 1, 2, 3, 4$ respectively. And the interaction between baseline examination, treatment and time were ruled out. The regression model is:

$$\text{logit}(p_{it}) = \beta_0 + \beta_1 x_i^{treat} + \beta_2 x_i^{base} + \beta_3 x_{it}^{time},$$

where $i = i, 2, ..., 111$, and $t = 1, 2, 3, 4$. The point estimates, standard errors and p-values for both the MP model and D-vine pair-copula model are presented in Table 31 as below.

Results in Table 31 are in agreement between the MP model and the D-vine pair-copula model, but the intercept with small value is insignificant. We run the model without intercept and the results is in Table 32.

Again, the MP model and the D-vine pair-copula model have similar results, although, AIC of the MP model is a little larger than the value of the D-vine pair-copula model. The p-values indicate that treatment, baseline response and time when repeated the measure are all significant, while the estimate of regression coefficient of time is negative, which means that the subject is less likely to get good respiratory response as time goes; and the estimates of regression coefficients of the baseline response is positive, which indicate that subject is more likely to get respiratory disease as they have the disease at baseline check; and the estimates of regression coefficients of the treatment is positive, which indicate that subjects of taking aggressive treatment is more likely to get good respiratory response than the subjects of taking placebo, or proves the effectiveness of the aggressive treatment.

Table 31: Parameter estimates for the respiratory illness data

	MP			**D-vine**		
Parameter	EST	SE	p-Value	EST	SE	p-Value
Intercept	0.187	0.163	0.250	0.169	0.123	0.169
Treatment	-1.308	0.178	< 0.001	-1.287	0.102	< 0.001
Baseline Response	2.199	0.182	< 0.001	2.199	0.065	< 0.001
Time	-0.385	0.140	0.001	-0.399	0.063	< 0.001
ρ	0.690	0.098	< 0.001	0.709	0.062	< 0.001
AIC	16.242			16.244		

NOTE: Range of ρ is (-0.283, 0.951).

Table 32: Parameter estimates for the respiratory illness data

	MP			**D-vine**		
Parameter	EST	SE	p-Value	EST	SE	p-Value
Treatment	-1.266	0.171	< 0.001	-1.255	0.134	< 0.001
Baseline Response	2.190	0.190	< 0.001	2.224	0.099	< 0.001
Time	-0.307	0.150	0.041	-0.351	0.066	< 0.001
ρ	0.657	0.118	< 0.001	0.692	0.083	< 0.001
AIC	14.451			14.348		

Or we can show the probability of good respiratory condition based on the estimation from Table 32. For example, to predict the probability of good respiratory response (p_{i4}) at time point 4 ($x_{i4}^{time} = 4$) for subject who has poor response at baseline ($x_i^{base} = 0$) and takes aggressive treatment ($x_i^{treat} = 0$) using D-vine pair-copula model is as below:

$$\begin{aligned} \text{logit}(p_{it}) &= -1.255x_i^{treat} + 2.224x_i^{base} - 0.351x_{it}^{time}, \\ \text{logit}(p_{i4}) &= -1.255 * 0 + 2.224 * 0 - 0.351 * 4, \\ \text{logit}(p_{i4}) &= -1.404, \\ p_{i4} &= \frac{1}{1 + e^{-1.404}} = 0.1972. \end{aligned}$$

Similarly, the probabilities of good respiratory response of other situations can be calculated and presented in the Table 33. The predicted probabilities of good response from MP model is given in the brackets.

The same conclusion we can draw with Table 33: respiratory condition gets worse as time goes; if the breath condition is good at baseline check, then it is more likely to have a good condition at follow-up checks; the active treatment is effective, because the probabilities of good response with active treatment are improved compared to the probabilities of good response with placebo.

Table 33: Probability of good respiratory condition

	Baseline Response	**Treatment**	**Time**			
			1	2	3	4
D-vine	Good	Placebo	0.6498	0.5664	0.4790	0.3929
		Treatment	0.8668	0.8208	0.7633	0.6942
	Poor	Placebo	0.1671	0.1238	0.0905	0.0654
		Treatment	0.4131	0.3314	0.2586	0.1972
MP	Good	Placebo	0.6495	0.5769	0.5007	0.4246
		Treatment	0.8680	0.8286	0.7806	0.7235
	Poor	Placebo	0.1718	0.1324	0.1009	0.0763
		Treatment	0.4238	0.3511	0.2848	0.2265

III.7 CONCLUSIONS

A popular method of analyzing high-dimensional longitudinal binary data is the multivariate probit model. In this chapter, we discussed an alternative to the MP model for analyzing AR(1) structured longitudinal binary data: D-vine pair Gaussian copula model. We conducted the comparison based on actual likelihoods. On the asymptotical simulation study, we showed that efficiency results from the D-vine Gaussian copula model and MP model are very close overall, except that the D-vine Gaussian copula works better to estimate the regression coefficients with larger correlation coefficients, while the MP model works better with correlation coefficients around lower boundaries. For the small samples, the MP works better to estimate the regression coefficients around the boundaries, while the D-vine Gaussian copula model works better for regression with the moderately correlated data. However, the D-vine Gaussian copula works more properly to estimate the correlation coefficients around the boundaries for both small or large samples.

CHAPTER IV

D-VINE PAIR-COPULA MODEL FOR MULTINOMIAL VARIABLES

IV.1 INTRODUCTION

We introduced the pair-copula Gaussian model to analyze longitudinal binary data in Chapter II and discussed parameter estimation for the associated regression model in the next Chapter III. In this chapter we extend the models from binary marginals to a more complicated situation where the marginal could be multinomial with ordered categories. For instance, in example 3 of Section II.6, the recorded responses were binary of the subjects who were treated treated with three different medicines at three different times. But suppose that the response was recorded as unfavorable, neutral, favorable, or even more specifically as extremely unfavorable, moderately unfavorable, neutral, moderately favorable, extremely favorable. This detailed response is an example of a multinomial distribution.

D-vine pair-copula model for longitudinal multinomial variables is one application of pair-copula constructions for multivariate discrete data by Panagiotelis et al. (2012). There are some R packages developed for generating associated multinomial responses, such as SimCorMultRes by Touloumis (2016), GenOrd by Barbiero and Ferrari (2015), MultiOrd by Amatya and Demirtas (2015), and a method using the convex combination by Ibrahim and Suliadi (2011).

Similar to AR(1) or equicorrelated structured longitudinal binary data, there is a feasible range for the correlation parameter for fixed success probabilities for longitudinal multinominal resposes. Even if the correlation coefficient value is within the feasible range, the MP model may still fail to generate a probability distribution for the longitudinal multinomial responses. In this chapter, we propose the D-Vine pair-copula Gaussian model to generate a proper multidimensional distribution. Our method works even when correlation parameter is around the boundaries, whereas other methods fail to generate a probability distribution. The organization of this chapter is as follows. Section IV.2 gives the background and detailed information of our D-vine pair-copula model. Section IV.3 contains some numerical examples, including ones that other methods don't work. Section IV.4 presents the results from simulation studies.

IV.2 DISTRIBUTION OF MULTINOMIAL VARIABLE USING PAIR-COPULA

It is well known that the multinomial distribution is a generalization of the Bernoulli distribution. The probability mass function of a multinomial variable Y, which has c possible outcomes with fixed success probability, can be written as

$$f(y; p_1, p_2, ..., p_c) = \sum_{j=1}^{c} p_j \, I(y = j),$$

where $I(\cdot)$ is the indicator function and $\Sigma_{j=1}^{c} p_j = 1$.

IV.2.1 MULTINOMIAL VARIABLE WITH THREE CATEGORIES

As we have done in Section II.2, we will start with the simplest situation. We first consider bivariate multinomial variables $Y = (Y_1,\ Y_2)$ with Y_j taking $c = 3$ possible outcomes: 1, 2, 3 with respective success probabilities $p_{i,1}$, $p_{i,2}$ and $p_{i,3}$, where $p_{1,1} + p_{1,2} + p_{1,3} = 1$ and $p_{2,1} + p_{2,2} + p_{2,3} = 1$, where $p_{i,j}$ represents the probability of $Y_i = j$, for $i = 1, 2$ and $j = 1, 2, 3$. The CDF of Y_i can be written as

$$F(y_i) = \begin{cases} 0 & \text{if} \quad y_i < 0, \\ p_{i,1} & \text{if} \quad 0 \le y_i < 1, \\ p_{i,1} + p_{i,2} & \text{if} \quad 1 \le y_i < 2, \\ 1 & \text{if} \quad 2 \le y_i. \end{cases}$$

The joint CDF of Y_1 and Y_2 using a copula C function would be, according to Sklar's Theorem proposed in Sklar (1959), $F(Y_1 = y_1,\ Y_2 = y_2) = C(F(y_1),\ F(y_2);\ \theta_{12})$, where θ_{12} is the copula parameter. The probability mass function of bivariate multinomial variables with 3 categories, according to Equation (2.2.5), is listed in Table 34.

Table 34: PMF of bivariate multinomial variables with three categories

$(Y_1,\ Y_2)$	**Probability**
(1, 1)	$C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(1, 2)	$C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12}) - C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(1, 3)	$p_{1,1} - C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(2, 1)	$C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12}) - C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(2, 2)	$C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12}) - C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})$ $-C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12}) + C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(2, 3)	$p_{1,2} - C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12}) + C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(3, 1)	$p_{2,1} - C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})$
(3, 2)	$p_{2,2} - C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12}) + C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})$
(3, 3)	$1 - p_{2,1} - p_{2,2} - p_{1,1} - p_{1,2} + C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12})$

NOTE: this distribution has 5 parameters: $p_{1,1}$, $p_{1,2}$, $p_{2,1}$, $p_{2,2}$, & copula parameter θ_{12}.

As in Section II.2.3, the D-vine pair-copula decomposes the probability distribution of trivariate (Y_1, Y_2, Y_3) multinomial variables as the product of the marginal distribution of Y_2 and conditional distribution of $(Y_1,\ Y_3)$ given Y_2, that is,

$$f(y_1, y_2, y_3) = f(y_2)f(y_1, y_3|y_2).$$

Let $p_{1|2,1|1} = P(Y_1 = 1|Y_2 = 1)$, the first two subscripts of p are the subscripts of the random variables and the second two subscripts are the values they take. Then

$$\begin{aligned}
p_{1|2,1|1} &= \frac{C(p_{1,1},\ p_{2,1};\ \theta_{12})}{p_{2,1}}, \\
p_{1|2,2|1} &= P(Y_1 = 2|Y_2 = 1) = \frac{C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12}) - C(p_{1,1},\ p_{2,1};\ \theta_{12})}{p_{2,1}}, \quad \text{and,} \\
p_{1|2,3|1} &= P(Y_1 = 3|Y_2 = 1) = \frac{1 - C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})}{p_{2,1}}.
\end{aligned}$$

Note that $p_{1|2,1|1} + p_{1|2,2|1} + p_{1|2,3|1} = \frac{1}{p_{2,1}}$. Similarly, we can write down the formulas for $p_{1|2,1|2},\ p_{1|2,2|2},\ \ldots,\ p_{1|2,3|2},\ p_{3|2,1|1},\ p_{3|2,2|1},\ \ldots,\ p_{3|2,2|3},\ p_{3|2,3|3}$, and we will have $p_{1|2,1|2} + p_{1|2,2|2} + p_{1|2,3|2} = \frac{1}{p_{2,2}}, \ldots,\ p_{3|2,1|3} + p_{3|2,2|3} + p_{3|2,3|3} = \frac{1}{p_{2,3}}$.

Table 35: PMF of trivariate multinomial variables with three categories

$(Y_1,\ Y_2,\ Y_3)$	**Probability**
(1, 1, 1)	$p_{2,1}C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1})$
(1, 1, 2)	$p_{2,1}(C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1})-C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1});\ \theta_{13\mid Y_2=1})$
(1, 1, 3)	$p_{2,1}(p_{1\mid 2,1\mid 1}-C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1}))$
(1, 2, 1)	$p_{2,2}C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2})$
(1, 2, 2)	$p_{2,2}(C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2})-C(p_{1\mid 2,1\mid 2},p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2}))$
(1, 2, 3)	$p_{2,2}(p_{1\mid 2,1\mid 2}-C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2}))$
(1, 3, 1)	$p_{2,3}C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3})$
(1, 3, 2)	$p_{2,3}(C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3})-C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3}))$
(1, 3, 3)	$p_{2,3}(p_{1\mid 2,1\mid 3}-C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3}))$
(2, 1, 1)	$p_{2,1}(C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1})-C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1}))$
(2, 1, 2)	$p_{2,1}(C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1})- C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1})- C(p_{1\mid 2,1\mid 1},p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1})+C(p_{1\mid 2,1\mid 1},p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1}))$
(2, 1, 3)	$p_{2,1}(p_{1\mid 2,2\mid 1}-C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1})+ C(p_{1\mid 2,1\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1}))$
(2, 2, 1)	$p_{2,2}(C(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2},\ p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2})-C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2}))$
(2, 2, 2)	$p_{2,2}(C(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2})- C(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2},\ p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2})- C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2})+C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2};\ \theta_{13\mid Y_2=2}))$
(2, 2, 3)	$p_{2,2}(p_{1\mid 2,2\mid 2}-C(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2})+ C(p_{1\mid 2,1\mid 2},\ p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2};\ \theta_{13\mid Y_2=2}))$
(2, 3, 1)	$p_{2,3}(C(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},\ p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3})-C(p_{1\mid 2,1\mid 3},p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3}))$
(2, 3, 2)	$p_{2,3}(C(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3})- C(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},\ p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3})- C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3})+C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3};\ \theta_{13\mid Y_2=3}))$
(2, 3, 3)	$p_{2,3}(p_{1\mid 2,2\mid 3}-C(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3})+ C(p_{1\mid 2,1\mid 3},\ p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3};\ \theta_{13\mid Y_2=3}))$
(3, 1, 1)	$p_{2,1}(p_{3\mid 2,1\mid 1}-C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1}))$
(3, 1, 2)	$p_{2,1}(p_{3\mid 2,2\mid 1}-C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1};\ \theta_{13\mid Y_2=1})+ C(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1},\ p_{3\mid 2,1\mid 1};\ \theta_{13\mid Y_2=1}))$

Continued

(3, 1, 3)	$p_{2,1}(1 - p_{3\|2,1\|1} - p_{3\|2,2\|1} - p_{1\|2,1\|1} - p_{1\|2,2\|1} +$ $C(p_{1\|2,1\|1} + p_{1\|2,2\|1},\ p_{3\|2,1\|1} + p_{3\|2,2\|1};\ \theta_{13\|Y_2=1}))$
(3, 2, 1)	$p_{2,2}(p_{3,1} - C(p_{1\|2,1\|2} + p_{1\|2,2\|2},\ p_{3\|2,1\|2};\ \theta_{13\|Y_2=2}))$
(3, 2, 2)	$p_{2,2}(p_{3\|2,2\|2} - C(p_{1\|2,1\|2} + p_{1\|2,2\|2},\ p_{3\|2,1\|2} + p_{3\|2,2\|2};\ \theta_{13\|Y_2=2}) +$ $C(p_{1\|2,1\|2} + p_{1\|2,2\|2},\ p_{3\|2,1\|2};\ \theta_{13\|Y_2=2}))$
(3, 2, 3)	$p_{2,2}(1 - p_{3\|2,1\|2} - p_{3\|2,2\|2} - p_{1\|2,1\|2} - p_{1\|2,2\|2} +$ $C(p_{1\|2,1\|2} + p_{1\|2,2\|2},\ p_{3\|2,1\|2} + p_{3\|2,2\|2};\ \theta_{13\|Y_2=2}))$
(3, 3, 1)	$p_{2,3}(p_{3,1} - C(p_{1\|2,1\|3} + p_{1\|2,2\|3},\ p_{3\|2,1\|3};\ \theta_{13\|Y_2=3}))$
(3, 3, 2)	$p_{2,3}(p_{3\|2,2\|3} - C(p_{1\|2,1\|3} + p_{1\|2,2\|3},\ p_{3\|2,1\|3} + p_{3\|2,2\|3};\ \theta_{13\|Y_2=3}) +$ $C(p_{1\|2,1\|3} + p_{1\|2,2\|3},\ p_{3\|2,1\|3};\ \theta_{13\|Y_2=3}))$
(3, 3, 3)	$p_{2,3}(1 - p_{3\|2,1\|3} - p_{3\|2,2\|3} - p_{1\|2,1\|3} - p_{1\|2,2\|3} +$ $C(p_{1\|2,1\|3} + p_{1\|2,2\|3},\ p_{3\|2,1\|3} + p_{3\|2,2\|3};\ \theta_{13\|Y_2=3}))$

NOTE: distribution has 11 parameters: $p_{1,1},\ p_{1,2},\ p_{2,1},\ p_{2,2},\ p_{3,1},\ p_{3,2}$, and copula parameters θ_{12}, θ_{23}, $\theta_{13|Y_2=1}$, $\theta_{13|Y_2=2}$ and $\theta_{13|Y_2=3}$.

