Copula in a multivariate mixed discrete-continuous model

Zilko, Aurelius A.; Kurowicka, Dorota

DOI
10.1016/j.csda.2016.02.017

Publication date
2016

Document Version
Final published version

Published in
Computational Statistics & Data Analysis

Citation (APA)
Zilko, A. A., & Kurowicka, D. (2016). Copula in a multivariate mixed discrete-continuous model. *Computational Statistics & Data Analysis*, *103*(C), 28-55. https://doi.org/10.1016/j.csda.2016.02.017

Computational Statistics and Data Analysis 103 (2016) 28–55

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Copula in a multivariate mixed discrete–continuous model

Aurelius A. Zilko *, Dorota Kurowicka

Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 11 June 2015
Received in revised form 15 February 2016
Accepted 15 February 2016
Available online 4 May 2016

Keywords:
Copula
Vine
Mixed models

ABSTRACT

The use of different copula-based models to represent the joint distribution of an eight-dimensional mixed discrete and continuous problem consisting of five discrete and three continuous variables is investigated. The discussion starts with the theoretical properties of the copula-based models. Four different models are constructed for the data collected for the purpose of predicting the length of disruption caused by problems with the train detection system in the Dutch railway network and their performance is tested. The more complex models turn out to represent the data better. Nevertheless, it is shown that the simpler eight dimensional Normal copula still constitutes a statistically sound model for the data.

0. Introduction

Copulas separate information present in the margins from the dependence in the joint distribution. They have been proven to be very attractive in many different applications where a joint distribution of continuous variables is of interest. However, when copulas are used for discrete models, Genest and Nešlehová (2007) show that the popular way of copula parameters estimation, through finding an empirical dependence measure and equating it to the theoretical one, is highly biased. Nevertheless, the maximum likelihood technique can still be used, even if it is much more computationally expensive.

Maximum likelihood estimation of copula parameters for discrete models requires an approximation of a multidimensional integral or evaluating 2^n finite differences of the copula to find the value of the probability mass function of an n-dimensional model. Due to computational costs, many copula applications of discrete models have only involved lower dimensional problems. Nikoloulopoulos and Karlis (2008) constructed a four-dimensional Bernoulli distribution with the help of several different copula families with three parameters and Song et al. (2009) built a trivariate discrete distribution with the Normal copula. In both cases, the copula models worked well and the authors highlighted that the dependence structure between the variables did not only come from the copula but also from the margins.

Nikoloulopoulos (2013) proposed computing the rectangle probabilities using the simulated maximum likelihood approach method. The new approach has been shown in Nikoloulopoulos (2015) to be applicable in dimension of up to 225, even though as dimension and sample size increase, computational burden becomes heavy. Another alternative technique to estimate the parameters uses the Bayesian methods as proposed by Smith and Khaled (2012). However, this technique is also computationally intensive.

The reduction in estimation cost of copula parameters has been achieved in Panagiotelis et al. (2012) by using the copula-vine approach. The multivariate discrete distribution has been constructed with a set of pairwise bivariate (conditional) copulas arranged according to a graphical structure called a regular vine (for more information about vines, see Kurowicka and Joe, 2011). The conditional copulas in this construction are assumed not to depend on the conditioning variables. The computation cost of calculating the probability mass function with this approach only grows as $2n(n-1)$, which makes this model applicable even for very high dimensional problems.

* Corresponding author.
E-mail address: A.A.Zilko@tudelft.nl (A.A. Zilko).

http://dx.doi.org/10.1016/j.csda.2016.02.017
0167-9473/

Similarly to purely discrete models, mixed discrete and continuous models with copulas encounter problems. Most applications of copulas to low dimensional problems are available in the literature. Song et al. (2009) model a bivariate mixed binary discrete (disposition) and continuous (severity of burn injury) variables with a Normal copula. De Leon and Wu (2010) proposed two strategies to compute the maximum likelihood for a bivariate mixed discrete and continuous distribution with a simulation study and an application to the same data set as in Song et al. (2009). He et al. (2012) used the Normal copula to construct two and three dimensional mixed discrete and continuous models each with one discrete variable to study the relationship between the genotype (discrete) and a few continuous phenotypes such as the cholesterol density and the protein concentration. Stöber et al. (2015) constructed a six-dimensional mixed discrete and continuous model with five binary variables and one continuous variable representing six chronic diseases by following the copula-vine approach with constant conditional copulas as described in Panagiotelis et al. (2012).

In the first part of this paper, we concentrate on theoretical issues concerning the use of copulas for purely discrete and mixed discrete–continuous models. A few simple results of the existence of a copula model for the joint distribution of binary variables are provided. This investigation provides a background for the exploration of copula models for a mixed discrete and continuous data presented in Zilko et al. (2015), where five binary and three continuous variables are used to construct a latency time model that is part of the railway disruption length model. The goal is to choose a model that allows fast and accurate prediction of the latency time for different combinations of values of the other variables in the model.

The rest of the paper is organized as follows. Section 1 introduces copula models for multivariate Bernoulli distribution. In Section 2 mixed discrete–continuous models with copulas are presented. Section 3 is concerned with the application of copula models to the latency time data. This section contains the results. Finally, conclusions and short discussions on how the model that is constructed in this paper will be used in practice are presented in Section 4.

1. Multivariate Bernoulli distribution with copulas

The aim of this section is to lay a theoretical background and discuss copula models for discrete and mixed discrete–continuous distributions. We start with the multivariate Bernoulli distribution and investigate the existence of a copula family that allows representation of such a distribution. We present a copula construction that allows to model any multivariate Bernoulli distribution.

Let $\mathbf{X} = (X_1, \dots, X_n)$ be a random vector taking values in $\{0, 1\}^n$ and $\mathbf{x} = (x_1, \dots, x_n)$ be a realization of $\mathbf{X}$. The joint probability is

$$\begin{aligned}\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) &= p(x_1, x_2, \dots, x_n) \\ &= p(0, 0, \dots, 0)^{\prod_{j=1}^{n}(1-x_j)} p(1, 0, \dots, 0)^{x_1 \prod_{j=2}^{n}(1-x_j)} \dots p(1, 1, \dots, 1)^{\prod_{j=1}^{n} x_j}\end{aligned} \tag{1}$$

where all the p's must sum up to 1. The marginal distribution of X_i is

$$\mathbb{P}(X_i = 0) = p_i = \sum_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_n \in \{0,1\}} p(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n).$$

Another popular representation of a multivariate Bernoulli distribution is the log-linear expansion. Taking the logarithm of the probability in (1) and collecting the appropriate terms leads to:

$$\begin{aligned}\log p(x_1, x_2, \dots, x_n) = {} & \log p(0, 0, \dots, 0) + \sum_i u_i x_i + \sum_{i,j} u_{ij} x_i x_j \\ & + \sum_{ijk} u_{ijk} x_i x_j x_k + \dots + u_{12\dots n} x_1 x_2 \dots x_n.\end{aligned} \tag{2}$$

The u-terms in (2) represent the two, three, $\dots$, n-way interactions between the variables (see e.g. Whittaker, 1990) and they can be obtained from the probabilities as follows:

$$\begin{aligned}u_1 &= \log \frac{p(1, 0, 0, \dots, 0)}{p(0, 0, 0, \dots, 0)}, \\ u_{12} &= \log \frac{p(1, 1, 0, \dots, 0)p(0, 0, 0, \dots, 0)}{p(1, 0, 0, \dots, 0)p(0, 1, 0, \dots, 0)}, \\ u_{123} &= \log \frac{p(1, 1, 1, 0, \dots, 0)p(1, 0, 0, 0, \dots, 0)p(0, 1, 0, 0, \dots, 0)p(0, 0, 1, 0, \dots, 0)}{p(1, 1, 0, 0, \dots, 0)p(1, 0, 1, 0, \dots, 0)p(0, 1, 1, 0, \dots, 0)p(0, 0, 0, 0, \dots, 0)}.\end{aligned} \tag{3}$$

The interactions between the variables contain information about dependence. The term u_{12} is also known as the log cross-product ratio (cpr) between variables X_1 and X_2. Notice that the cross product ratio $cpr(X_1, X_2)$ can be rewritten in terms of conditional probabilities of variables X_1 and X_2 given all remaining variables $X_3, \dots, X_n$ equal zero. Moreover, u_{123}

is the logarithm of the ratio of the cross product ratio of variables X_1, X_3 given $X_2 = 1$, and the cross product ratio of X_1, X_3 given $X_2 = 0$.

The symbol u_{ij} represents the two-way dependence between the variables X_i and X_j. However the dependence between X_i and X_j is also affected by all higher order interactions containing these variables.

Example 1.1 (*The Trivariate Bernoulli Distribution*). The trivariate Bernoulli distribution of (X_1, X_2, X_3) is

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, X_3 = x_3) = p(0, 0, 0)^{(1-x_1)(1-x_2)(1-x_3)} \dots p(1, 1, 1)^{x_1 x_2 x_3} \tag{4}$$

for $x_1, x_2, x_3 \in \{0, 1\}$. Its log-linear expansion is:

$$\log p(x_1, x_2, x_3) = u_\emptyset + u_1x_1 + u_2x_2 + u_3x_3 + u_{12}x_1x_2 + u_{13}x_1x_3 + u_{23}x_2x_3 + u_{123}x_1x_2x_3 \tag{5}$$

where $u_\emptyset = \log p(0, 0, 0)$ and the u-terms are as presented in (3).

The conditional distribution of X_1 and X_3 given $X_2 = x_2$ is a bivariate Bernoulli distribution. Let the log cross product ratio of this conditional distribution be denoted as $u_{13|2=x_2}$. When $u_{13|2=x_2} = 0$, the variables $X_1|X_2 = x_2$ and $X_3|X_2 = x_2$ are independent. Moreover, when $u_{13|2=x_2} = 0$ for both realizations of $x_2 = 0$ and $x_2 = 1$, the variables X_1 and X_3 are conditionally independent given variable X_2. Notice that $u_{13} = u_{13|2=0}$ and $u_{123} = \log\left(\frac{cpr(X_1,X_3|X_2=1)}{cpr(X_1,X_3|X_2=0)}\right)$. Hence X_1 and X_3 are conditionally independent given variable X_2 if and only if $u_{123} = 0$ and $u_{13} = 0$.

The above relationships can be generalized for higher-order interactions and allow independencies and conditional independencies to be read from the log-linear expansion by examining the u-terms. Moreover, if the random vector $(X_1, \dots, X_n)$ has the Bernoulli distribution, then it is easy to see that the conditional distributions are also Bernoulli.

The dependences between variables with Bernoulli distributions are contained in the u-terms of its log-linear expansion. A very popular way of examining dependence in a joint distribution, especially for continuous distributions, is to study its corresponding *copula*.