For four dimensions or higher, the PMF can be constructed similarly by first obtaining the new success probabilities, followed by the decomposition of the joint PMF. For example,

$$f(y_1,\ y_2,\ y_3,\ y_4) = f(y_2, y_3) f(y_1,\ y_4 | y_2,\ y_3),$$

where $f(y_2, y_3)$ is similar to Table 34, and construction of $f(y_1, y_4 | Y_2 = y_2, Y_3 = y_3)$ would need the following conditional probabilities

$$\begin{aligned}
p_{1|23,1|11} &= P(Y_1 = 1 | Y_2 = 1,\ Y_3 = 1) = \frac{C(p_{1|2,1|1},\ p_{3|2,1|1};\ \theta_{13|Y_2=1})}{p_{3|2,1|1}}, \\
p_{1|23,2|11} &= P(Y_1 = 2 | Y_2 = 1,\ Y_3 = 1) = \frac{C(p_{1|2,2|1},\ p_{3|2,1|1};\ \theta_{13|Y_2=1})}{p_{3|2,1|1}}, \\
& \quad .... \\
p_{4|23,2|33} &= P(Y_4 = 2 | Y_2 = 3, Y_3 = 3) = \frac{C(p_{1|2,1|1},\ p_{3|2,1|1};\ \theta_{24|Y_3=3})}{p_{2|3,3|3}}.
\end{aligned}$$

The conditional bivariate distribution of $(Y_1,\ Y_4)$ for given values of (y_2, y_3) is similar to Table 34 with the these new success probabilities, for example when $(y_2 = 1,\ y_3 = 1)$ is given in Table 36.

Table 36: Conditional PMF of bivariate multinomial variables with three categories given $(Y_2 = 1, Y_3 = 1)$

$(Y_1,\ Y_4 \vert Y_2 = 1,\ Y_3 = 1)$	**Probability**
(1, 1)	$C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(1, 2)	$C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(1, 3)	$p_{1\vert 23,1\vert 11} - C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(2, 1)	$C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(2, 2)	$C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $+C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(2, 3)	$p_{1\vert 23,2\vert 11} + C(p_{1\vert 23,1\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(3, 1)	$p_{4\vert 23,1\vert 11} - C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(3, 2)	$p_{4\vert 23,2\vert 11} + C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$ $-C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$
(3, 3)	$1 - p_{1\vert 23,2\vert 11} - p_{4\vert 23,2\vert 11} +$ $C(p_{1\vert 23,1\vert 11} + p_{1\vert 23,2\vert 11},\ p_{4\vert 23,1\vert 11} + p_{4\vert 23,2\vert 11};\ \theta_{14\vert Y_2=1,Y_3=1})$

NOTE: distribution has 12 parameters: $p_{1,1}$, $p_{1,2}$, $p_{2,1}$, $p_{2,2}$, $p_{3,1}$, $p_{3,2}$, $p_{4,1}$, $p_{4,2}$, and copula parameters θ_{12}, θ_{23}, θ_{34}, $\theta_{14|Y_2=1,Y_3=1}$.

IV.2.2 RELATIONSHIP BETWEEN COPULA PARAMETER AND CORRELATION OF MULTINOMIAL VARIABLES WITH THREE CATEGORIES

In Section II.2.1 we studied the relation between the correlation of the binary variables and Gaussian copula parameter, which happens to be the correlation of the latent variables. In this section we will discuss the corresponding relation for multinomial variables. Let Y_i be multinomial with three categories and corresponding probabilities $p_{i,1}, p_{i,2}$, and $p_{i,3}$, for $i = 1, 2$. The mean of Y_i is

$$\mu_i = E(Y_i) \;=\; p_{i,1} * 1 + p_{i,2} * 2 + p_{i,3} * 3 = 3 - 2p_{i,1} - p_{i,2},$$

and the variance is

$$\begin{aligned} Var(Y_i) &= p_{i,1} * (1 - \mu_i)^2 + p_{i,2} * (2 - \mu_i)^2 + p_{i,3} * (3 - \mu_i)^2 \\ &= 4p_{i,1} + p_{i,2} - (2p_{i,1} + p_{i,2})^2, \end{aligned}$$

for $i = 1, 2$. Assume that the joint distribution of $(Y_1,\, Y_2)$ is given as in Table 34 with copula parameter θ_{12}. Then

$$\begin{aligned} E(Y_1, Y_2) &= 1 * f(1,1) + 2 * f(1,2) + ... + 9 * f(3,3) \\ &= C(p_{1,1},\; p_{2,1};\; \theta_{12}) + 2 * C(p_{1,1},\; p_{2,1} + p_{2,2};\; \theta_{12}) + ... \\ &= C(p_{1,1},\; p_{2,1};\; \theta_{12}) + C(p_{1,1},\; p_{2,1} + p_{2,2};\; \theta_{12}) + C(p_{1,1} + p_{1,2},\; p_{2,1};\; \theta_{12}) \\ &\quad + C(p_{1,1} + p_{1,2},\; p_{2,1} + p_{2,2};\; \theta_{12}) - 6p_{1,1} - 3p_{1,2} - 6p_{2,1} - 3p_{2,2} + 9. \end{aligned}$$

And the correlation betweeen (Y_1, Y_2) is

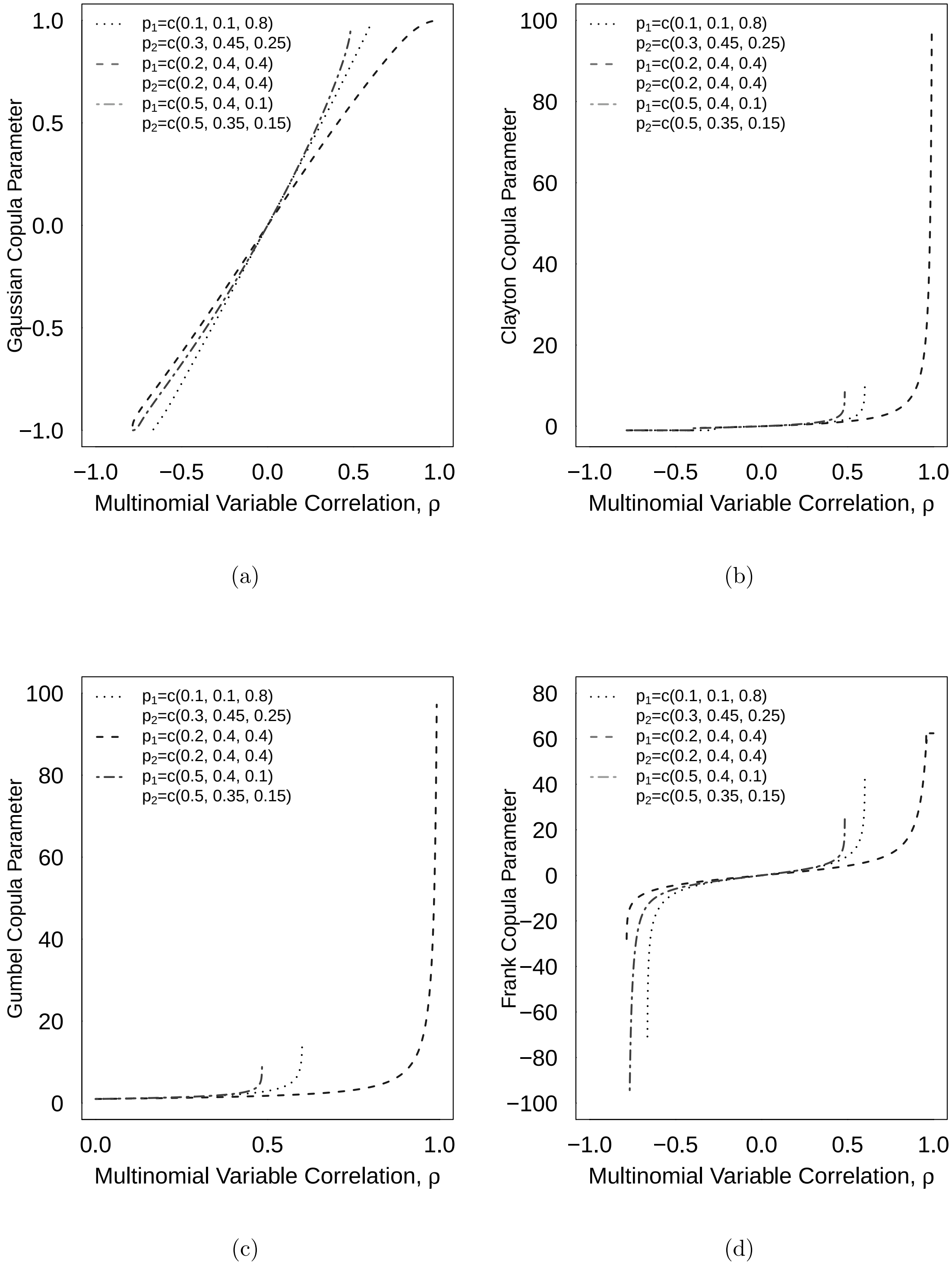
$$\text{Corr}(Y_1, Y_2) \;=\; \rho_{12} = \frac{E(Y_1, Y_2) - E(Y_1)E(Y_2)}{\sqrt{Var(Y_1)Var(Y_2)}}. \tag{4.2.30}$$

When we use the Gaussian copula, $C(u_1, u_2; \theta_{12}) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \gamma_{12})$ in $E(Y_1, Y_2)$, the above formula (4.2.30) reduces to

$$\rho_{12} \;=\; \frac{E(Y_1, Y_2) - (3 - 2p_{1,1} - p_{1,2}) * (3 - 2p_{2,1} - p_{2,2})}{\sqrt{(4p_{1,1} + p_{1,2} - (2p_{1,1} + p_{1,2})^2)(4p_{2,1} + p_{2,2} - (2p_{2,1} + p_{2,2})^2)}}, \tag{4.2.31}$$

where

$$\begin{aligned} E(Y_1, Y_2) &= \Phi_2(\Phi^{-1}(p_{1,1}), \Phi^{-1}(p_{2,1}); \gamma_{12}) + \Phi_2(\Phi^{-1}(p_{1,1}), \Phi^{-1}(p_{2,1} + p_{2,2}); \gamma_{12}) \\ &\quad + \Phi_2(\Phi^{-1}(p_{1,1} + p_{1,2}), \Phi^{-1}(p_{2,1}); \gamma_{12}) \\ &\quad + \Phi_2(\Phi^{-1}(p_{1,1} + p_{1,2}), \Phi^{-1}(p_{2,1} + p_{2,2}); \gamma_{12}) - 6p_{1,1} - 3p_{1,2} \\ &\quad - 6p_{2,1} - 3p_{2,2} + 9. \end{aligned}$$



(a)

(b)

(c)

(d)

Figure 11: Plots of relationship between copula parameter α or γ and multinomial variable correlation. (a) Gaussian copula; (b) Clayton copula; (c) Gumbel copula; (d) Frank copula

Similarly, the functional relationship between ρ_{12} and the parameter α of Clayton, Frank and Gumbel copulas can be obtained if we were to use those copulas instead of the Gaussian copula. A plot of these functions is given in Figure 11. Clearly, the plots depend on the copula and vary with the marginal probabilities. The plots also show that as the correlation varies, the copula parameter is also feasible, for Gaussian copula parameter is between -1 and 1, Clayton copula has parameter α greater than -1, Frank copula has α not equal to zero, Gumbel copula has α greater than 1.

We will continue using the independence copulas in D-vine starting with tree 2, which we hope the end result would be an AR(1) structure as in Chapter 3. To get an equicorrelated stucture, as before, we assume the conditional correlations are fixed and equal to $\rho/(1+(k-1)\rho)$ for tree k.

IV.2.3 MULTINOMIAL VARIABLE WITH FOUR CATEGORIES

Although it is very similar to build the PMF of correlated multinomial variables with more categories, the bivariate multinomial variables with four categories is presented in this section, to show the pattern. Consider multinomial variables Y_i, $i = 1, 2$, with 4 possible outcomes: 1, 2, 3, 4 with probabilities as $p_{i,1}$, $p_{i,2}$, $p_{i,3}$ and $p_{i,4}$, with $p_{i,1} + p_{i,2} + p_{i,3} + p_{i,4} = 1$, i=1,2. The CDF of Y_i is

$$F(y_i) = \begin{cases} 0 & \text{if} \quad y_i < 0, \\ p_{i,1} & \text{if} \quad 0 \le y_i < 1, \\ p_{i,1} + p_{i,2} & \text{if} \quad 1 \le y_i < 2, \\ p_{i,1} + p_{i,2} + p_{i,3} & \text{if} \quad 2 \le y_i < 3, \\ 1 & \text{if} \quad 3 \le y_i. \end{cases}$$

The joint CDF of Y_1 and Y_2 using a copula C would be $F(Y_1 = y_1,\ Y_2 = y_2) = C(F(y_1),\ F(y_2))$. The relationship between copula parameter and correlation of multinomial variables with four categories can be obtained as in Section IV.2.2. Probability distribution of bivariate multinomial variables with 4 categories is listed in Table 37.

Table 37: PMF of bivariate multinomial variables with four categories

$(Y_1,\ Y_2)$	**Probability**
(1, 1)	$C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(1, 2)	$C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})-C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(1, 3)	$C(p_{1,1},;\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})-C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(1, 4)	$p_{1,1}-C(p_{1,1},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})$
(2, 1)	$C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})-C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(2, 2)	$C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12})-C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})-$ $C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})+C(p_{1,1},\ p_{2,1};\ \theta_{12})$
(2, 3)	$C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})-C(p_{1,1},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})$ $-C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12})+C(p_{1,1},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(2, 4)	$p_{1,2}-C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})+C(p_{1,1},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})$
(3, 1)	$C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1};\ \theta_{12})-C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})$
(3, 2)	$C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2};\ \theta_{12})-C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1};\ \theta_{12})$ $-C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12})+C(p_{1,1}+p_{1,2},\ p_{2,1};\ \theta_{12})$
(3, 3)	$C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})-$ $C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2};\ \theta_{12})$ $-C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})+C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(3, 4)	$p_{1,3}-C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})+$ $C(p_{1,1}+p_{1,2},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})$
(4, 1)	$p_{2,1}-C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1};\ \theta_{12})$
(4, 2)	$p_{2,2}-C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2};\ \theta_{12})+C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1};\ \theta_{12})$
(4, 3)	$p_{2,3}-C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})+$ $C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2};\ \theta_{12})$
(4, 4)	$1-p_{1,1}-p_{1,2}-p_{1,3}-p_{2,1}-p_{2,2}-p_{2,3}+$ $C(p_{1,1}+p_{1,2}+p_{1,3},\ p_{2,1}+p_{2,2}+p_{2,3};\ \theta_{12})$

NOTE: distribution has 7 parameters: $p_{1,1},\ p_{1,2},\ p_{1,3},\ p_{2,1},\ p_{2,2},\ p_{2,3}$, and copula parameter θ_{12}.

The joint distributions of trivariate or higher multinomial variables can be developed as in Section IV.2.1, and they are presented in Appendix C.

IV.3 NUMERICAL EXAMPLES

In this section, we present some numerical PMF's of dependent longitudinal multinomial variables, and compare them with two others obtained using algorithms: (i) the R package GenOrd which generates discrete variables using multivariate Gaussian copula developed by Barbiero and Ferrari (2015), (ii) the R package MultiOrd developed by Amatya and Demirtas (2015), by simulating the correlated ordinal discrete variables by collapsing the ordinal variables to the binary ones in the process. There are other methods to generate correlated multinomial variables, for instance, the method by Ibrahim and Suliadi (2011), but it uses Goodman-Kruskal's τ instead of Pearson's correlation coefficient, and therefore not comparable. However, we may include this algorithm when we develop regression methods in future research.

Let Y_1, Y_2 and Y_3 be three categorical variables taking values $1, 2, 3$ with parameters $p_1 = (0.2, 0.3, 0.5)$, $p_2 = (0.4, 0.25, 0.35)$ and $p_3 = (0.16, 0.48, 0.36)$ respectively. Our goal is to construct a joint distribution with using the methods described in Section IV.2.1. We are interested in AR(1) with parameter $\rho = 0.3$. First, we will construct joint distribution for $f(y_1,\, y_2)$ using the Gaussian copula and correlation $\rho = 0.3$. The parameter γ_{12} of the Gaussian copula is obtained solving Equation 4.2.31. In this case it reduces to

$$0.3 \;=\; \frac{E(Y_1, Y_2) - (3 - 2*0.2 - 0.3)*(3 - 2*0.4 - 0.25)}{\sqrt{(4*0.2 + 0.3 - (2*0.2 + 0.3)^2)(4*0.4 + 0.25 - (2*0.4 + 0.25)^2)}},$$

where,

$$\begin{aligned} E(Y_1, Y_2) \;=\;& \Phi_2(\Phi^{-1}(0.2), \Phi^{-1}(0.4); \gamma_{12}) + \Phi_2(\Phi^{-1}(0.2), \Phi^{-1}(0.65); \gamma_{12}) \\ & + \Phi_2(\Phi^{-1}(0.5), \Phi^{-1}(0.4); \gamma_{12}) \\ & + \Phi_2(\Phi^{-1}(0.5), \Phi^{-1}(0.65); \gamma_{12}) \\ & - 6p_{1,1} - 3p_{1,2} - 6p_{2,1} - 3p_{2,2} + 9, \end{aligned}$$

and solving this equation gives $\gamma_{12} = 0.389$. Similarly, we obtain $\gamma_{23} = 0.381$ that is needed for the construction of $f(y_2,\, y_3)$. Plugging in these values in Table 34, we get $f(y_1,\, y_2)$ and $f(y_2,\, y_3)$ as given in Table 38.

Table 38: PMF of bivariate multinomial variables

	Probability
$\boldsymbol{(Y_1,\ Y_2)}$	
(1, 1)	$C(p_{1,1}, p_{2,1}) = \Phi_2(\Phi^{-1}(0.2), \Phi^{-1}(0.4); 0.389) = 0.1243$
(1, 2)	$\Phi_2(\Phi^{-1}(0.2), \Phi^{-1}(0.4 + 0.25)) - 0.1243 = 0.1678 - 0.1243 = 0.0435$
(1, 3)	$0.2 - 0.1678 = 0.0322$
(2, 1)	$\Phi_2(\Phi^{-1}(0.2 + 0.3), \Phi^{-1}(0.4)) - 0.1243 = 0.2615 - 0.1243 = 0.1372$
(2, 2)	$\Phi_2(\Phi^{-1}(0.2 + 0.3), \Phi^{-1}(0.4 + 0.25)) - 0.2615 - 0.1678 + 0.1243$
	$= 0.3838 - 0.2615 - 0.1678 + 0.1243 = 0.0788$
(2, 3)	$0.3 - 0.3838 + 0.1678 = 0.084$
(3, 1)	$0.4 - 0.2615 = 0.1385$
(3, 2)	$0.25 - 0.3838 + 0.2615 = 0.1277$
(3, 3)	$1 - 0.2 - 0.3 - 0.4 - 0.25 + 0.3838 = 0.2338$
$(Y_2,\ Y_3)$	
(1, 1)	$C(p_{2,1}, p_{3,1}) = \Phi_2(\Phi^{-1}(0.4), \Phi^{-1}(0.16); 0.381) = 0.1016$
(1, 2)	$\Phi_2(\Phi^{-1}(0.4), \Phi^{-1}(0.16 + 0.48)) - 0.1016 = 0.3112 - 0.1016 = 0.2096$
(1, 3)	$0.4 - 0.3112 = 0.0888$
(2, 1)	$\Phi_2(\Phi^{-1}(0.4 + 0.25), \Phi^{-1}(0.16)) - 0.1016 = 0.1356 - 0.1016 = 0.0340$
(2, 2)	$\Phi_2(\Phi^{-1}(0.4 + 0.25), \Phi^{-1}(0.16 + 0.48)) - 0.1356 - 0.3112 + 0.1016$
	$= 0.4713 - 0.1356 - 0.3112 + 0.1016 = 0.1261$
(2, 3)	$0.25 - 0.4713 + 0.3112 = 0.0899$
(3, 1)	$0.16 - 0.1356 = 0.0244$
(3, 2)	$0.48 - 0.4713 + 0.1356 = 0.1443$
(3, 3)	$1 - 0.4 - 0.25 - 0.16 - 0.48 + 0.4713 = 0.1813$

Using Table 38, we can get the parameters of the conditional distributons. For example, $Y_1|Y_2 = 1$ has parameters

$$\begin{aligned}
p_{1|2,1|1} &= \frac{f(Y_1 = 1, Y_2 = 1)}{f(Y_2 = 1)} = \frac{0.1243}{0.4} = 0.3108 \\
p_{1|2,2|1} &= \frac{f(Y_1 = 2, Y_2 = 1)}{f(Y_2 = 1)} = \frac{0.1372}{0.4} = 0.343 \\
p_{1|2,3|1} &= \frac{f(Y_1 = 3, Y_2 = 1)}{f(Y_2 = 1)} = \frac{0.1385}{0.4} = 0.3463.
\end{aligned}$$

Similarly, we get the parameters of the other conditional distributions as $p_{1|2,.|2} = (0.1740, 0.3152, 0.5108)$, $p_{1|3,.|2} = (0.092, 0.240, 0.668)$, $p_{3|2,.|1} = (0.254, 0.524, 0.222)$, $p_{3|2,.|2} = (0.1360, 0.5044, 0.3596)$, $p_{3|2,.|3} = (0.0697, 0.4123, 0.518)$. Now using these values and the Gaussian copula we can get the joint PMF of (Y_1, Y_2, Y_3) as in Table 35. For example,

$$\begin{aligned}
f(1,1,1) &= f(Y_2 = 1)f(Y_1 = 1, Y_3 = 1|Y_2 = 1) \\
&= p_{2,1} * (p_{1|2,1|1} * p_{3|2,1|1}) \\
&= 0.4 * 0.3108 * 0.254 \\
&= 0.0316 \\
f(1,1,2) &= f(Y_2 = 1)f(Y_1 = 1, Y_3 = 2|Y_2 = 1) \\
&= p_{2,1} * (p_{1|2,1|1} * p_{3|2,2|1}) \\
&= 0.4 * 0.3108 * 0.524 \\
&= 0.0651
\end{aligned}$$

The result is listed in Table 39 along with the PMF created by the R-packages GenOrd and MultiOrd.

A check of the marginal probabilities and correlations is as follows. The PMF from D-vine pair-copula model has marginals: $p_1 = (0.2000, 0.3001, 0.5001)$, $p_2 = (0.4001, 0.2500, 0.3501)$, $p_3 = (0.1601, 0.4800, 0.3601)$, $\text{Corr}(Y_1, Y_2) = 0.300$, $\text{Corr}(Y_2, Y_3) = 0.300$ and $\text{Corr}(Y_1, Y_3) = 0.090$. The PMF from R package GenOrd has marginal parameters: $p_1 = (0.2019, 0.2984, 0.4995)$, $p_2 = (0.4007, 0.2469, 0.3522)$, $p_3 = (0.1592, 0.4789, 0.3617)$, $\text{Corr}(Y_1, Y_2) = 0.302$, $\text{Corr}(Y_2, Y_3) = 0.303$ and $\text{Corr}(Y_1, Y_3) = 0.089$; and the PMF from R package MultiOrd has marginal parameters: $p_1 = (0.2008, 0.3004, 0.4987)$, $p_2 = (0.4005, 0.2483, 0.3511)$, $p_3 = (0.1596, 0.4805, 0.3598)$, $\text{Corr}(Y_1, Y_2) = 0.300$, $\text{Corr}(Y_2, Y_3) = 0.297$ and $\text{Corr}(Y_1, Y_3) = 0.093$. All models generates the PMF of multinominal variables with an AR(1) correlation structure and required marginal parameters, but the D-vine pair-copula model is more accurate to the desired marginal parameters and correlations.

Following similar steps, PMF of equicorrelated multinomial variables can be generated, except that the conditional correlation is set to $\rho/(1+(k-1)\rho)$ at tree k. For example, the correlation of new variables $Y_1|Y_2=1$ and $Y_3|Y_2=1$ is assumed as $\rho/(1+\rho)=0.3/1.3=0.23$. We solve for $\gamma_{13|Y_2=1}$ from Equation (4.2.31), that is,

$$0.23 = \frac{E(Y_1|Y_2=1, Y_3|Y_2=1)-(3-2*0.3108-0.343)*(3-2*0.254-0.524)}{\sqrt{(4*0.3108+0.343-(2*0.3108+0.343)^2)(4*0.254+0.524-(2*0.254+0.524)^2)}},$$

where,

$$\begin{aligned} E(Y_1|Y_2, Y_3|Y_2) &= \Phi_2(\Phi^{-1}(0.3108), \Phi^{-1}(0.254); \gamma_{13|Y_2=1}) + \\ &\quad \Phi_2(\Phi^{-1}(0.3108), \Phi^{-1}(0.254+0.524); \gamma_{13|Y_2=1}) \\ &\quad +\Phi_2(\Phi^{-1}(0.3108+0.343), \Phi^{-1}(0.254); \gamma_{13|Y_2=1}) \\ &\quad +\Phi_2(\Phi^{-1}(0.3108+0.343), \Phi^{-1}(0.254+0.524); \gamma_{13|Y_2=1}) \\ &\quad -6p_{1|2,1|1}-3p_{1|2,2|1}-6p_{3|2,1|1}-3p_{3|2,2|1}+9. \end{aligned}$$

This gives us $\gamma_{13|Y_2=1}=0.287$. We can get $\gamma_{13|Y_2=2}$, $\gamma_{13|Y_2=3}$, similarly. The resulting PMF are listed in the Table 40. Checking back the parameters we have: PMF from D-vine pair-copula model has marginal parameters: $p_1=(0.2, 0.3001, 0.4999)$, $p_2=(0.4, 0.2499, 0.3501)$, $p_3=(0.16, 0.4799, 0.3601)$, $\text{Corr}(Y_1, Y_2)=0.3$, $\text{Corr}(Y_2, Y_3)=0.3$ and $\text{Corr}(Y_1, Y_3)=0.3$; PMF from R package GenOrd has marginal parameters: $p_1=(0.2006, 0.2985, 0.5011)$, $p_2=(0.4023, 0.2472, 0.3507)$, $p_3=(0.1593, 0.4785, 0.3624)$, $\text{Corr}(Y_1, Y_2)=0.3$, $\text{Corr}(Y_2, Y_3)=0.306$ and $\text{Corr}(Y_1, Y_3)=0.304$; and PMF from R package MultiOrd has marginal parameters: $p_1=(0.1993, 0.3013, 0.4995)$, $p_2=(0.3987, 0.2507, 0.3507)$, $p_3=(0.1605, 0.4807, 0.3589)$, $\text{Corr}(Y_1, Y_2)=0.302$, $\text{Corr}(Y_2, Y_3)=0.302$ and $\text{Corr}(Y_1, Y_3)=0.301$. All models generates PMF of multinominal variables with equicorrelated correlation structure and required success probabilities.

Table 39: PMF of AR(1) trivariate multinomial variables with three categories

$(Y_1,\ Y_2,\ Y_3)$	PMF_{D-Vine}	PMF_{GenOrd}	$PMF_{MultiOrd}$
(1, 1, 1)	0.0316	0.0334	0.0247
(1, 1, 2)	0.0651	0.0658	0.0749
(1, 1, 3)	0.0276	0.0273	0.0174
(1, 2, 1)	0.0059	0.0054	0.0044
(1, 2, 2)	0.0220	0.0210	0.0123
(1, 2, 3)	0.0156	0.0164	0.0180
(1, 3, 1)	0.0022	0.0024	0.0059
(1, 3, 2)	0.0133	0.0132	0.0180
(1, 3, 3)	0.0167	0.0170	0.0252
(2, 1, 1)	0.0349	0.0338	0.0378
(2, 1, 2)	0.0719	0.0725	0.1123
(2, 1, 3)	0.0315	0.0304	0.0260
(2, 2, 1)	0.0107	0.0099	0.0060
(2, 2, 2)	0.0398	0.0395	0.0183
(2, 2, 3)	0.0283	0.0281	0.0267
(2, 3, 1)	0.0059	0.0057	0.0088
(2, 3, 2)	0.0346	0.0348	0.0266
(2, 3, 3)	0.0435	0.0437	0.0379
(3, 1, 1)	0.0352	0.0337	0.0243
(3, 1, 2)	0.0725	0.0739	0.0721
(3, 1, 3)	0.0308	0.0299	0.0110
(3, 2, 1)	0.0174	0.0183	0.0196
(3, 2, 2)	0.0644	0.0636	0.0617
(3, 2, 3)	0.0459	0.0447	0.0813
(3, 3, 1)	0.0163	0.0166	0.0281
(3, 3, 2)	0.0942	0.0946	0.0843
(3, 3, 3)	0.1212	0.1242	0.1163

Table 40: PMF of equicorrelated trivariate multinomial variables with three categories

$(Y_1,\ Y_2,\ Y_3)$	PMF_{D-Vine}	PMF_{GenOrd}	$PMF_{MultiOrd}$
(1, 1, 1)	0.0453	0.0467	0.0266
(1, 1, 2)	0.0628	0.0619	0.0800
(1, 1, 3)	0.0162	0.0154	0.0095
(1, 2, 1)	0.0107	0.0100	0.0062
(1, 2, 2)	0.0239	0.0248	0.0181
(1, 2, 3)	0.0089	0.0091	0.0102
(1, 3, 1)	0.0055	0.0052	0.0083
(1, 3, 2)	0.0173	0.0173	0.0251
(1, 3, 3)	0.0094	0.0102	0.0153
(2, 1, 1)	0.0343	0.0338	0.0400
(2, 1, 2)	0.0743	0.0749	0.1209
(2, 1, 3)	0.0287	0.0292	0.0143
(2, 2, 1)	0.0124	0.0119	0.0096
(2, 2, 2)	0.0425	0.0416	0.0275
(2, 2, 3)	0.0239	0.0239	0.01540
(2, 3, 1)	0.0085	0.0086	0.01300
(2, 3, 2)	0.0411	0.0398	0.0383
(2, 3, 3)	0.0344	0.0348	0.0223
(3, 1, 1)	0.0220	0.0213	0.0198
(3, 1, 2)	0.0724	0.0740	0.0596
(3, 1, 3)	0.0440	0.0451	0.02800
(3, 2, 1)	0.0109	0.0111	0.0148
(3, 2, 2)	0.0597	0.0591	0.0465
(3, 2, 3)	0.0570	0.0557	0.1024
(3, 3, 1)	0.0104	0.0107	0.0222
(3, 3, 2)	0.0859	0.0851	0.0647
(3, 3, 3)	0.1376	0.1390	0.1415

NOTE: Equicorrelated structure coefficient is $\rho = 0.3$, success probabilities are $p_1 = (0.2, 0.3, 0.5)$, $p_2 = (0.4, 0.25, 0.35)$ and $p_3 = (0.16, 0.48, 0.36)$. Consequently, correlation coefficient for D-vine pair-copula model using Gaussian copula are $\gamma_{12} = 0.389$, $\gamma_{23} = 0.381$, $\gamma_{13|y_2=1} = 0.287$, $\gamma_{13|y_2=2} = 0.296$, and $\gamma_{13|y_2=3} = 0.318$.