A copula is the joint distribution of n uniform variables in the n-dimensional unit hypercube. Due to the Sklar's theorem (Sklar, 1959), any joint cumulative distribution of variables $(X_1, \dots, X_n)$, denoted as $F_{1,\dots,n}$, can be rewritten in terms of the corresponding copula C as

$$F_{1,\dots,n}(x_1, \dots, x_n) = \mathbb{P}(X_1 \le x_1, \dots, X_n \le x_n) = C(\mathbb{P}(X_1 \le x_1), \dots, \mathbb{P}(X_n \le x_n)). \tag{6}$$

The copula satisfying Eq. (6) is unique if the variables are continuous. When one or more variables are discrete, however, the copula satisfying (6) is no longer unique.

1.1. Bivariate Bernoulli distribution with copulas

To illustrate how copulas are used to model discrete distributions and to present the graphical interpretation of Eq. (6), a bivariate Bernoulli random vector (X_1, X_2) with margins p_1, p_2 is considered. Moreover, let U_{X_1} and U_{X_2} be uniform random variables with copula C. The probability mass function of (X_1, X_2) can be represented in terms of latent variables U_{X_1} and U_{X_2} with copula C as follows:

$$\mathbb{P}(X_1 = x_1, X_2 = x_2) = \begin{cases} p(0, 0), & U_{X_1} \le p_1, \quad U_{X_2} \le p_2; \\ p(0, 1), & U_{X_1} \le p_1, \quad U_{X_2} > p_2; \\ p(1, 0), & U_{X_1} > p_1, \quad U_{X_2} \le p_2; \\ p(1, 1), & U_{X_1} > p_1, \quad U_{X_2} > p_2. \end{cases} \tag{7}$$

Fig. 1 shows the above construction graphically. The two axes in Fig. 1 correspond to the latent vector (U_{X_1}, U_{X_2}). The range $U_{X_1} \in (0, p_1]$ in the vertical axes corresponds to the realization $X_1 = 0$; and $U_{X_2} \in (0, p_2]$ on the horizontal axes to $X_2 = 0$. The mass in the bottom left rectangle is $\mathbb{P}(X_1 = 0, X_2 = 0) = p(0, 0)$.

In this case, Eq. (6) takes the form of

$$\mathbb{P}(X_1 \le x_1, X_2 \le x_2) = C(\mathbb{P}(X_1 \le x_1), \mathbb{P}(X_2 \le x_2)). \tag{8}$$

Since $\mathbb{P}(X_1 \le 1, X_2 \le x_2) = \mathbb{P}(X_2 \le x_2) = C(1, \mathbb{P}(X_2 \le x_2))$ for $x_2 \in \{0, 1\}$ and $\mathbb{P}(X_1 \le x_1, X_2 \le 1) = \mathbb{P}(X_1 \le x_1) = C(\mathbb{P}(X_1 \le x_1), 1)$ for $x_1 \in \{0, 1\}$ hold for any copula, the only constraint on C to realize the distribution of (X_1, X_2) is:

$$p(0, 0) = \mathbb{P}(X_1 \le 0, X_2 \le 0) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0)) = C(p_1, p_2). \tag{9}$$

For continuous random vectors, there exists a unique copula that models the dependence of the joint distribution. However, the copula is constrained to satisfy the Sklar's theorem in every point of the unit hypercube. In the bivariate Bernoulli case, where the constraint is at only one point in the unit square, any copula satisfying (9) will be appropriate to model the dependence of (X_1, X_2). The bounds on copulas satisfying (9) have been presented in Carley (2002). The upper (lower) Carley bound belongs to the family of copulas constructed as a shuffle of the upper (M) and lower (W) Fréchet bounds, where $M = \min(u, v)$ and $W(u, v) = \max(u + v - 1, 0)$ for $(u, v) \in (0, 1)^2$ (Nelsen, 2006). The mass in M and W

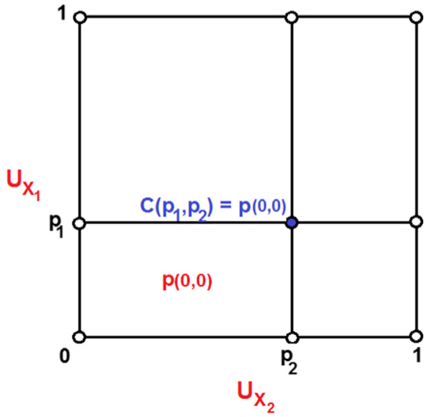


Fig. 1. The unit square corresponding to the latent variables (U_{X_1}, U_{X_2}).

is concentrated uniformly on the diagonal and anti-diagonal of the unit square, respectively. For (U_{X_1}, U_{X_2}) with copula M (W), the Spearman's correlation is $\rho(U_{X_1}, U_{X_2}) = 1(-1)$.

If a one-parametric copula family is already chosen to work with, the non-uniqueness problem of the copulas satisfying (9) is avoided. However, then one might face the problem of non-existence of a copula in the chosen class that satisfies Eq. (9). In the theorem below, the conditions that a copula has to satisfy to be able to recover a bivariate Bernoulli distribution are given.