Table 41: PMF of trivariate multinomial variables with three categories when only D-vine pair-copula model works

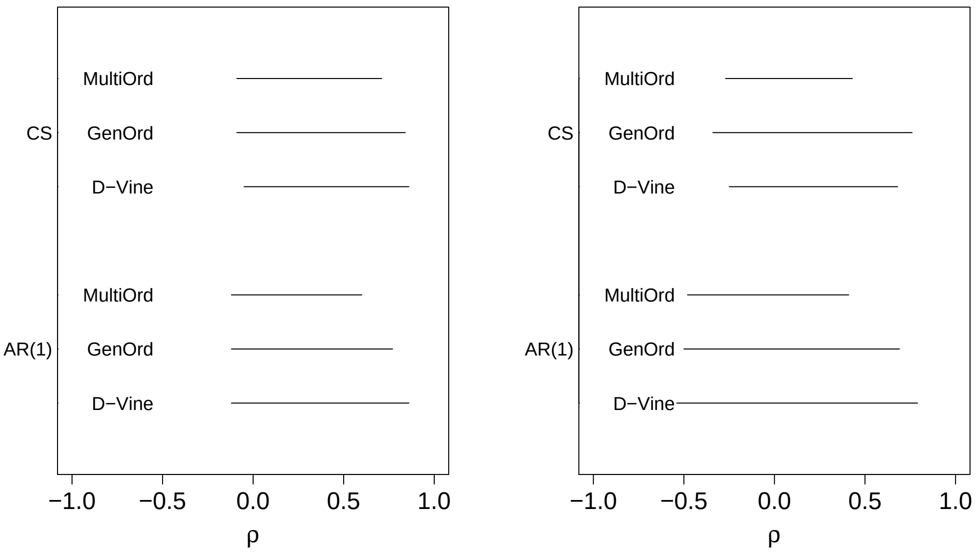
$(Y_1,\ Y_2,\ Y_3)$	**AR(1)** $\rho = -0.851$	**AR(1)** $\rho = 0.83$	**equicorrelated** $\rho = 0.78$
(1, 1, 1)	0	0.0800	0.1198
(1, 1, 2)	0	0.1200	0.0767
(1, 1, 3)	0	0	0
(1, 2, 1)	0	0	0.0005
(1, 2, 2)	0.0009	0	0.0028
(1, 2, 3)	0	0	0
(1, 3, 1)	0.0910	0	0
(1, 3, 2)	0.1081	0	0
(1, 3, 3)	0	0	0
(2, 1, 1)	0	0.0785	0.0386
(2, 1, 2)	0.0007	0.1178	0.1409
(2, 1, 3)	0.0065	0	0.0011
(2, 2, 1)	0	0	0.0004
(2, 2, 2)	0.1439	0.0894	0.1046
(2, 2, 3)	0.0001	0.0135	0.0024
(2, 3, 1)	0.0680	0	0
(2, 3, 2)	0.0807	0	0.0114
(2, 3, 3)	0	0.0007	0.0007
(3, 1, 1)	0	0.0015	0.0007
(3, 1, 2)	0.0395	0.0022	0.0205
(3, 1, 3)	0.3532	0	0.0016
(3, 2, 1)	0	0	0
(3, 2, 2)	0.1049	0.1277	0.0853
(3, 2, 3)	0.0001	0.0193	0.0540
(3, 3, 1)	0.0010	0	0
(3, 3, 2)	0.0012	0.0228	0.0377
(3, 3, 3)	0	0.3265	0.3002

NOTE: With the same success probabilities as previous table, AR(1) structure with parameter -0.85 has correlations for Gaussian copula as $\gamma_{12} = -0.9709$, and $\gamma_{23} = -0.9984$; with parameter 0.83 has correlations for Gaussian copula as $\gamma_{12} = 0.9804$, and $\gamma_{23} = 0.9829$; equicorrelated structure with parameter 0.78 has correlations for Gaussian copula as $\gamma_{12} = 0.9283$, $\gamma_{23} = 0.9345$, $\gamma_{13|y_2=1} = 0.6279$, $\gamma_{13|y_2=2} = 0.7916$, and $\gamma_{13|y_2=3} = 0.8814$.

Some general comments are in order. From Tables 39 and 40, we notice that for both AR(1) and equicorrelated correlation structures, the PMF values created by D-Vine and GenOrd seem closer than the ones from MultiOrd. For instance, for $y = (1, 1, 1)$, PMF with AR(1) correlation structure from D-Vine is 0.0334, which is close to 0.0334, from GenOrd, while MultiOrd has value 0.0247; PMF with equicorrelated correlation structure from D-Vine is 0.0453, which is close to 0.0467 from GenOrd, while MultiOrd assigns probability 0.0266. Also, MultiOrd has a narrower feasible range of the correlation coefficient than the other two, for example, it doesn't work for $\rho = 0.5$, but both D-Vine and GenOrd generate a distribution. This is because MultiOrd has to generate binary data in intermediate steps, and some specified correlations cannot be generated due to the feasible boundaries range of correlation for binary variables.

Furthermore, our D-vine pair-copula module works even when the package GenOrd fails. Table 41 shows the PMF generated by D-vine pair-copula module, while the other two packages fail, although these ρ values are within the feasible ranges. The correlations are checked directly with the generated PMF as well: for the AR(1) with $\rho = -0.851$ and $\rho^2 = 0.724$ case, $\text{Corr}(Y_1, Y_2) = -0.853$, $\text{Corr}(Y_2, Y_3) = -0.854$, $\text{Corr}(Y_1, Y_3) = 0.705$ shows an approximate AR(1) struture; for AR(1) with $\rho = 0.83$ and $\rho^2 = 0.689$ case, $\text{Corr}(Y_1, Y_2) = 0.829$, $\text{Corr}(Y_2, Y_3) = 0.829$, $\text{Corr}(Y_1, Y_3) = 0.672$ shows an approximate AR(1) struture; for equicorrelated structure with $\rho = 0.78$ case, $\text{Corr}(Y_1, Y_2) = 0.778$, $\text{Corr}(Y_2, Y_3) = 0.779$, $\text{Corr}(Y_1, Y_3) = 0.766$ shows an approximate equicorrelated struture.

However, the GenOrd package is useful for negative correlation values. With the same marginal parameters as above with equicorrelated structure and $\rho = -0.36$, GenOrd package could generate a joint distribution whereas Multiord package and D-vine pair-copula with Gaussian copula fail. In order to compare the ability of generating joint PMF, we checked the feasible range of ρ for different success probabilities, and put two examples in Figure 12. When correlation structure is AR(1), most of the time, Genord and D-vine pair-copula module has the same lower bound, but D-Vine has a larger upper bound; When correlation structure is equicorrelated, most of the time, GenOrd works better around the lower bound, but around the upper bound, sometimes D-Vine works while GenOrd fails, sometimes D-Vine fails while GenOrd works. To sum up, D-Vine works better for AR(1) struture, GenOrd works better for equicorrelated structure, and MultiOrd always has the narrowest feasible range.



(a) (b)

Figure 12: Feasible range of ρ for D-Vine pair-copula model, GenOrd and MultiOrd with AR(1) or equicorrelated correlation structure, (a)p_1=(0.05, 0.05, 0.9), p_2=(0.1, 0.05, 0.85), p_3=(0.05, 0.05, 0.9); (b)p_1=(0.2, 0.3, 0.5), p_2=(0.1, 0.2, 0.7), p_3=(0.3, 0.3, 0.4)

IV.4 SIMULATION

To compare the performance of the D-vine pair-copula model with (i) R package GenOrd by Barbiero and Ferrari (2015), and (ii) R package MultiOrd by Amatya and Demirtas (2015), we have conducted simulations to compare the performance of estimating the correlation coefficients and success probabilities.

For simulations we took $p_1 = (0.25, 0.15, 0.6)$, $p_2 = (0.4, 0.5, 0.1)$, $p_3 = (0.34, 0.26, 0.4)$, and the correlation $\rho = 0.45$, the structure is either AR(1) or equicorrelated. Known the true parameters as above, we generated a sample with size $n = 100$ using D-vine model and estimated the parameters using the method of moments. Then we repeated the procedure 500 times, thus 500 estimates of ρ's and p's were obtained to calculate the mean and standard deviation to form the 95% confidence interval, and then the coverage of standard errors (CSE), which means the proportion of replicates where true parameter was contained in the 95% confidence interval. Then we generated samples using GenOrd and MultiOrd, respectively, and repeated the procedure as above. The CSE results are listed in Tables 42.

Table 42: CSE from simulated three-dimensional correlated multinomial variables

	AR(1)			**Equicorrelated**		
	D-vine	GenOrd	MultiOrd	D-vine	GenOrd	MultiOrd
ρ_{12}	0.958	0.952	0.954	0.958	0.944	0.952
ρ_{23}	0.956	0.952	0.946	0.956	0.950	0.954
ρ_{13}	0.950	0.948	0.942	0.956	0.950	0.956
$p_{1,1}$	0.952	0.960	0.942	0.952	0.954	0.956
$p_{1,2}$	0.948	0.932	0.944	0.950	0.946	0.944
$p_{1,3}$	0.950	0.942	0.944	0.950	0.946	0.948
$p_{2,1}$	0.952	0.956	0.962	0.956	0.942	0.950
$p_{2,2}$	0.942	0.954	0.952	0.952	0.954	0.944
$p_{2,3}$	0.956	0.938	0.954	0.964	0.958	0.934
$p_{3,1}$	0.938	0.946	0.942	0.946	0.940	0.950
$p_{3,2}$	0.966	0.938	0.942	0.944	0.960	0.948
$p_{3,3}$	0.948	0.952	0.944	0.960	0.954	0.954

NOTE: With the success probabilities $p_1 = c(0.25, 0.15, 0.6)$, $p_2 = c(0.4, 0.5, 0.1)$, $p_3 = c(0.34, 0.26, 0.4)$, with AR(1) structure, feasible range of ρ is (-0.72, 0.75) for D-vine pair-copula, (-0.71,0.72) for GenOrd, (-0.34, 0.47) for MultiOrd; with equicorrelated structure, feasible range of ρ is (-0.3, 0.6) for D-vine pair-copula, (-0.37, 0.81) for GenOrd, (-0.27, 0.55) for MultiOrd.

Looking at Tables 42, although the differences are very small among all three models, D-vine pair-copula model performs the best for generating the dependence among multinomial variables, while all three models appear pretty same effective for generating the marginal parameters. For example, the CSE of ρ_{12} in Table 42 is 0.958 for D-vine pair-copula model, which is larger than 0.952 from GenOrd package, or 0.954 from MultiOrd package.

CHAPTER V

CONCLUSIONS AND FUTURE WORK

Longitudinal discrete data occur frequently in scientific disciplines. The most widely used methodologies for analysis of such data are the generalized estimating equations (GEE) approach (see Liang and Zeger (1986)), Gaussian method of estimation (see (Crowder, 2001)), the quasi-least squares method of estimation for the correlation (see (Chaganty, 1997)), the first-order Markov chains model and the multivariate probit model (see (Chaganty and Joe, 2004), (Yang and Chaganty, 2014).)

In this dissertation, we presented a D-vine pair-copula model for longitudinal discrete binary data, and extended it for longitudinal multinomial data. In Chapter II, using the D-vine pair-copula model, we have shown the procedure of generating the PMF for the multivariate binary variables with AR(1) or equicorrelated structure. We have also shown that MP model is different from vine Gaussian pair-copula starting three dimensions. The main advantage of the D-vine pair-copula model is that it can produce the PMF around correlation feasible boundaries where MP model fails. In Chapter III, we conducted the comparison between D-vine pair-copula regression model and MP model based on actual likelihoods. We shown that the D-vine Gaussian copula works more properly to estimate the correlation coefficients close to the correlation feasible boundaries for both small or large samples. In Chapter IV, we show that D-Vine pair-copula Gaussian model generated a proper multidimensional distribution, and it worked even when correlation parameter is around the correlation feasible boundaries, whereas the two R package GenOrd (developed by Barbiero and Ferrari (2015)) and MultiOrd (developed by Amatya and Demirtas (2015)) fail to generate a probability distribution. We performed some simulations to contrast our method with other competing established methods. Simulation results indicated that although the differences were very small among all three models, D-vine pair-copula models was the most efficient for estimating the correlation coefficient, while all three models appeared pretty close for estimating the marginal parameters.

As a potential future research, we will extend the D-vine pair-copula model to the regression setting assuming the data consists of covariates associated with the multinomial responses.

REFERENCES

Amatya, A. and Demirtas, H. (2015), "MultiOrd: An R package for generating correlated ordinal data," *Communications in Statistics-Simulation and Computation*, 44, 1683–1691.

Barbiero, A. and Ferrari, P. A. (2015), "GenOrd: simulation of discrete random variables with given correlation matrix and marginal distributions," *R package version*, 1.

Bedford, T. and Cooke, R. M. (2002), "A new graphical model for dependent random variables," *Annals of Statistics*, 30, 1031–1068.

Brechmann, E. and Schepsmeier, U. (2013), "Cdvine: Modeling dependence with c-and d-vine copulas in r," *Journal of statistical software*, 52, 1–27.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995), "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, 16, 1190–1208.

Chaganty, N. R. (1997), "An alternative approach to the analysis of longitudinal data via generalized estimating equations," *Journal of Statistical Planning and Inference*, 63, 39–54.

Chaganty, N. R. and Joe, H. (2004), "Efficiency of generalized estimating equations for binary responses," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66, 851–860.

— (2006), "Range of correlation matrices for dependent Bernoulli random variables," *Biometrika*, 93, 197–206.

Crowder, M. (2001), "On repeated measures analysis with misspecified covariance structure," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 55–62.

Czado, C. (2019), *Analyzing Dependent Data with Vine Copulas*, Springer Nature.

Emrich, L. J. and Piedmonte, M. R. (1991), "A method for generating high-dimensional multivariate binary variates," *The American Statistician*, 45, 302–304.

Escarela, G., Perez-Ruiz, L. C., and Bowater, R. J. (2009), “A copula-based Markov chain model for the analysis of binary longitudinal data,” *Journal of Applied Statistics*, 36, 647–657.

Frees, E. W. and Wang, P. (2006), “Copula credibility for aggregate loss models,” *Insurance: Mathematics and Economics*, 38, 360–373.

Gilbert, P. and Varadhan, R. (2012), “numDeriv: accurate numerical derivatives,” *R package version*, 1.

Grizzle, J. E., Starmer, C. F., and Koch, G. G. (1969), “Analysis of categorical data by linear models,” *Biometrics*, 489–504.

Ibrahim, N. A. and Suliadi, S. (2011), “Generating correlated discrete ordinal data using R and SAS IML,” *Computer methods and programs in biomedicine*, 104, e122–e132.

Joe, H. (2014), *Dependence modeling with copulas*, Chapman and Hall/CRC.

Lennon, H. (2016), “Gaussian copula modelling for integer-valued time series,” Ph.D. thesis, The University of Manchester (United Kingdom).

Liang, K.-Y. and Zeger, S. L. (1986), “Longitudinal data analysis using generalized linear models,” *Biometrika*, 73, 13–22.

Mobily, P. R., Herr, K. A., Clark, M. K., and Wallace, R. B. (1994), “An epidemiologic analysis of pain in the elderly: The Iowa 65+ rural health study,” *Journal of Aging and Health*, 6, 139–154.

Nagler, T. and Vatter, T. (2018), “rvinecopulib: High performance algorithms for vine copula modeling,” *R package version*, 2, 0.

Panagiotelis, A., Czado, C., and Joe, H. (2012), “Pair copula constructions for multivariate discrete data,” *Journal of the American Statistical Association*, 107, 1063–1072.

Panagiotelis, A., Czado, C., Joe, H., and Stöber, J. (2017), “Model selection for discrete regular vine copulas,” *Computational Statistics & Data Analysis*, 106, 138–152.

Plackett, R. L. (1954), “A reduction formula for normal multivariate integrals,” *Biometrika*, 41, 351–360.

Poddar, A. (2016), “Analysis off Dependent Discrete Choices Using Gaussian Copula,” Ph.D. thesis, Old Dominion University.

Radice, R., Marra, G., and Wojtyś, M. (2016), “Copula regression spline models for binary outcomes,” *Statistics and Computing*, 26, 981–995.

Sabo, R. T. and Chaganty, N. R. (2010), “What can go wrong when ignoring correlation bounds in the use of generalized estimating equations,” *Statistics in Medicine*, 29, 2501–2507.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., et al. (2015), “Package ‘VineCopula’,” *R package version*, 2.

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

Smith, M., Min, A., Almeida, C., and Czado, C. (2010), “Modeling longitudinal data using a pair-copula decomposition of serial dependence,” *Journal of the American Statistical Association*, 105, 1467–1479.

Smith, M. S. and Khaled, M. A. (2012), “Estimation of copula models with discrete margins via Bayesian data augmentation,” *Journal of the American Statistical Association*, 107, 290–303.

Stock, J. R., Weaver, J., Ray, H., Brink, J., Sadof, M. G., et al. (1983), “Evaluation of safe performance secondary school driver education curriculum demonstration project,” Tech. rep., United States. National Highway Traffic Safety Administration.

Stokes, M. E., Davis, C. S., and Koch, G. G. (1995), “Categorical data analysis using the SAS system. SAS Institute,” *Inc., Cary, NC*, 34–35.

Touloumis, A. (2016), “Simulating correlated binary and multinomial responses with simcormultres,” *The Comprehensive R Archive Network*, 1–5.

Viskov, O. (2008), "On the Mehler formula for Hermite polynomials," in *Doklady Mathematics*, Springer, vol. 77, pp. 1–4.

Winkelmann, R. (2012), "Copula bivariate probit models: with an application to medical expenditures," *Health economics*, 21, 1444–1455.

Woolson, R. F. and Clarke, W. R. (1984), "Analysis of categorical incomplete longitudinal data," *Journal of the Royal Statistical Society: Series A (General)*, 147, 87–99.

Xiao, Q. and Zhou, S. (2019), "Matching a correlation coefficient by a Gaussian copula," *Communications in Statistics-Theory and Methods*, 48, 1728–1747.

Xu, J. J. (1996), "Statistical modelling and inference for multivariate and longitudinal discrete response data," Ph.D. thesis, University of British Columbia.

Yang, W. and Chaganty, N. R. (2014), "A contrasting study of likelihood methods for the analysis of longitudinal binary data," *Communications in Statistics-Theory and Methods*, 43, 3027–3046.