Theorem 1.1. *Let (X_1, X_2) be Bernoulli distributed random vector. Let C_θ be a one-parametric copula that is continuous with respect to θ and satisfies:*

$$\lim_{\theta\to\theta_L} C_\theta(u, v) = W(u, v) \quad \textit{and} \quad \lim_{\theta\to\theta_U} C_\theta(u, v) = M(u, v) \quad \textit{for } (u, v) \in (0, 1)^2$$

for some θ_L and θ_U, where $W(u, v)$ and $M(u, v)$ are the lower and upper Fréchet bounds, respectively. Then, there exists a θ which satisfies

$$C_\theta(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0)) = C_\theta(p_1, p_2) = p(0, 0) = \mathbb{P}(X_1 \le 0, X_2 \le 0). \tag{10}$$

Proof. Since any copula at point (p_1, p_2) has to lie between the lower and upper Fréchet bounds at (p_1, p_2), $p(0, 0)$ has to satisfy

$$W(p_1, p_2) \le p(0, 0) \le M(p_1, p_2). \tag{11}$$

Since $\lim_{\theta\to\theta_L} C_\theta(p_1, p_2) = W(p_1, p_2)$, $\lim_{\theta\to\theta_U} C_\theta(p_1, p_2) = M(p_1, p_2)$, and C_θ is continuous with respect to θ, inequality (11) together with the Intermediate Value Theorem guarantee the existence of $\theta \in [\theta_L, \theta_U]$ such that Eq. (10) is satisfied. □

Corollary 1.1. *The bivariate Normal copula is defined as:*

$$C_\rho(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \quad (u, v) \in (0, 1)^2, \ \rho \in (-1, 1). \tag{12}$$

This copula is known to be a continuous function of the parameter ρ and

$$\lim_{\rho\to -1} C_\rho(u, v) = W(u, v) \quad \textit{and} \quad \lim_{\rho\to 1} C_\rho(u, v) = M(u, v).$$

Therefore, according to Theorem 1.1, the solution to Eq. (10) exists for the Normal copula.

The corollary above states that one can always find a Normal copula which corresponds to a bivariate Bernoulli random variable. This paper mainly concentrates on the Normal copula as this is the copula that is intended to be used in the application part further on.

In Fig. 2 (left), the relationship between the parameter of the Normal copula and the probability $p(0, 0) = \mathbb{P}(X_1 = 0, X_2 = 0)$ of Bernoulli distribution with margins $p_1 = 0.4$ and $p_2 = 0.8$ is shown. The joint probability $p(0, 0)$ is bounded by $\max(p_1 + p_2 - 1, 0)$ and $\min(p_1, p_2)$. When $p(0, 0) = 0.37$, the parameter of the Normal copula is $\rho = 0.4868$. Fig. 2 (right) illustrates the relationship between the parameter of the Normal copula and $p(0, 0)$ in case of different univariate margins.

The Normal copula is often applied in practice. However, the relationships between the margins, $p(0, 0)$, and the parameters of other copulas can be found as well. These relationships are not available in nice analytic form but in the bivariate case they can be calculated easily. Fig. 3 illustrates the relationship between $p(0, 0)$ and Spearman's correlations realized

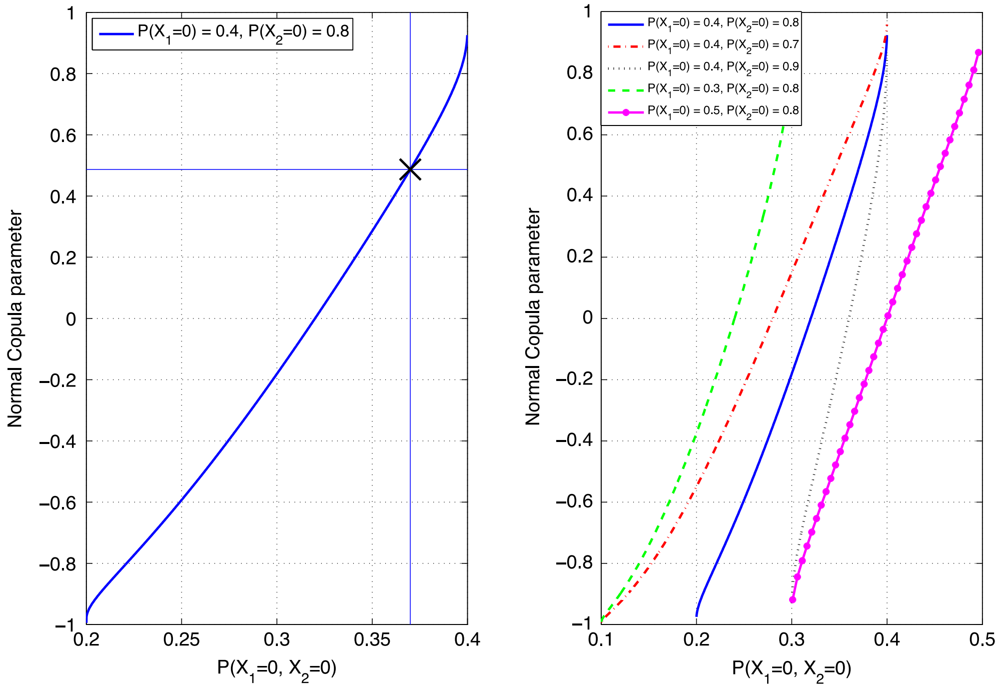


Fig. 2. The plot of the parameter of the Normal copula versus the joint probability of variables X_1 and X_2.

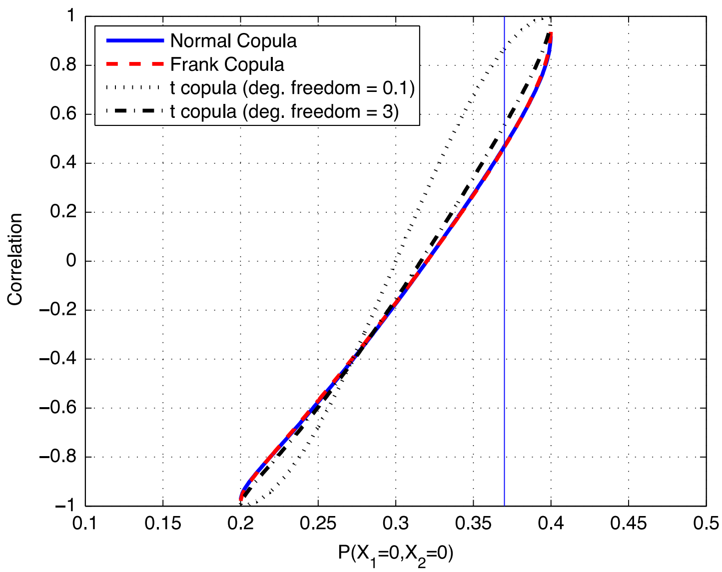


Fig. 3. The plot of the Spearman's correlation of the copula versus the joint probability between variables X_1 and X_2.

by Normal, Frank's and Student t copula in case of the bivariate Bernoulli distribution as in the example depicted in Fig. 2 (left). Fig. 3 shows that for different copula families, the relationships differ slightly.

Next, it will be checked whether the properties of the latent random vector (U_{X_1}, U_{X_2}) translate to equivalent properties of its corresponding Bernoulli distributed variables (X_1, X_2). The example below shows that, in contrast to continuous variables, two Bernoulli distributions constructed with the same copula have very different dependences.

Example 1.2. Let a bivariate Normal copula with parameter $\rho = 0.4868$ be a distribution of latent variables U_{X_1}, U_{X_2}. We fix the first margin to be $p_1 = 0.4$ and the second margin can vary $p_2 \in (0, 1)$. For each p_2, the log cross product ratio and rank correlation of the corresponding Bernoulli distribution are calculated. It turns out that both values are different for different choice of p_2 as presented in Fig. 4. Moreover, the minimum log cross product ratio is obtained at $p_2 = 0.4459$ and the maximum rank correlation is obtained at $p_2 = 0.4375$.

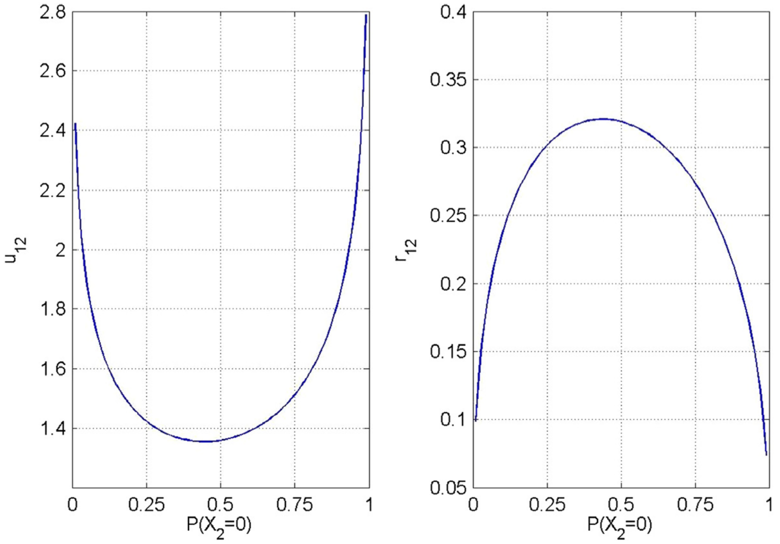


Fig. 4. The log cross product ratio and rank correlation of bivariate Bernoulli with varying $p_2 \in (0, 1)$.

Example 1.2 shows that dependences of Bernoulli distributed random vector depend not only on the copula, but also on the marginal distributions. This observation is in line with previous contributions in Denuit and Lambert (2005), Mesfioui and Tajar (2005) and Nešlehová (2007).

1.2. Trivariate Bernoulli distribution with copulas

In the case of a three-dimensional Bernoulli distribution, the Sklar's equality (6) takes the following form:

$$\mathbb{P}(X_1 \le x_1, X_2 \le x_2, X_3 \le x_3) = C(\mathbb{P}(X_1 \le x_1), \mathbb{P}(X_2 \le x_2), \mathbb{P}(X_3 \le x_3)).$$

Given that the univariate margins of the Bernoulli random vector (X_1, X_2, X_3) are fixed and all probabilities have to sum to one, a copula C needs to satisfy the following four equations

$$\begin{cases} \mathbb{P}(X_1 \le 0, X_2 \le 0, X_3 \le 0) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0), \mathbb{P}(X_3 \le 0)), \\ \mathbb{P}(X_1 \le 0, X_2 \le 0, X_3 \le 1) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0), \mathbb{P}(X_3 \le 1)) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0), 1), \\ \mathbb{P}(X_1 \le 0, X_2 \le 1, X_3 \le 0) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 1), \mathbb{P}(X_3 \le 0)) = C(\mathbb{P}(X_1 \le 0), 1, \mathbb{P}(X_3 \le 1)), \quad \text{and} \\ \mathbb{P}(X_1 \le 1, X_2 \le 0, X_3 \le 0) = C(\mathbb{P}(X_1 \le 1), \mathbb{P}(X_2 \le 0), \mathbb{P}(X_3 \le 0)) = C(1, \mathbb{P}(X_2 \le 0), \mathbb{P}(X_3 \le 1)), \end{cases} \tag{13}$$

to model the dependence of (X_1, X_2, X_3). The second, third, and fourth equations of (13) correspond to the three bivariate margins of the copula. The first equation completes the information needed to construct a trivariate Bernoulli distribution.