APPENDIX A

EXAMPLES OF RELATIONSHIP BETWEEN BINARY CORRELATION AND COPULA PARAMETER

Table 43 presents the parameter of Clayton copula, Table 44 has the parameter of Frank copula, and Table 45 shows the parameter of Gumbel copula. NA in the table means given correlation of binary variables is not in the feasible range with the marginal proportions. For example, the first cell in Table 43 is NA, because $p = (0.1, 0.1)$ gives feasible range of both AR(1) and exchangeable structure coefficient as $(-0.111, 1)$, which doesn't cover $\rho = -0.8$. In the two dimension situation, AR(1) and exchangeable structure have the same boundaries.

Table 43: Copula parameter α for Clayton copula

		p_2								
ρ	p_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
-0.8	0.1	NA	NA	NA	NA	NA	NA	NA	NA	-0.9311
	0.2	NA	NA	NA	NA	NA	NA	NA	-0.9194	NA
	0.3	NA	NA	NA	NA	NA	-0.9990	-0.8608	NA	NA
	0.4	NA	NA	NA	NA	-0.9903	-0.9091	-0.8578	NA	NA
	0.5	NA	NA	NA	-0.9903	-0.9080	-0.8560	NA	NA	NA
	0.6	NA	NA	-0.860	-0.9091	-0.8560	NA	NA	NA	NA
	0.7	NA	NA	-0.9127	-0.8578	NA	NA	NA	NA	NA
	0.8	NA	-0.9194	NA	NA	NA	NA	NA	NA	NA
	0.9	-0.9311	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	0.1	NA	NA	NA	NA	NA	NA	NA	-0.9546	-0.8337
	0.2	NA	NA	NA	NA	NA	-0.9890	-0.8886	-0.8048	-0.7436
	0.3	NA	NA	NA	NA	-0.9523	-0.8608	-0.7885	-0.7345	NA
	0.4	NA	NA	NA	-0.9420	-0.8493	-0.7799	-0.7291	-0.7111	NA
	0.5	NA	NA	-0.9523	-0.8493	-0.7771	-0.7263	-0.6993	NA	NA
	0.6	NA	-0.9890	-0.8608	-0.7799	-0.7263	-0.6970	NA	NA	NA
	0.7	NA	-0.8886	-0.7885	-0.7291	-0.6993	NA	NA	NA	NA
	0.8	-0.9546	-0.8048	-0.7885	-0.7111	NA	NA	NA	NA	NA
	0.9	-0.8337	-0.7436	NA	NA	NA	NA	NA	NA	NA
-0.2	0.1	NA	NA	-0.9329	-0.7915	-0.6928	-0.6183	-0.5592	-0.5107	-0.4714
	0.2	NA	-0.8346	-0.6803	-0.5854	-0.5194	-0.4702	-0.4323	-0.4034	-0.3856
	0.3	-0.9329	-0.6803	-0.5606	-0.4869	-0.4359	-0.3984	-0.3703	-0.3506	-0.3440
	0.4	-0.7915	-0.5854	-0.4869	-0.4262	-0.3844	-0.3540	-0.3321	-0.3185	-0.3204
	0.5	-0.6928	-0.5194	-0.4359	-0.3844	-0.3491	-0.3239	-0.3067	-0.2980	-0.3082
	0.6	-0.6183	-0.4702	-0.3984	-0.3540	-0.3239	-0.3031	-0.2898	-0.2860	-0.3075
	0.7	-0.5592	-0.4323	-0.3703	-0.3321	-0.3067	-0.2898	-0.2808	-0.2832	-0.3312

Continued										
	0.8	-0.5107	-0.4034	-0.3506	-0.3185	-0.2980	-0.2860	-0.2832	-0.2990	NA
	0.9	-0.4714	-0.3856	-0.3440	-0.3204	-0.3082	-0.3075	-0.3312	NA	NA
0.1	0.1	1.1109	0.7946	0.6631	0.5916	0.5510	0.5330	0.5407	0.6013	1.0423
	0.2	0.7946	0.5570	0.4581	0.4029	0.3691	0.3495	0.3430	0.3568	0.4411
	0.3	0.6631	0.4581	0.3734	0.3259	0.2962	0.2778	0.2690	0.2735	0.3159
	0.4	0.5916	0.4029	0.3259	0.2827	0.2554	0.2380	0.2285	0.2291	0.2557
	0.5	0.5510	0.3691	0.2962	0.2554	0.2296	0.2127	0.2029	0.2013	0.2193
	0.6	0.5330	0.3495	0.2778	0.2380	0.2127	0.1960	0.1857	0.1825	0.1950
	0.7	0.5407	0.3430	0.2690	0.2285	0.2029	0.1857	0.1747	0.1700	0.1783
	0.8	0.6013	0.3568	0.2735	0.2291	0.2013	0.1825	0.1700	0.1635	0.1678
	0.9	1.0423	0.4411	0.3159	0.2557	0.2193	0.1950	0.1783	0.1678	0.1664
0.3	0.1	4.2397	3.1454	2.8436	2.9049	3.6497	NA	NA	NA	NA
	0.2	3.1454	2.1219	1.7708	1.6190	1.5866	1.7013	2.3579	NA	NA
	0.3	2.8436	1.7708	1.4185	1.2486	1.1666	1.1541	1.2471	1.8433	NA
	0.4	2.9049	1.6190	1.2486	1.0693	0.9717	0.9271	0.9394	1.0840	NA
	0.5	3.6497	1.5866	1.1666	0.9717	0.8626	0.8021	0.7835	0.8337	1.3316
	0.6	NA	1.7013	1.1541	0.9271	0.8021	0.7282	0.6911	0.7000	0.8758
	0.7	NA	2.3579	1.2471	0.9394	0.7835	0.6911	0.6369	0.6193	0.6917
	0.8	NA	NA	1.8433	1.0840	0.8337	0.7000	0.6193	0.5755	0.5891
	0.9	NA	NA	NA	NA	1.3316	0.8758	0.6917	0.5891	0.5409
0.5	0.1	9.6011	7.9355	13.6175	NA	NA	NA	NA	NA	NA
	0.2	7.9355	4.7915	4.1623	4.3653	NA	NA	NA	NA	NA
	0.3	13.6175	4.1623	3.1905	2.8783	2.9425	3.9933	NA	NA	NA
	0.4	NA	4.3653	2.8783	2.3920	2.2151	2.2878	3.2030	NA	NA
	0.5	NA	NA	2.9425	2.2151	1.9150	1.8113	1.9327	NA	NA
	0.6	NA	NA	3.9933	2.2878	1.8113	1.5993	1.5461	1.7935	NA
	0.7	NA	NA	NA	3.2030	1.9327	1.5461	1.3765	1.3755	2.9472
	0.8	NA	NA	NA	NA	NA	1.7935	1.3755	1.2127	1.3416
	0.9	NA	NA	NA	NA	NA	NA	2.9472	1.3416	1.0900
0.7	0.1	20.8866	NA	NA	NA	NA	NA	NA	NA	NA
	0.2	NA	10.3860	11.1582	NA	NA	NA	NA	NA	NA
	0.3	NA	11.1582	6.8846	6.9955	NA	NA	NA	NA	NA
	0.4	NA	NA	6.9955	5.1328	5.2218	NA	NA	NA	NA
	0.5	NA	NA	NA	5.2218	4.0804	4.2345	NA	NA	NA
	0.6	NA	NA	NA	NA	4.2345	3.3774	3.6618	NA	NA
	0.7	NA	NA	NA	NA	NA	3.6618	2.8733	3.5224	NA
	0.8	NA	NA	NA	NA	NA	NA	3.5224	2.4926	NA
	0.9	NA	NA	NA	NA	NA	NA	NA	NA	2.1923

Table 44: Copula parameter α for Frank copula

		p_2								
ρ	p_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
-0.8	0.1	NA	NA	NA	NA	NA	NA	NA	NA	-37.88
	0.2	NA	NA	NA	NA	NA	NA	NA	-21.45	NA
	0.3	NA	NA	NA	NA	NA	-41.01	-16.42	NA	NA
	0.4	NA	NA	NA	NA	-23.20	-14.41	-41.01	NA	NA
	0.5	NA	NA	NA	-23.20	-13.84	-23.20	NA	NA	NA
	0.6	NA	NA	-41.01	-14.41	-23.20	NA	NA	NA	NA
	0.7	NA	NA	-16.42	-41.01	NA	NA	NA	NA	NA
	0.8	NA	-21.45	NA	NA	NA	NA	NA	NA	NA
	0.9	-37.88	NA	NA	NA	NA	NA	NA	NA	NA
-0.6	0.1	NA	NA	NA	NA	NA	NA	NA	-17.18	-16.43
	0.2	NA	NA	NA	NA	NA	-15.95	-9.48	-9.65	-17.18
	0.3	NA	NA	NA	NA	-9.85	-7.65	-7.60	-9.48	NA
	0.4	NA	NA	NA	-9.03	-7.03	-6.79	-7.65	-15.95	NA
	0.5	NA	NA	-9.85	-7.03	-6.56	-7.03	-9.85	NA	NA
	0.6	NA	-15.95	-7.65	-6.79	-7.03	-9.03	NA	NA	NA
	0.7	NA	-9.48	-7.60	-7.65	-9.85	NA	NA	NA	NA
	0.8	-17.18	-9.65	-9.48	-15.95	NA	NA	NA	NA	NA
	0.9	-16.43	-17.18	NA	NA	NA	NA	NA	NA	NA
-0.2	0.1	NA	NA	-5.34	-3.62	-3.09	-2.87	-2.84	-3.03	-3.72
	0.2	NA	-3.73	-2.68	-2.30	-2.13	-2.08	-2.14	-2.36	-3.03
	0.3	-5.34	-2.68	-2.14	-1.91	-1.81	-1.81	-1.90	-2.14	-2.84
	0.4	-3.62	-2.30	-1.91	-1.74	-1.68	-1.70	-1.81	-2.08	-2.87
	0.5	-3.09	-2.13	-1.81	-1.68	-1.64	-1.68	-1.81	-2.13	-3.09
	0.6	-2.87	-2.08	-1.81	-1.70	-1.68	-1.74	-1.91	-2.30	-3.62
	0.7	-2.84	-2.14	-1.90	-1.81	-1.81	-1.91	-2.14	-2.68	-5.34
	0.8	-3.03	-2.36	-2.14	-2.08	-2.13	-2.30	-2.68	-3.73	NA
	0.9	-3.72	-3.03	-2.84	-2.87	-3.09	-3.62	-5.34	NA	NA
0.1	0.1	1.93	1.53	1.40	1.35	1.38	1.47	1.66	2.11	4.18
	0.2	1.53	1.19	1.06	1.01	1.01	1.06	1.17	1.40	2.11
	0.3	1.40	1.06	0.94	0.89	0.88	0.91	0.99	1.17	1.66
	0.4	1.35	1.01	0.89	0.84	0.82	0.84	0.91	1.06	1.47
	0.5	1.38	1.01	0.88	0.82	0.81	0.82	0.88	1.01	1.38
	0.6	1.47	1.06	0.91	0.84	0.82	0.84	0.89	1.01	1.35
	0.7	1.66	1.17	0.99	0.91	0.88	0.89	0.94	1.06	1.40
	0.8	2.11	1.40	1.17	1.06	1.01	1.01	1.06	1.19	1.53
	0.9	4.18	2.11	1.66	1.47	1.38	1.35	1.40	1.53	1.93

Continued

0.3	0.1	5.74	4.73	4.63	5.04	6.58	NA	NA	NA	NA
	0.2	4.73	3.62	3.32	3.29	3.50	4.05	5.91	NA	NA
	0.3	4.63	3.32	2.93	2.82	2.87	3.11	3.70	5.91	NA
	0.4	5.04	3.29	2.82	2.64	2.63	2.76	3.11	4.05	NA
	0.5	6.58	3.50	2.87	2.63	2.56	2.63	2.87	3.50	6.58
	0.6	NA	4.05	3.11	2.76	2.63	2.64	2.82	3.29	5.04
	0.7	NA	5.91	3.70	3.11	2.87	2.82	2.93	3.32	4.63
	0.8	NA	NA	5.91	4.05	3.50	3.29	3.32	3.62	4.73
	0.9	NA	NA	NA	NA	6.58	5.04	4.63	4.73	5.74
0.5	0.1	11.62	10.32	17.97	NA	NA	NA	NA	NA	NA
	0.2	10.32	7.01	6.60	7.34	49.94	NA	NA	NA	
	0.3	17.97	6.60	5.60	5.49	6.04	8.56	NA	NA	NA
	0.4	NA	7.34	5.49	5.03	5.10	5.73	8.56	NA	NA
	0.5	NA	49.94	6.04	5.10	4.88	5.10	6.04	NA	NA
	0.6	NA	NA	8.56	5.73	5.10	5.03	5.49	7.34	NA
	0.7	NA	NA	NA	8.56	6.04	5.49	5.60	6.60	17.97
	0.8	NA	NA	NA	NA	100.00	7.34	6.60	7.01	10.32
	0.9	NA	NA	NA	NA	NA	NA	17.97	10.32	11.62
0.7	0.1	23.94	NA	NA	NA	NA	NA	NA	NA	NA
	0.2	NA	13.77	15.63	NA	NA	NA	NA	NA	NA
	0.3	NA	15.63	10.67	11.60	NA	NA	NA	NA	NA
	0.4	NA	NA	11.60	9.44	10.41	NA	NA	NA	NA
	0.5	NA	NA	NA	10.41	9.10	10.41	NA	NA	NA
	0.6	NA	NA	NA	NA	10.41	9.44	11.60	NA	NA
	0.7	NA	NA	NA	NA	NA	11.60	10.67	15.63	NA
	0.8	NA	NA	NA	NA	NA	NA	15.63	13.77	NA
	0.9	NA	NA	NA	NA	NA	NA	NA	NA	23.94

Table 45: Copula parameter α for Gumbel copula

		p_2								
ρ	p_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.1	1.08	1.09	1.10	1.11	1.13	1.15	1.19	1.26	1.61
	0.2	1.09	1.09	1.09	1.10	1.11	1.13	1.15	1.19	1.31
	0.3	1.10	1.09	1.10	1.10	1.11	1.12	1.14	1.17	1.25
	0.4	1.11	1.10	1.10	1.10	1.11	1.12	1.13	1.16	1.23
	0.5	1.13	1.11	1.11	1.11	1.11	1.12	1.14	1.16	1.22
	0.6	1.15	1.13	1.12	1.12	1.12	1.13	1.14	1.16	1.21
	0.7	1.19	1.15	1.14	1.13	1.14	1.14	1.15	1.17	1.22
	0.8	1.26	1.19	1.17	1.16	1.16	1.16	1.17	1.19	1.23
	0.9	1.61	1.31	1.25	1.23	1.22	1.21	1.22	1.23	1.28
0.3	0.1	1.32	1.36	1.44	1.57	1.90	NA	NA	NA	NA
	0.2	1.36	1.34	1.37	1.41	1.49	1.64	2.09	NA	NA
	0.3	1.44	1.37	1.37	1.39	1.43	1.51	1.66	2.22	NA
	0.4	1.57	1.41	1.39	1.40	1.42	1.47	1.56	1.79	NA
	0.5	1.90	1.49	1.43	1.42	1.43	1.47	1.53	1.68	2.42
	0.6	NA	1.64	1.51	1.47	1.47	1.49	1.53	1.64	2.01
	0.7	NA	2.09	1.66	1.56	1.53	1.53	1.56	1.64	1.88
	0.8	NA	NA	2.22	1.79	1.68	1.64	1.64	1.68	1.85
	0.9	NA	NA	NA	NA	2.42	2.01	1.88	1.85	1.93
0.7	0.1	2.73	NA	NA	NA	NA	NA	NA	NA	NA
	0.2	NA	2.83	3.63	NA	NA	NA	NA	NA	NA
	0.3	NA	3.63	2.95	3.46	NA	NA	NA	NA	NA
	0.4	NA	NA	3.46	3.10	3.57	NA	NA	NA	NA
	0.5	NA	NA	NA	3.57	3.29	3.81	NA	NA	NA
	0.6	NA	NA	NA	NA	3.81	3.54	4.26	NA	NA
	0.7	NA	NA	NA	NA	NA	4.26	3.88	5.30	NA
	0.8	NA	NA	NA	NA	NA	NA	5.30	4.40	NA
	0.9	NA	NA	NA	NA	NA	NA	NA	NA	5.41
0.9	0.1	7.16	NA	NA	NA	NA	NA	NA	NA	NA
	0.2	NA	8.00	NA	NA	NA	NA	NA	NA	NA
	0.3	NA	NA	8.46	NA	NA	NA	NA	NA	NA
	0.4	NA	NA	NA	9.02	NA	NA	NA	NA	NA
	0.5	NA	NA	NA	NA	9.71	NA	NA	NA	NA
	0.6	NA	NA	NA	NA	NA	10.61	NA	NA	NA
	0.7	NA	NA	NA	NA	NA	NA	11.84	NA	NA
	0.8	NA	NA	NA	NA	NA	NA	NA	13.72	NA
	0.9	NA	NA	NA	NA	NA	NA	NA	NA	17.27