Consider a Normal copula C_R with a symmetric and positive definite matrix R of bivariate correlations

$$R = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}.$$

It is easy to see why there might be a problem with the existence of a Normal copula that realizes a given trivariate Bernoulli distribution. The three correlations in the correlation matrix are determined by the second, third, and fourth equations in (13), and each can be computed as in Section 1.1. These three correlations have to (1) form a positive definite matrix R. If this is the case, additionally (2) the first equation in (13) has to be satisfied. In the following example, these problems are highlighted.

Example 1.3. Consider a trivariate Bernoulli distribution with margins $\mathbb{P}(X_1 = 0) = 0.4$, $\mathbb{P}(X_2 = 0) = 0.8$, and $\mathbb{P}(X_3 = 0) = 0.2$ and joint probabilities

- $\mathbb{P}(X_1 = 0, X_2 = 0, X_3 = 0) = 0.01$
- $\mathbb{P}(X_1 = 1, X_2 = 0, X_3 = 0) = 0.16$
- $\mathbb{P}(X_1 = 0, X_2 = 1, X_3 = 0) = 0.02$
- $\mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 0) = 0.01$
- $\mathbb{P}(X_1 = 0, X_2 = 0, X_3 = 1) = 0.36$
- $\mathbb{P}(X_1 = 1, X_2 = 0, X_3 = 1) = 0.27$
- $\mathbb{P}(X_1 = 0, X_2 = 1, X_3 = 1) = 0.01$
- $\mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 1) = 0.16.$

The bivariate margins of this distribution are:

- $\mathbb{P}(X_1 = 0, X_2 = 0) = 0.37$
- $\mathbb{P}(X_2 = 0, X_3 = 0) = 0.17$
- $\mathbb{P}(X_1 = 0, X_3 = 0) = 0.03$.

Notice that the bivariate margin of (X_1, X_2) of the Bernoulli has been discussed in Section 1.1. From the last three equations in (13), we find $\rho_{12} = 0.4868$, $\rho_{13} = -0.4868$, $\rho_{23} = 0.1340$. These form a positive definite matrix. However with the obtained matrix R, $C_R(0.4, 0.8, 0.2) = 0.0298 \neq 0.01 = \mathbb{P}(X_1 = 0, X_2 = 0, X_3 = 0)$. Hence the Bernoulli distribution cannot be recovered with a Normal copula.

A Bernoulli distribution inherits some properties from its corresponding copula. However, it is not easy to specify conditions under which a Bernoulli distribution can be constructed with a given copula. For some special cases we can give conditions under which the construction is possible.

Proposition 1.1. *If the trivariate Bernoulli distribution (Y_1, Y_2, Y_3) obtained from a trivariate Normal copula C_R has margins $\mathbb{P}(Y_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$, then the three-way interaction u_{123} is zero.*

Proof. Since the Normal copula realizes the trivariate Bernoulli distribution of (Y_1, Y_2, Y_3) and $\forall y_1, y_2, y_3 \in \{0, 1\}$ when $\mathbb{P}(Y_i = 0) = 0.5$, the radial symmetry of the trivariate Normal distribution implies:

$$p(y_1, y_2, y_3) = p(1 - y_1, 1 - y_2, 1 - y_3).$$

Therefore, in this case we get

$$\begin{aligned} u_{123} &= \log\left(\frac{p(1,1,1)p(1,0,0)p(0,1,0)p(0,0,1)}{p(1,1,0)p(1,0,1)p(0,1,1)p(0,0,0)}\right) \\ &= \log\left(\frac{p(0,0,0)p(0,1,1)p(1,0,1)p(1,1,0)}{p(1,1,0)p(1,0,1)p(0,1,1)p(0,0,0)}\right) = 0. \quad \square \end{aligned}$$

By the construction of the Bernoulli distribution from a copula, we can immediately get the following result.

Proposition 1.2. *If the margins of the Bernoulli distribution are $\mathbb{P}(X_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$ and the distribution has a radial symmetry, i.e. $p(x_1, x_2, x_3) = p(1 - x_1, 1 - x_2, 1 - x_3)$ for $x_i \in \{0, 1\}$, (X_1, X_2, X_3) can be realized with a trivariate Normal copula.*

Proof. Since the margins are equal to 0.5, writing $p(1, 1, 1)$ in terms of $p(0, 0, 0)$ leads to:

$$p(1,1,1) = -0.5 + p(0,0,0) + p(0,0,1) + p(0,0,0) + p(0,1,0) + p(0,0,0) + p(1,0,0) - p(0,0,0).$$

Because of the radial symmetry, the above equation becomes:

$$p(0,0,0) = \frac{\mathbb{P}(X_1 \le 0, X_2 \le 0, X_3 \le 1) + \mathbb{P}(X_1 \le 0, X_2 \le 1, X_3 \le 0) + \mathbb{P}(X_1 \le 1, X_2 \le 0, X_3 \le 0) - 0.5}{2}.$$

The numerator of the right hand side of the above equation is determined from the second, third, and fourth equations of (13). Therefore, the first equation of (13) is automatically satisfied by the parameters obtained from the other three equations. $\square$

Proposition 1.3. *If the three-way interaction of a trivariate Bernoulli distribution (X_1, X_2, X_3) is zero and $\mathbb{P}(X_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$, a trivariate Normal copula is able to realize (X_1, X_2, X_3).*

Proof. The proof can be found in the Appendix.

Proposition 1.3 implies that a zero three-way interaction does not guarantee the existence of a Normal copula that corresponds to the given Bernoulli distribution.