APPENDIX B

JOINT PMF OF FOUR OR FIVE DIMENSIONAL BINARY VARIABLES

The joint PMF of four dimensions, using D-vine is

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, Y_4 = y_4) &= P(Y_1 = y_1, Y_4 = y_4 | Y_2 = y_2, Y_3 = y_3) \\ & * P(Y_2 = y_2, Y_3 = y_3) \end{aligned}$$

Table 46: Conditional probability of binary variables using D-vine pair-copula

(y_2,y_3)	(y_1,y_4)	$\mathbf{P}(y_1,\ y_4 \vert y_2,\ y_3)$
(0, 0)	(0, 0)	$C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})$
	(0, 1)	$q_{1\vert 00} - C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})$
	(1, 0)	$q_{4\vert 00} - C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})$
	(1, 1)	$1 - q_{1\vert 00} - q_{4\vert 00} + C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})$
(0, 1)	(0, 0)	$C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})$
	(0, 1)	$q_{1\vert 01} - C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})$
	(1, 0)	$q_{4\vert 01} - C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})$
	(1, 1)	$1 - q_{1\vert 01} - q_{4\vert 01} + C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})$
(1, 0)	(0, 0)	$C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})$
	(0, 1)	$q_{1\vert 10} - C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})$
	(1, 0)	$q_{4\vert 10} - C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})$
	(1, 1)	$1 - q_{1\vert 10} - q_{4\vert 10} + C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})$
(1, 1)	(0, 0)	$C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})$
	(0, 1)	$q_{1\vert 11} - C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})$
	(1, 0)	$q_{4\vert 11} - C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})$
	(1, 1)	$1 - q_{1\vert 11} - q_{4\vert 11} + C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})$

Let $q_{1|00} = P(y_1 = 0|y_2 = 0, y_3 = 0) = \frac{q_2 * C_{13|0}(q_{1|0}, q_{3|0})}{C_{23}(q_2,q_3)}$, $q_{1|01} = P(y_1 = 0|y_2 = 0, y_3 = 1) = \frac{C_{12}(q_1,q_2) - q_2 * C_{13|0}(q_{1|0}, q_{3|0})}{q_2 - C_{23}(q_2,q_3)}$, $q_{1|10} = P(y_1 = 0|y_2 = 1, y_3 = 0) = \frac{p_2 * C_{13|1}(q_{1|1}, q_{3|1})}{q_3 - C_{23}(q_2,q_3)}$, and $q_{1|11} = P(y_1 = 0|y_2 = 1, y_3 = 1) = \frac{q_1 - C_{12}(q_1,q_2) - p_2 * C_{13|1}(q_{1|1}, q_{3|1})}{1 - q_2 - q_3 + C_{23}(q_2,q_3)}$. Also, $q_{4|00} = P(y_4 = 0|y_2 = 0, y_3 = 0)$, $q_{4|01} = P(y_4 = 0|y_2 = 0, y_3 = 1)$, $q_{4|10} = P(y_4 = 0|y_2 = 1, y_3 = 0)$, and $q_{4|11} = P(y_4 = 0|y_2 = 1, y_3 = 1)$ can be obtained similarly. Here, $q_{1|0}$, $q_{1|1}$ and so on are from Section II.2.3.

Based on the conditional probabilities in Table 46, the joint PMF is shown in Table 47.

Table 47: Joint PMF of four-dimension binary variables using D-vine pair-copula

(y_1, y_2, y_3, y_4)	**Probability**
(0, 0, 0, 0)	$C_{14|00}(q_{1|00}, q_{4|00}) * C_{23}(q_2, q_3)$
(0, 0, 0, 1)	$(q_{1|00} - C_{14|00}(q_{1|00}, q_{4|00})) * C_{23}(q_2, q_3)$
(0, 0, 1, 0)	$C_{14|01}(q_{1|01}, q_{4|01}) * (q_2 - C_{23}(q_2, q_3))$
(0, 0, 1, 1)	$(q_{1|01} - C_{14|01}(q_{1|01}, q_{4|01})) * (q_2 - C_{23}(q_2, q_3))$
(0, 1, 0, 0)	$C_{14|10}(q_{1|10}, q_{4|10}) * (q_3 - C_{23}(q_2, q_3))$
(0, 1, 0, 1)	$(q_{1|10} - C_{14|10}(q_{1|10}, q_{4|10})) * (q_3 - C_{23}(q_2, q_3))$
(0, 1, 1, 0)	$C_{14|11}(q_{1|11}, q_{4|11}) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))$
(0, 1, 1, 1)	$(q_{1|11} - C_{14|11}(q_{1|11}, q_{4|11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))$
(1, 0, 0, 0)	$(q_{4|00} - C_{14|00}(q_{1|00}, q_{4|00})) * C_{23}(q_2, q_3)$
(1, 0, 0, 1)	$(1 - q_{1|00} - q_{4|00} + C_{14|00}(q_{1|00}, q_{4|00})) * C_{23}(q_2, q_3)$
(1, 0, 1, 0)	$(q_{4|01} - C_{14|01}(q_{1|01}, q_{4|01})) * (q_2 - C_{23}(q_2, q_3))$
(1, 0, 1, 1)	$(1 - q_{1|01} - q_{4|01} + C_{14|01}(q_{1|01}, q_{4|01})) * (q_2 - C_{23}(q_2, q_3))$
(1, 1, 0, 0)	$(q_{4|10} - C_{14|10}(q_{1|10}, q_{4|10})) * (q_3 - C_{23}(q_2, q_3))$
(1, 1, 0, 1)	$(1 - q_{1|10} - q_{4|10} + C_{14|10}(q_{1|10}, q_{4|10})) * (q_3 - C_{23}(q_2, q_3))$
(1, 1, 1, 0)	$(q_{4|11} - C_{14|11}(q_{1|11}, q_{4|11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))$
(1, 1, 1, 1)	$(1 - q_{1|11} - q_{4|11} + C_{14|11}(q_{1|11}, q_{4|11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))$

Similarly, The joint PMF of five dimensions, using D-vine is

$$\begin{aligned} P(Y_1 &= y_1, Y_2 = y_2, Y_3 = y_3, Y_4 = y_4, Y_5 = y_5) = \\ & P(Y_1 = y_1, Y_5 = y_5|Y_2 = y_2, Y_3 = y_3, Y_4 = y_4) \\ & *P(Y_2 = y_2, Y_4 = y_4|Y_3 = y_3) * P(Y_3 = y_3) \end{aligned}$$

Let $q_{1|000} = P(y_1 = 0|y_2 = 0, y_3 = 0, y_4 = 0) = \frac{C_{14|00}(q_{1|00},q_{4|00})*C_{23}(q_2,q_3)}{q_2*C_{34|0}(q_{3|0},q_{4|0})}$, $q_{1|001} = P(y_1 = 0|y_2 = 0, y_3 = 0, y_4 = 1) = \frac{(q_{1|00}-C_{14|00}(q_{1|00},q_{4|00}))*C_{23}(q_2,q_3)}{C_{23}(q_2,q_3)-q_2*C_{34|0}(q_{3|0},q_{4|0})}$, $q_{1|010} = P(y_1 = 0|y_2 = 0, y_3 = 1, y_4 = 0) = \frac{C_{14|01}(q_{1|01},q_{4|01})*(q_2-C_{23}(q_2,q_3))}{C_{24}(q_2,q_4)-q_2*C_{34|0}(q_{3|0},q_{4|0})}$, $q_{1|011} = P(y_1 = 0|y_2 = 0, y_3 = 1, y_4 = 1) = \frac{(q_{1|01}-C_{14|01}(q_{1|01},q_{4|01}))*(q_2-C_{23}(q_2,q_3))}{q_2-C_{23}(q_2,q_3)-C_{24}(q_2,q_4)+q_2*C_{34|0}(q_{3|0},q_{4|0})}$, $q_{1|100} = P(y_1 = 0|y_2 = 1, y_3 = 0, y_4 = 0) = \frac{C_{14|10}(q_{1|10},q_{4|10})*(q_3-C_{23}(q_2,q_3))}{p_2*C_{34|1}(q_{3|1},q_{4|1})}$, $q_{1|101} = P(y_1 = 0|y_2 = 1, y_3 = 0, y_4 = 1) = \frac{(q_{1|10}-C_{14|10}(q_{1|10},q_{4|10}))*(q_3-C_{23}(q_2,q_3))}{q_3-C_{23}(q_2,q_3)-p_2*C_{34|1}(q_{3|1},q_{4|1})}$, $q_{1|110} = P(y_1 = 0|y_2 = 1, y_3 = 1, y_4 = 0) = \frac{C_{14|11}(q_{1|11},q_{4|11})*(1-q_2-q_3+C_{23}(q_2,q_3))}{q_4-C_{24}(q_2,q_4)-p_2*C_{34|1}(q_{3|1},q_{4|1})}$, and $q_{1|111} = P(y_1 = 0|y_2 = 1, y_3 = 1, y_4 = 1) = \frac{(q_{1|11}-C_{14|11}(q_{1|11},q_{4|11}))*(1-q_2-q_3+C_{23}(q_2,q_3))}{1-q_2-q_3-q_4+C_{24}(q_2,q_4)+C_{23}(q_2,q_3)+p_2*C_{34|1}(q_{3|1},q_{4|1})}$.
Similarly, $q_{5|000} = P(y_5 = 0|y_2 = 0, y_3 = 0, y_4 = 0) = \frac{C_{25|00}(q_{2|00},q_{5|00})*C_{34}(q_3,q_4)}{q_2*C_{34|0}(q_{3|0},q_{4|0})}$, and so on.

Table 48: Conditional probability of binary variables using D-vine pair-copula

(y_2,y_3,y_4)	(y_1,y_5)	$\mathbf{P}(y_1,\ y_5 \vert y_2,\ y_3,\ y_4)$
(0, 0, 0)	(0, 0)	$C_{15\vert 000}(q_{1\vert 000}, q_{5\vert 000})$
	(0, 1)	$q_{1\vert 000} - C_{15\vert 000}(q_{1\vert 000}, q_{5\vert 000})$
	(1, 0)	$q_{5\vert 000} - C_{15\vert 000}(q_{1\vert 000}, q_{5\vert 000})$
	(1, 1)	$1 - q_{1\vert 000} - q_{5\vert 000} + C_{15\vert 000}(q_{1\vert 000}, q_{5\vert 000})$
(0, 0, 1)	(0, 0)	$C_{15\vert 001}(q_{1\vert 001}, q_{5\vert 001})$
	(0, 1)	$q_{1\vert 001} - C_{15\vert 001}(q_{1\vert 001}, q_{5\vert 001})$
	(1, 0)	$q_{5\vert 001} - C_{15\vert 001}(q_{1\vert 001}, q_{5\vert 001})$
	(1, 1)	$1 - q_{1\vert 001} - q_{5\vert 001} + C_{15\vert 001}(q_{1\vert 001}, q_{5\vert 001})$
(0, 1, 0)	(0, 0)	$C_{15\vert 010}(q_{1\vert 010}, q_{5\vert 010})$
	(0, 1)	$q_{1\vert 010} - C_{15\vert 010}(q_{1\vert 010}, q_{5\vert 010})$
	(1, 0)	$q_{5\vert 010} - C_{15\vert 010}(q_{1\vert 010}, q_{5\vert 010})$
	(1, 1)	$1 - q_{1\vert 010} - q_{5\vert 010} + C_{15\vert 010}(q_{1\vert 010}, q_{5\vert 010})$

Continued

(0, 1, 1)	(0, 0)	$C_{15\|011}(q_{1\|011}, q_{5\|011})$
	(0, 1)	$q_{1\|011} - C_{15\|011}(q_{1\|011}, q_{5\|011})$
	(1, 0)	$q_{5\|011} - C_{15\|011}(q_{1\|011}, q_{5\|011})$
	(1, 1)	$1 - q_{1\|011} - q_{5\|011} + C_{15\|011}(q_{1\|011}, q_{5\|011})$
(1, 0, 0)	(0, 0)	$C_{15\|100}(q_{1\|100}, q_{5\|100})$
	(0, 1)	$q_{1\|100} - C_{15\|100}(q_{1\|100}, q_{5\|100})$
	(1, 0)	$q_{5\|100} - C_{15\|100}(q_{1\|100}, q_{5\|100})$
	(1, 1)	$1 - q_{1\|100} - q_{5\|100} + C_{15\|100}(q_{1\|100}, q_{5\|100})$
(1, 0, 1)	(0, 0)	$C_{15\|101}(q_{1\|101}, q_{5\|101})$
	(0, 1)	$q_{1\|101} - C_{15\|101}(q_{1\|101}, q_{5\|101})$
	(1, 0)	$q_{5\|101} - C_{15\|101}(q_{1\|101}, q_{5\|101})$
	(1, 1)	$1 - q_{1\|101} - q_{5\|101} + C_{15\|101}(q_{1\|101}, q_{5\|101})$
(1, 1, 0)	(0, 0)	$C_{15\|110}(q_{1\|110}, q_{5\|110})$
	(0, 1)	$q_{1\|110} - C_{15\|110}(q_{1\|110}, q_{5\|110})$
	(1, 0)	$q_{5\|110} - C_{15\|110}(q_{1\|110}, q_{5\|110})$
	(1, 1)	$1 - q_{1\|110} - q_{5\|110} + C_{15\|110}(q_{1\|110}, q_{5\|110})$
(1, 1, 1)	(0, 0)	$C_{15\|111}(q_{1\|111}, q_{5\|111})$
	(0, 1)	$q_{1\|111} - C_{15\|111}(q_{1\|111}, q_{5\|111})$
	(1, 0)	$q_{5\|111} - C_{15\|111}(q_{1\|111}, q_{5\|111})$
	(1, 1)	$1 - q_{1\|111} - q_{5\|111} + C_{15\|111}(q_{1\|111}, q_{5\|111})$

Based on the conditional probabilities in Table 48, the joint PMF is shown in Table 49.

Table 49: Joint PMF of five-dimension binary variables using D-vine pair-copula

$(y_1, y_2, y_3, y_4, y_5)$	**Probability**
(0, 0, 0, 0, 0)	$C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00}) * C_{23}(q_2, q_3) * q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0})$
(0, 0, 0, 0, 1)	$(q_{1\vert 00} - C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})) * C_{23}(q_2, q_3) * q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0})$
(0, 0, 0, 1, 0)	$C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01}) * (q_2 - C_{23}(q_2, q_3)) * (C_{24}(q_2, q_4)$ $-q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 0, 0, 1, 1)	$(q_{1\vert 01} - C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})) * (q_2 - C_{23}(q_2, q_3)) * (C_{24}(q_2, q_4)$ $-q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 0, 1, 0, 0)	$C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10}) * (q_3 - C_{23}(q_2, q_3)) * p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1})$
(0, 0, 1, 0, 1)	$(q_{1\vert 10} - C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})) * (q_3 - C_{23}(q_2, q_3)) * p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1})$
(0, 0, 1, 1, 0)	$C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11}) * (1 - q_2 - q_3 + C_{23}(q_2, q_3)) * (q_2 - C_{23}(q_2, q_3)$ $-p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(0, 0, 1, 1, 1)	$(q_{1\vert 11} - C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))$ $*(q_2 - C_{23}(q_2, q_3) - p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(0, 1, 0, 0, 0)	$(q_{4\vert 00} - C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})) * C_{23}(q_2, q_3)$ $*(C_{34}(q_3, q_4) - q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 1, 0, 0, 1)	$(1 - q_{1\vert 00} - q_{4\vert 00} + C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00})) * C_{23}(q_2, q_3)*$ $(C_{34}(q_3, q_4) - q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 1, 0, 1, 0)	$(q_{4\vert 01} - C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})) * (q_2 - C_{23}(q_2, q_3))*$ $(q_3 - C_{34}(q_3, q_4) - C_{23}(q_2, q_3) + q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 1, 0, 1, 1)	$(1 - q_{1\vert 01} - q_{4\vert 01} + C_{14\vert 01}(q_{1\vert 01}, q_{4\vert 01})) * (q_2 - C_{23}(q_2, q_3))*$ $(q_3 - C_{34}(q_3, q_4) - C_{23}(q_2, q_3) + q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0}))$
(0, 1, 1, 0, 0)	$(q_{4\vert 10} - C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})) * (q_3 - C_{23}(q_2, q_3))*$ $(q_4 - C_{34}(q_3, q_4) - p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(0, 1, 1, 0, 1)	$(1 - q_{1\vert 10} - q_{4\vert 10} + C_{14\vert 10}(q_{1\vert 10}, q_{4\vert 10})) * (q_3 - C_{23}(q_2, q_3))*$ $(q_4 - C_{34}(q_3, q_4) - p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(0, 1, 1, 1, 0)	$(q_{4\vert 11} - C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))*$ $(1 - q_2 - q_3 - q_4 + C_{23}(q_2, q_3) + C_{34}(q_3, q_4) + p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(0, 1, 1, 1, 1)	$(1 - q_{1\vert 11} - q_{4\vert 11} + C_{14\vert 11}(q_{1\vert 11}, q_{4\vert 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))*$ $(1 - q_2 - q_3 - q_4 + C_{23}(q_2, q_3) + C_{34}(q_3, q_4) + p_3 * C_{24\vert 1}(q_{2\vert 1}, q_{4\vert 1}))$
(1, 0, 0, 0, 0)	$C_{14\vert 00}(q_{1\vert 00}, q_{4\vert 00}) * C_{23}(q_2, q_3) * q_3 * C_{24\vert 0}(q_{2\vert 0}, q_{4\vert 0})$