Let the latent vector $(U_{Y_1}, U_{Y_2}, U_{Y_3})$ be joined by a Normal copula C_R such that U_{Y_1} and U_{Y_3} are independent conditionally on U_{Y_2}. This happens when the correlations in R are such that $\rho_{13} = \rho_{12} \cdot \rho_{23}$ (Whittaker, 1990). The example below shows that the conditional independence of latent variables does not translate to the corresponding Bernoulli random vector (Y_1, Y_2, Y_3).

Example 1.4. Let (Y_1, Y_2, Y_3) be a trivariate Bernoulli distribution realized by a bivariate Normal copula C with parameters $\rho_{12} = 0.5$, $\rho_{13} = -0.5$, and $\rho_{23} = \rho_{12} \cdot \rho_{13} = -0.25$ and let $\mathbb{P}(Y_1 = 0) = 0.4$, $\mathbb{P}(Y_2 = 0) = 0.8$, and $\mathbb{P}(Y_3 = 0) = 0.2$. We have that U_{Y_2}, U_{Y_3} are conditionally independent given U_{Y_1}.

The variables $Y_2|Y_1$ and $Y_3|Y_1$ where (Y_1, Y_2, Y_3) is constructed with copula C are not conditionally independent. In fact, the u-terms of the distribution of (Y_1, Y_2, Y_3) are: $u_{123} = -0.1408 \neq 0$ and $u_{23} = -0.7485 \neq 0$.

Even in the case when $\mathbb{P}(Y_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$, the conditional independence of the latent variables does not translate to the conditional independence of the Y's. With the same correlation matrix as above and all margins equal to 0.5, the u-term are: $u_{23} = -0.6491 \neq 0$ and $u_{123} = 0$.

Corollary 1.2. *From Proposition 1.3, if X_1 and X_3 are conditionally independent given X_2 and the margins are $\mathbb{P}(X_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$, then (X_1, X_2, X_3) can be recovered with a trivariate Normal copula.*

1.3. Non-constant conditional copula-vine

In the previous section, we have investigated the existence of Normal copula for a given Bernoulli distribution. It was shown that in general it is not easy to give conditions which assert the existence or non-existence of the Normal copula for a specified Bernoulli distribution. This section examines a copula-vine approach to construct a multivariate discrete distribution proposed by Panagiotelis et al. (2012), which is based on the copula-vine specification (Kurowicka and Cooke, 2006).

In the copula-vine approach, the conditional distributions in the standard factorization of the joint probability mass function of $(X_1, \dots, X_n)$, namely

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2|X_1 = x_1)\dots\mathbb{P} \times (X_n = x_n|X_1 = x_1, \dots, X_{n-1} = x_{n-1}) \quad (14)$$

are decomposed with specially chosen bivariate conditional distributions which are represented with conditional copulas. Following Panagiotelis et al. (2012), let $\mathbf{V}$ denote the conditioning set of the jth term on the right hand side of (14), and let $\mathbf{V}_{\setminus i}$ represent $\mathbf{V}$ without X_i. Each term $\mathbb{P}(X_j = x_j|\mathbf{V})$ for all j and $i < j$ on the right hand side of (14) becomes

$$\mathbb{P}(X_j = x_j|\mathbf{V} = \mathbf{v}) = \frac{\mathbb{P}(X_j = x_j, X_i = x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})}{\mathbb{P}(X_i = x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})}. \quad (15)$$

The numerator of the right-hand side of (15) can be rewritten and then expressed with copula

$$\begin{aligned}\mathbb{P}(X_j = x_j, X_i = x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) &= \sum_{s_j=0}^{x_j}\sum_{s_i=0}^{x_i}(-1)^{(s_j+s_i)}\mathbb{P}(X_j \le x_j - s_j, X_i \le x_i - s_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) \\ &= \sum_{s_j=0}^{x_j}\sum_{s_i=0}^{x_i}(-1)^{(s_j+s_i)}C_{X_j,X_i|\mathbf{V}_{\setminus i}}\left(\mathbb{P}(X_j \le x_j - s_j|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}), \mathbb{P}(X_i \le x_i - s_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})\right). \quad (16)\end{aligned}$$

For $i \ne k < j$, the arguments of the copula on the right-hand side of the above expression can also be expressed with copula:

$$\mathbb{P}(X_j \le x_j - s_j|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) = \frac{\sum_{s_k=0}^{x_k}(-1)^{s_k}C_{X_j,X_k|\mathbf{V}_{\setminus i,k}}\left(\mathbb{P}(X_j \le x_j - s_j|\mathbf{V}_{\setminus i,k}), \mathbb{P}(X_k \le x_k - s_k|\mathbf{V}_{\setminus i,k})\right)}{\mathbb{P}(X_k = x_k|\mathbf{V}_{\setminus i,k})}. \quad (17)$$

Note that the decomposition above is performed sequentially until the conditioning set $\mathbf{V}$ is empty. All (conditional) copulas appearing in this decomposition can be organized using a graphical structure called regular vine (Kurowicka and Cooke, 2006).

In Panagiotelis et al. (2012), the assumption has been made that all conditional copulas do not depend on the conditioning variables. The conditional copulas in general do not have to be constant, so the following conditions have been proposed in Panagiotelis et al. (2012) to ensure the existence of constant conditional copulas for the multivariate Bernoulli case.