Continue

(1, 0, 0, 0, 1)	$(q_{1\mid 00} - C_{14\mid 00}(q_{1\mid 00}, q_{4\mid 00})) * C_{23}(q_2, q_3) * q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0})$
(1, 0, 0, 1, 0)	$C_{14\mid 01}(q_{1\mid 01}, q_{4\mid 01}) * (q_2 - C_{23}(q_2, q_3)) * C_{24}(q_2, q_4) -q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0})$
(1, 0, 0, 1, 1)	$(q_{1\mid 01} - C_{14\mid 01}(q_{1\mid 01}, q_{4\mid 01})) * (q_2 - C_{23}(q_2, q_3)) * C_{24}(q_2, q_4) -q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0})$
(1, 0, 1, 0, 0)	$C_{14\mid 10}(q_{1\mid 10}, q_{4\mid 10}) * (q_3 - C_{23}(q_2, q_3)) * p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1})$
(1, 0, 1, 0, 1)	$(q_{1\mid 10} - C_{14\mid 10}(q_{1\mid 10}, q_{4\mid 10})) * (q_3 - C_{23}(q_2, q_3)) * p_3 *C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1})$
(1, 0, 1, 1, 0)	$C_{14\mid 11}(q_{1\mid 11}, q_{4\mid 11}) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))* (q_2 - C_{23}(q_2, q_3) - p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$
(1, 0, 1, 1, 1)	$(q_{1\mid 11} - C_{14\mid 11}(q_{1\mid 11}, q_{4\mid 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))* (q_2 - C_{23}(q_2, q_3) - p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$
(1, 1, 0, 0, 0)	$(q_{4\mid 00} - C_{14\mid 00}(q_{1\mid 00}, q_{4\mid 00})) * C_{23}(q_2, q_3)* (C_{34}(q_3, q_4) - q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0}))$
(1, 1, 0, 0, 1)	$(1 - q_{1\mid 00} - q_{4\mid 00} + C_{14\mid 00}(q_{1\mid 00}, q_{4\mid 00})) * C_{23}(q_2, q_3)* (C_{34}(q_3, q_4) - q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0}))$
(1, 1, 0, 1, 0)	$(q_{4\mid 01} - C_{14\mid 01}(q_{1\mid 01}, q_{4\mid 01})) * (q_2 - C_{23}(q_2, q_3))* (q_3 - C_{34}(q_3, q_4) - C_{23}(q_2, q_3) + q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0}))$
(1, 1, 0, 1, 1)	$(1 - q_{1\mid 01} - q_{4\mid 01} + C_{14\mid 01}(q_{1\mid 01}, q_{4\mid 01})) * (q_2 - C_{23}(q_2, q_3))* (q_3 - C_{34}(q_3, q_4) - C_{23}(q_2, q_3) + q_3 * C_{24\mid 0}(q_{2\mid 0}, q_{4\mid 0}))$
(1, 1, 1, 0, 0)	$(q_{4\mid 10} - C_{14\mid 10}(q_{1\mid 10}, q_{4\mid 10})) * (q_3 - C_{23}(q_2, q_3))* (q_4 - C_{34}(q_3, q_4) - p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$
(1, 1, 1, 0, 1)	$(1 - q_{1\mid 10} - q_{4\mid 10} + C_{14\mid 10}(q_{1\mid 10}, q_{4\mid 10})) * (q_3 - C_{23}(q_2, q_3))* (q_4 - C_{34}(q_3, q_4) - p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$
(1, 1, 1, 1, 0)	$(q_{4\mid 11} - C_{14\mid 11}(q_{1\mid 11}, q_{4\mid 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))* (1 - q_2 - q_3 - q_4 + C_{23}(q_2, q_3) + C_{34}(q_3, q_4) + p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$
(1, 1, 1, 1, 1)	$(1 - q_{1\mid 11} - q_{4\mid 11} + C_{14\mid 11}(q_{1\mid 11}, q_{4\mid 11})) * (1 - q_2 - q_3 + C_{23}(q_2, q_3))* (1 - q_2 - q_3 - q_4 + C_{23}(q_2, q_3) + C_{34}(q_3, q_4) + p_3 * C_{24\mid 1}(q_{2\mid 1}, q_{4\mid 1}))$

APPENDIX C

JOINT PMF OF FOUR DIMENSIONAL MULTINOMIAL VARIABLES WITH FOUR CATEGORIES

Continuing the work in Section IV.2.3, the multinomial variables y_i has 4 possible outcomes: 1, 2, 3, 4 with a fixed success probability as $p_{i,1}$, $p_{i,2}$, $p_{i,3}$ and $p_{i,4}$, with $p_{i,1} + p_{i,2} + p_{i,3} + p_{i,4} = 1$, i=1,2,3,4. The CDF of y_i would be

$$F(y_i = j) \quad = \quad \begin{cases} 0 \quad if \quad y_i = 0; \\ p_{i,1} \quad if \quad y_i = 1; \\ p_{i,1} + p_{i,2} \quad if \quad y_i = 2; \\ p_{i,1} + p_{i,2} + p_{i,3} \quad if \quad y_i = 3; \\ 1 \quad if \quad y_i = 4; \end{cases}$$

The joint PMF of three dimensions multinomial variables with four categories, using D-vine is

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3) = P(Y_1 = y_1, Y_3 = y_3 | Y_2 = y_2) * P(Y_2 = y_2)$$

Same with notations in Chapter IV, $p_{1|2,1|1} = P(Y_1 = 1|Y_2 = 1) = \frac{C_{12}(p_{1,1},p_{2,1})}{p_{2,1}}$,..., $p_{3|2,3|4} = P(Y_3 = 3|Y_2 = 4) = \frac{C_{23}(p_{2,4},p_{3,3})}{p_{2,4}}$.

Table 50: PMF of trivariate multinomial variables with four categories

(y_1, y_2, y_3)	**Probability**
(1, 1, 1)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1})$
(1, 1, 2)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,2\vert 1} + p_{3\vert 2,1\vert 1}) - p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1})$
(1, 1, 3)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, 1 - p_{3\vert 2,4\vert 1}) -$ $p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,2\vert 1} + p_{3\vert 2,1\vert 1})$
(1, 1, 4)	$p_{2,1}p_{1\vert 2,1\vert 1} - p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, 1 - p_{3\vert 2,4\vert 1})$
(1, 2, 1)	$p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, p_{3\vert 2,1\vert 2})$
(1, 2, 2)	$p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, p_{3\vert 2,2\vert 2} + p_{3\vert 2,1\vert 2}) - p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, p_{3\vert 2,1\vert 2})$
(1, 2, 3)	$p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, 1 - p_{3\vert 2,4\vert 2}) -$ $p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, p_{3\vert 2,2\vert 2} + p_{3\vert 2,1\vert 2})$
(1, 2, 4)	$p_{2,2}p_{1\vert 2,1\vert 2} - p_{2,2}C_{\theta_{13\vert Y_2=2}}(p_{1\vert 2,1\vert 2}, 1 - p_{3\vert 2,4\vert 2})$
(1, 3, 1)	$p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, p_{3\vert 2,1\vert 3})$
(1, 3, 2)	$p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, p_{3\vert 2,2\vert 3} + p_{3\vert 2,1\vert 3}) - p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, p_{3\vert 2,1\vert 3})$
(1, 3, 3)	$p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, 1 - p_{3\vert 2,4\vert 3}) -$ $p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, p_{3\vert 2,2\vert 3} + p_{3\vert 2,1\vert 3})$
(1, 3, 4)	$p_{2,3}p_{1\vert 2,1\vert 3} - p_{2,3}C_{\theta_{13\vert Y_2=3}}(p_{1\vert 2,1\vert 3}, 1 - p_{3\vert 2,4\vert 3})$
(1, 4, 1)	$p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, p_{3\vert 2,1\vert 4})$
(1, 4, 2)	$p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, p_{3\vert 2,2\vert 4} + p_{3\vert 2,1\vert 4}) - p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, p_{3\vert 2,1\vert 4})$
(1, 4, 3)	$p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, 1 - p_{3\vert 2,3\vert 4}) -$ $p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, p_{3\vert 2,2\vert 4} + p_{3\vert 2,1\vert 4})$
(1, 4, 4)	$p_{2,4}p_{1\vert 2,1\vert 4} - p_{2,4}C_{\theta_{13\vert Y_2=4}}(p_{1\vert 2,1\vert 4}, 1 - p_{3\vert 2,3\vert 4})$
(2, 1, 1)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, p_{3\vert 2,1\vert 1}) - p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1})$
(2, 1, 2)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, p_{3\vert 2,1\vert 1} + p_{3\vert 2,2\vert 1}) -$ $p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1} + p_{3\vert 2,2\vert 1})$ $-p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, p_{3\vert 2,1\vert 1}) + p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1})$
(2, 1, 3)	$p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, 1 - p_{3\vert 2,4\vert 1}) -$ $p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, p_{3\vert 2,1\vert 1} + p_{3\vert 2,2\vert 1}) -$ $p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, 1 - p_{3\vert 2,4\vert 1}) + p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, p_{3\vert 2,1\vert 1} + p_{3\vert 2,2\vert 1})$
(2, 1, 4)	$p_{2,1}p_{1\vert 2,2\vert 1} - p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1} + p_{1\vert 2,2\vert 1}, 1 - p_{3\vert 2,4\vert 1})$ $+p_{2,1}C_{\theta_{13\vert Y_2=1}}(p_{1\vert 2,1\vert 1}, 1 - p_{3\vert 2,4\vert 1})$

Continued

(2, 2, 2)	$p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}+p_{1\|2,2\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2})- p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2}) -p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}+p_{1\|2,2\|2}, p_{3\|2,1\|2}) + p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}, p_{3\|2,1\|2})$
(2, 2, 3)	$p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}+p_{1\|2,2\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2}+p_{3\|2,3\|2})- p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}+p_{1\|2,2\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2})- p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2}+p_{3\|2,3\|2})+ p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}, p_{3\|2,1\|2}+p_{3\|2,2\|2})$
(2, 2, 4)	$p_{2,2}p_{1\|2,2\|2} - p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}+p_{1\|2,2\|2}, 1-p_{3\|2,4\|2})+ p_{2,2}C_{\theta_{13\|Y_2=2}}(p_{1\|2,1\|2}, 1-p_{3\|2,4\|2})$
(2, 3, 1)	$p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, p_{3\|2,1\|3}) - p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, p_{3\|2,1\|3})$
(2, 3, 2)	$p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, p_{3\|2,1\|3}+p_{3\|2,2\|3})- p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, p_{3\|2,1\|3}+p_{3\|2,2\|3}) -p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, p_{3\|2,1\|3}) + p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, p_{3\|2,1\|3})$
(2, 3, 3)	$p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, 1-p_{3\|2,4\|3})- p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, p_{3\|2,1\|3}+p_{3\|2,2\|3})- p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, 1-p_{3\|2,4\|3})+ p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, p_{3\|2,1\|3}+p_{3\|2,2\|3})$
(2, 3, 4)	$p_{2,3}p_{1\|2,2\|3} - p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}+p_{1\|2,2\|3}, 1-p_{4\|2,1\|3})- +p_{2,3}C_{\theta_{13\|Y_2=3}}(p_{1\|2,1\|3}, 1-p_{3\|2,4\|3})$
(2, 4, 1)	$p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, p_{3\|2,1\|4}) - p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, p_{3\|2,1\|4})$
(2, 4, 2)	$p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, p_{3\|2,1\|4}+p_{3\|2,2\|4})- p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, p_{3\|2,1\|4}+p_{3\|2,2\|4}) -p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, p_{3\|2,1\|4}) + p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, p_{3\|2,1\|4})$
(2, 4, 3)	$p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, 1-p_{3\|2,3\|4})- p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, p_{3\|2,1\|4}+p_{3\|2,2\|4})- p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, 1-p_{3\|2,3\|4})+ p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, p_{3\|2,1\|4}+p_{3\|2,2\|4})$
(2, 4, 4)	$p_{2,4}p_{1\|2,2\|4} - p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}+p_{1\|2,2\|4}, 1-p_{3\|2,3\|4}) +p_{2,4}C_{\theta_{13\|Y_2=4}}(p_{1\|2,1\|4}, 1-p_{3\|2,3\|4})$
(3, 1, 1)	$p_{2,1}C_{\theta_{13\|Y_2=1}}(1-p_{1\|2,4\|1}, p_{3\|2,1\|1})- p_{2,1}C_{\theta_{13\|Y_2=1}}(p_{1\|2,1\|1}+p_{1\|2,2\|1}, p_{3\|2,1\|1})$

Continued

(3, 1, 2)	$p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1}, p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})-$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1}, p_{3\mid 2,1\mid 1})-$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1}, p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})+$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1}, p_{3\mid 2,1\mid 1})$
(3, 1, 3)	$p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1}, 1-p_{3\mid 2,4\mid 1})-$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1}, 1-p_{3\mid 2,4\mid 1})-$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1}, p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})+$ $p_{2,1}C_{\theta_{13\mid Y_2=1}}(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1}, p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})$
(3, 1, 4)	$p_{2,1}(p_{1\mid 2,3\mid 1})$ $-p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1}, 1-p_{3\mid 2,4\mid 1})$ $+p_{2,1}C_{\theta_{13\mid Y_2=1}}(p_{1\mid 2,1\mid 1}+p_{1\mid 2,2\mid 1}, 1-p_{3\mid 2,4\mid 1})$
(3, 2, 1)	$p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2})$
(3, 2, 2)	$p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2})+$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2})$
(3, 2, 3)	$p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2}+p_{3\mid 2,3\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2}+p_{3\mid 2,3\mid 2})+$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2})$
(3, 2, 4)	$p_{2,2}(p_{1\mid 2,3\mid 2})-$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2}+p_{3\mid 2,3\mid 2})+$ $p_{2,2}C_{\theta_{13\mid Y_2=2}}(p_{1\mid 2,1\mid 2}+p_{1\mid 2,2\mid 2}, p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2}+p_{3\mid 2,3\mid 2})$
(3, 3, 1)	$p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3}, p_{3\mid 2,1\mid 3})-$ $p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3}, p_{3\mid 2,1\mid 3})$
(3, 3, 2)	$p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3}, p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3})-$ $p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3}, p_{3\mid 2,1\mid 3})-$ $p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3}, p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3})+$ $p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3}, p_{3\mid 2,1\mid 3})$

Continued

(3, 3, 3)	$p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3},1-p_{3\mid 2,4\mid 3})- p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3},p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3})- p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},1-p_{3\mid 2,4\mid 3})+ p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},p_{3\mid 2,1\mid 3}+p_{3\mid 2,2\mid 3})$
(3, 3, 4)	$p_{2,3}(p_{1\mid 2,3\mid 3})- p_{2,3}C_{\theta_{13\mid Y_2=3}}(1-p_{1\mid 2,4\mid 3},1-p_{3\mid 2,4\mid 3})+ p_{2,3}C_{\theta_{13\mid Y_2=3}}(p_{1\mid 2,1\mid 3}+p_{1\mid 2,2\mid 3},1-p_{3\mid 2,4\mid 3})$
(3, 4, 1)	$p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},p_{3\mid 2,1\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},p_{3\mid 2,1\mid 4})$
(3, 4, 2)	$p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},p_{3\mid 2,1\mid 4}+p_{3\mid 2,2\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},p_{3\mid 2,1\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},p_{3\mid 2,1\mid 4}+p_{3\mid 2,2\mid 4})+ p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},p_{3\mid 2,1\mid 4})$
(3, 4, 3)	$p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},1-p_{3\mid 2,3\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},p_{3\mid 2,1\mid 4}+p_{3\mid 2,2\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},1-p_{3\mid 2,3\mid 4})+ p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},p_{3\mid 2,1\mid 4}+p_{3\mid 2,2\mid 4})$
(3, 4, 4)	$p_{2,4}(p_{1\mid 2,3\mid 4})-p_{2,4}C_{\theta_{13\mid Y_2=4}}(1-p_{1\mid 2,3\mid 4},1-p_{3\mid 2,3\mid 4})- p_{2,4}C_{\theta_{13\mid Y_2=4}}(p_{1\mid 2,1\mid 4}+p_{1\mid 2,2\mid 4},1-p_{3\mid 2,3\mid 4})$
(4, 1, 1)	$p_{2,1}p_{3\mid 2,1\mid 1}-p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},p_{3\mid 2,1\mid 1})$
(4, 1, 2)	$p_{2,1}p_{3\mid 2,2\mid 1}-p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})+ p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},p_{3\mid 2,1\mid 1})$
(4, 1, 3)	$p_{2,1}p_{3\mid 2,3\mid 1}-p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},1-p_{3\mid 2,4\mid 1})+ p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},p_{3\mid 2,1\mid 1}+p_{3\mid 2,2\mid 1})$
(4, 1, 4)	$1-p_{2,1}p_{3\mid 2,3\mid 1}-p_{2,1}p_{1\mid 2,3\mid 1}+ p_{2,1}C_{\theta_{13\mid Y_2=1}}(1-p_{1\mid 2,4\mid 1},1-p_{3\mid 2,4\mid 1})$
(4, 2, 1)	$p_{2,2}p_{3\mid 2,1\mid 2}-p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2},p_{3\mid 2,1\mid 2})$
(4, 2, 2)	$p_{2,2}p_{3\mid 2,2\mid 2}-p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2},p_{3\mid 2,1\mid 2}+p_{3\mid 2,2\mid 2})+ p_{2,2}C_{\theta_{13\mid Y_2=2}}(1-p_{1\mid 2,4\mid 2},p_{3\mid 2,1\mid 2})$

Continued

(4, 2, 3)	$p_{2,2}p_{3\vert 2,3\vert 2} - p_{2,2}C_{\theta_{13\vert Y_2=2}}(1 - p_{1\vert 2,4\vert 2}, 1 - p_{3\vert 2,4\vert 2})+$ $p_{2,2}C_{\theta_{13\vert Y_2=2}}(1 - p_{1\vert 2,4\vert 2}, p_{3\vert 2,1\vert 2} + p_{3\vert 2,2\vert 2})$
(4, 2, 4)	$1 - p_{2,2}p_{3\vert 2,3\vert 2} - p_{2,2}p_{1\vert 2,3\vert 2}+$ $p_{2,2}C_{\theta_{13\vert Y_2=2}}(1 - p_{1\vert 2,4\vert 2}, 1 - p_{3\vert 2,4\vert 2})$
(4, 3, 1)	$p_{2,3}p_{3\vert 2,1\vert 3} - p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, p_{3\vert 2,1\vert 3})$
(4, 3, 2)	$p_{2,3}p_{3\vert 2,2\vert 3} - p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, p_{3\vert 2,1\vert 3} + p_{3\vert 2,2\vert 3})+$ $p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, p_{3\vert 2,1\vert 3})$
(4, 3, 3)	$p_{2,3}p_{3\vert 2,3\vert 3} - p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, 1 - p_{3\vert 2,4\vert 3})+$ $p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, p_{3\vert 2,1\vert 3} + p_{3\vert 2,2\vert 3})$
(4, 3, 4)	$1 - p_{2,3}p_{3\vert 2,3\vert 3} - p_{2,3}p_{1\vert 2,3\vert 3}+$ $p_{2,3}C_{\theta_{13\vert Y_2=3}}(1 - p_{1\vert 2,4\vert 3}, 1 - p_{3\vert 2,4\vert 3})$
(4, 4, 1)	$p_{2,4}p_{3\vert 2,1\vert 4} - p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, p_{3\vert 2,1\vert 4})$
(4, 4, 2)	$p_{2,4}p_{3\vert 2,2\vert 4} - p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, p_{3\vert 2,1\vert 4} + p_{3\vert 2,2\vert 4})+$ $p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, p_{3\vert 2,1\vert 4})$
(4, 4, 3)	$p_{2,4}p_{3\vert 2,3\vert 4} - p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, 1 - p_{3\vert 2,3\vert 4})+$ $p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, p_{3\vert 2,1\vert 4} + p_{3\vert 2,2\vert 4})$
(4, 4, 4)	$1 - p_{2,4}p_{3\vert 2,3\vert 4} - p_{2,4}p_{1\vert 2,3\vert 4}+$ $p_{2,4}C_{\theta_{13\vert Y_2=4}}(1 - p_{1\vert 2,3\vert 4}, 1 - p_{3\vert 2,3\vert 4})$

Now, For the joint PMF of four dimensions multinomial variables with four categories, using D-vine is:

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3, Y_4 = y_4) =& P(Y_1 = y_1, Y_4 = y_4 | Y_2 = y_2, Y_3 = y_3) \\ & * P(Y_2 = y_2, Y_3 = y_3) \end{aligned}$$

Thus, instead of showing the PMF with $4^4 = 256$ outcomes, we present one table for the conditional probability $P(Y_1 = y_1, Y_4 = y_4 | Y_2 = y_2, Y_3 = y_3)$, the other for the bivariate probability $P(Y_2 = y_2, Y_3 = y_3)$, which is adjusted from Table 37.

Table 51: PMF of bivariate multinomial variables with four categories

$(y_2,\, y_3)$	**Probability**
(1, 1)	$C_\theta(p_{2,1}, p_{3,1})$
(1, 2)	$C_\theta(p_{2,1}, p_{3,1} + p_{3,2}) - C_\theta(p_{2,1}, p_{3,1})$
(1, 3)	$C_\theta(p_{2,1}, p_{3,1} + p_{3,2} + p_{3,3}) - C_\theta(p_{2,1}, p_{3,1} + p_{3,2})$
(1, 4)	$p_{2,1} - C_\theta(p_{2,1}, p_{3,1} + p_{3,2} + p_{3,3})$
(2, 1)	$C_\theta(p_{2,1} + p_{2,2}, p_{3,1}) - C_\theta(p_{2,1}, p_{3,1})$
(2, 2)	$C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2}) - C_\theta(p_{2,1} + p_{2,2}, p_{3,1}) - C_\theta(p_{2,1}, p_{3,1} + p_{3,2}) +$ $C_\theta(p_{2,1}, p_{3,1})$
(2, 3)	$C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2} + p_{3,3}) - C_\theta(p_{2,1}, p_{3,1} + p_{3,2} + p_{3,3})$ $-C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2}) + C_\theta(p_{2,1}, p_{3,1} + p_{3,2})$
(2, 4)	$p_{2,2} - C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2} + p_{3,3}) + C_\theta(p_{2,1}, p_{3,1} + p_{3,2} + p_{3,3})$
(3, 1)	$C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1}) - C_\theta(p_{2,1} + p_{2,2}, p_{3,1})$
(3, 2)	$C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2}) - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1})$ $-C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2}) + C_\theta(p_{2,1} + p_{2,2}, p_{3,1})$
(3, 3)	$C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2} + p_{3,3}) - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2})$ $-C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2} + p_{3,3}) + C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2})$
(3, 4)	$p_{2,3} - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2} + p_{3,3})$ $+C_\theta(p_{2,1} + p_{2,2}, p_{3,1} + p_{3,2} + p_{3,3})$
(4, 1)	$p_{3,1} - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1})$
(4, 2)	$p_{3,2} - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2}) + C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1})$
(4, 3)	$p_{3,3} - C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2} + p_{3,3})$ $+C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2})$
(4, 4)	$1 - p_{2,1} - p_{2,2} - p_{2,3} - p_{3,1} - p_{3,2} - p_{3,3}$ $+C_\theta(p_{2,1} + p_{2,2} + p_{2,3}, p_{3,1} + p_{3,2} + p_{3,3})$

Table 52: PMF of quadvariate multinomial variables with four categories

$(y_1,\, y_4 \mid y_2,\, y_3)$	Probability
$(1,\, 1 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(1,\, 2 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(1,\, 3 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})$
$(1,\;\, 4 \mid y_2, y_3)$	$p_{1\mid 23,1\mid y_2y_3}-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})$
$(2,\, 1 \mid y_2,\, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(2,\, 2 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})+$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(2,\, 3 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})+$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})$
$(2,\;\, 4 \mid y_2, y_3)$	$p_{1\mid 23,2\mid y_2y_3}-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})+$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})$
$(3,\, 1 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(1 - p_{1\mid 23,4\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(3,\, 2 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(1 - p_{1\mid 23,4\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(1 - p_{1\mid 23,4\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})+$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3})$
$(3,\, 3 \mid y_2, y_3)$	$C_{\theta_{14\mid Y_2Y_3}}(1 - p_{1\mid 23,4\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})$
	$-C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, 1 - p_{4\mid 23,4\mid y_2y_3})-$
	$C_{\theta_{14\mid Y_2Y_3}}(1 - p_{1\mid 23,4\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})+$
	$C_{\theta_{14\mid Y_2Y_3}}(p_{1\mid 23,1\mid y_2y_3} + p_{1\mid 23,2\mid y_2y_3}, p_{4\mid 23,1\mid y_2y_3} + p_{4\mid 23,2\mid y_2y_3})$

Continued

$(3,\ 4\vert y_2,y_3)$	$p_{1\vert 23,3\vert y_2y_3}-$ $C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},1-p_{4\vert 23,4\vert y_2y_3})$ $+C_{\theta_{14\vert Y_2Y_3}}(p_{1\vert 23,1\vert y_2y_3}+p_{1\vert 23,2\vert y_2y_3},1-p_{4\vert 23,4\vert y_2y_3})$
$(4,1\vert y_2,y_3)$	$p_{4\vert 23,1\vert y_2y_3}-$ $C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},p_{4\vert 23,1\vert y_2y_3})$
$(4,2\vert y_2,y_3)$	$p_{4\vert 23,2\vert y_2y_3})-$ $C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},p_{4\vert 23,1\vert y_2y_3+p_{4\vert 23,2\vert y_2y_3}})+$ $C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},p_{4\vert 23,1\vert y_2y_3})$
$(4,3\vert y_2,y_3)$	$p_{4\vert 23,3\vert y_2y_3}-C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},1-p_{4\vert 23,4\vert y_2y_3})+$ $C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},p_{4\vert 23,1\vert y_2y_3}+p_{4\vert 23,2\vert y_2y_3})$
$(4,\ 4\vert y_2,y_3)$	$1-p_{1\vert 23,1\vert y_2y_3}-p_{1\vert 23,2\vert y_2y_3}-p_{1\vert 23,3\vert y_2y_3}-p_{4\vert 23,1\vert y_2y_3}$ $-p_{4\vert 23,2\vert y_2y_3}-p_{4\vert 23,3\vert y_2y_3}+C_{\theta_{14\vert Y_2Y_3}}(1-p_{1\vert 23,4\vert y_2y_3},1-p_{4\vert 23,4\vert y_2y_3})$

Process above shows the procedure of building joint PMF of multinomial Variables, and it's similar to construct higher dimension joint PMF or joint PMF of multinomial variables with more categories.

APPENDIX D

CONDITIONAL CORRELATION FOR EQUICORRELATED STRUCTURE

From Joe (2014), page 40, formula (2.19) the partial correlation is given by the formula

$$\rho_{m-1,m|1,\dots,m-2} = \frac{\sigma_{m-1,m} - a_{m-1}^T \Sigma_{11}^{-1} a_m}{\sqrt{\sigma_{m-1,m-1} - a_{m-1}^T \Sigma_{11}^{-1} a_{m-1}} \sqrt{\sigma_{m,m} - a_m^T \Sigma_{11}^{-1} a_m}} \qquad (4.0.32)$$

where, $\sigma_{i,j}$ is the standard deviation of normal variable Y_i and Y_j, $a_{m-1} = \{\sigma_{1,m-1}, \dots, \sigma_{m-2,m-1}\}^T$ and $a_m = \{\sigma_{1,m}, \dots, \sigma_{m-2,m}\}^T$. Let $\Sigma = R$ be a correlation matrix, then $\sigma_{i,j} = \rho_{i,j}$ and $\sigma_{i,i} = 1$.

For m-dimensional multivariate normal variables with equicorrelated structure, the correlation matrix is

$$R = \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \dots & \dots & \dots & \dots & \dots \\ \rho & \rho & \rho & \dots & 1 \end{bmatrix}_{m \times m}$$

The partial correlation in Equation (4.0.32) becomes

$$\rho_{m-1,m|1,\dots,m-2} = \frac{\rho - \{\rho, \dots, \rho\}_{m-2} \Sigma_{11}^{-1} \{\rho, \dots, \rho\}_{m-2}^T}{\sqrt{1 - \{\rho, \dots, \rho\}_{m-2} \Sigma_{11}^{-1} \{\rho, \dots, \rho\}_{m-2}^T}^{\,2}} \qquad (4.0.33)$$

Since

$$\Sigma_{11} = \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \dots & \dots & \dots & \dots & \dots \\ \rho & \rho & \rho & \dots & 1 \end{bmatrix}_{(m-2) \times (m-2)} = (1-\rho) I_{m-2} + \rho e e^T$$

therefore,

$$\Sigma_{11}^{-1} = \frac{1}{1-\rho} \{ I_{m-2} - \frac{\rho}{1 + (m-2-1)\rho} e e^T \}$$

where, e is the unit vector. Therefore, Equation (4.0.33) become:

$$\rho_{m-1,m|1,...,m-2} = \frac{\rho - \{\rho,...,\rho\}_{m-2}\Sigma_{11}^{-1}\{\rho,...,\rho\}_{m-2}^T}{1 - \{\rho,...,\rho\}_{m-2}\Sigma_{11}^{-1}\{\rho,...,\rho\}_{m-2}^T}$$

where,

$$\begin{aligned}
\{\rho,...,\rho\}_{m-2} \quad \Sigma_{11}^{-1} & \quad \{\rho,...,\rho\}_{m-2}^T \\
&= \{\rho,...,\rho\}_{m-2}\frac{1}{1-\rho}\{I_{m-2} - \frac{\rho}{1+(m-2-1)\rho}ee^T\}\{\rho,...,\rho\}_{m-2}^T \\
&= \frac{1}{1-\rho}((m-2)\rho^2 - \frac{\rho(m-2)}{1+(m-3)\rho}(m-2)^2\rho^2) \\
&= \frac{(m-2)\rho^2}{1-\rho}(1 - \frac{\rho(m-2)}{1+(m-3)\rho}) \\
&= \frac{(m-2)\rho^2}{1-\rho}\frac{1-\rho}{1+(m-3)\rho} \\
&= \frac{(m-2)\rho^2}{1+(m-3)\rho}
\end{aligned}$$

therefore,

$$\begin{aligned}
\rho_{m-1,m|1,...,m-2} &= \frac{\rho - \frac{(m-2)\rho^2}{1+(m-3)\rho}}{1 - \frac{(m-2)\rho^2}{1+(m-3)\rho}} \\
&= \frac{\rho(1+(m-3)\rho) - (m-2)\rho^2}{1+(m-3)\rho - (m-2)\rho^2} \\
&= \frac{\rho - \rho^2}{m\rho(1-\rho) + (1-\rho)(1-2\rho)} \\
&= \frac{\rho}{1+(m-2)\rho}
\end{aligned}$$

This is the conditional correlation of Y_{m-1}, Y_m conditioned on the other $m-2$ variables Y_1,...,Y_{m-2}. For the D-vine pair-copula, we need $cor(Y_i, Y_i + k|Y_{i+1}, ...Y_{i+k-1})$, we can get it the same way except the first step is to permute indices i and $i+k-1$, and $cor(Y_i, Y_{i+k}|Y_{i+1}, ...Y_{i+k-1}) = \frac{\rho}{1+(k-1)\rho}$.

VITA

Huihui Lin
Department of Computational and Applied Mathematics
Old Dominion University
Norfolk, VA 23529

Education

Ph.D Old Dominion Universtiy, Norfok, VA. (Auguest 2020)
Major: Computational & Applied Mathematics (Statistics).
MS Michigan Technological University, Houghton, MI. (December 2014)
Major: Mathematical Sciences (Statistics).
MS Michigan Technological University, Houghton, MI. (December 2012)
Major: Environmental Engineering.
BS East China University of Science and Technology, China. (June 2006)
Major: Environmental Engineering.

Experience

Graduate Teaching Assistant, Old Dominion University, Norfolk, VA, (01/2016-06/2020).
Graduate Teaching Assistant, Michigan Technological University, Houghton, MI, (01/2012-12/2014).
Graduate Research Assistant, Michigan Technological University, Houghton, MI, (09/2010-12/2011).

Typeset using LaTeX.