Proposition 1.4. *Let $p_{j,(1)}, \dots, p_{j,(\kappa_1)}$ denote the ordered κ_1 distinct values of $\mathbb{P}(X_j \le 0|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})$ and $p_{i,(1)}, \dots, p_{i,(\kappa_2)}$ denote the ordered κ_2 distinct values of $\mathbb{P}(X_i \le 0|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})$. A constant bivariate copula C exists over the conditioning set $\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}$ if it solves*

$$\mathbb{P}(X_j \le 0, X_i \le 0|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) = C\left(\mathbb{P}(X_j \le 0|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}), \mathbb{P}(X_i \le 0|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})\right)$$

for each member in the conditioning set. For this to happen, all of the $(\kappa_1+1)(\kappa_2+1)$ values of $\mathbb{P}(p_{j,(a)}, p_{i,(b)}) - \mathbb{P}(p_{j,(a-1)}, p_{i,(b)}) - \mathbb{P}(p_{j,(a)}, p_{i,(b-1)}) + \mathbb{P}(p_{j,(a-1)}, p_{i,(b-1)})$ must be non-negative.

Even if a constant conditional copula exists for the above construction, the copula does not have to be a Normal copula. To the best of the authors' knowledge, there is no result ensuring when the constant Normal copula exists.

For large models the assumption of constant conditional copula is understandable. In case of moderate sized models with variables that do not contain many states, it might be not prohibitive to consider the non-constant conditional copula-vine model. In such cases, different copulas can be specified for each combination of conditioning variables in (16).

Theorem 1.2. *Any multivariate Bernoulli random variables can be represented with the bivariate Normal copulas with the non-constant conditional copula-vine model.*

Proof. Since the conditional distribution of a multivariate Bernoulli distribution is also multivariate Bernoulli and for each combination of conditioning variables in Eq. (16) the copulas are allowed to be different, then the result follows immediately from Theorem 1.1. □

To illustrate how the non-constant conditional copula-vine model works, a simple example on the trivariate Bernoulli variable is presented.

Example 1.5. Let (X_1, X_2, X_3) be a trivariate Bernoulli distribution. First, two bivariate unconditional copulas are fixed, say C_{12} and C_{23}. Then, conditional copulas need to be specified for both realizations of variable X_2 which are denoted as: $C_{13|2=0}$ and $C_{13|2=1}$.

The parameters of four bivariate Normal copulas have to be such that Eqs. (13) are all satisfied. The parameters ρ_{12} and ρ_{23} of the Normal copula C_{12} and C_{23} are found from the second and fourth equations of (13), respectively. One finds the parameter of copula $C_{13|2=0}$ to satisfy the first equation and $C_{13|2=1}$ the fourth equation as follows:

$$\begin{aligned}\mathbb{P}(X_1 \le 0, X_2 \le 0, X_3 \le 0) &= \mathbb{P}(X_1 \le 0, X_2 = 0, X_3 \le 0) = \mathbb{P}(X_1 \le 0, X_3 \le 0|X_2 = 0)\mathbb{P}(X_2 = 0)\\ &= C_{13|2=0}(\mathbb{P}(X_1 \le 0|X_2 = 0), \mathbb{P}(X_3 \le 0|X_2 = 0))\mathbb{P}(X_2 = 0)\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(X_1 \le 0, X_2 \le 1, X_3 \le 0) &= \mathbb{P}(X_1 \le 0, X_2 = 1, X_3 \le 0) + \mathbb{P}(X_1 \le 0, X_2 = 0, X_3 \le 0)\\ &= C_{13|2=1}(\mathbb{P}(X_1 \le 0|X_2 = 1), \mathbb{P}(X_3 \le 0|X_2 = 1))\mathbb{P}(X_2 = 1) + \mathbb{P}(X_1 \le 0, X_2 = 0, X_3 \le 0).\end{aligned}$$

Since there is no constraint on copulas in the above representation, any copula can be chosen and all four equations in (13) are satisfied. These conditional copulas correspond to $\mathbb{P}(X_1 \le 0, X_3 \le 0|X_2 = 0)$ and $\mathbb{P}(X_1 \le 0, X_3 \le 0|X_2 = 1)$; both are bivariate Bernoulli distributions.

For the Bernoulli distribution in Example 1.3, the parameters of the Normal copulas can be calculated as above and are equal to: $\rho_{12} = 0.4868$, $\rho_{23} = 0.1340$, $\rho_{13|2=0} = -0.7612$, and $\rho_{13|2=1} = 0.8552$. We see that the Normal copula parameters are very different for different realization of variable X_2.

Moreover, $(X_1, X_3|X_2)$ cannot be represented with any constant conditional copula. In this example, $\mathbb{P}(X_1 = 0|X_2 = 0) = 0.4625 > \mathbb{P}(X_1 = 0|X_2 = 1) = 0.1500$, $\mathbb{P}(X_3 = 0|X_2 = 0) = 0.2125 > \mathbb{P}(X_3 = 0|X_2 = 1) = 0.1500$, and $\mathbb{P}(X_1 = 0, X_3 = 0|X_2 = 0) = 0.0125 < \mathbb{P}(X_1 = 0, X_3 = 0|X_2 = 1) = 0.1000$. This results in the probability in the region $([0, \mathbb{P}(X_1 = 0|X_2 = 0)] \times [0, \mathbb{P}(X_3 = 0|X_2 = 0)]) \setminus ([0, \mathbb{P}(X_1 = 0|X_2 = 1)] \times [0, \mathbb{P}(X_3 = 0|X_2 = 1)])$ to be negative which is a violation of the condition in Proposition 1.4.

For a joint Normal copula, each marginal as well as conditional copula are Normal and the conditional copulas do not depend on the conditioning variables. This property does not translate to the trivariate Bernoulli distribution (Y_1, Y_2, Y_3) implied by the Normal copula. It is not always the case that the conditional copulas $C_{13|2=0}$ and $C_{13|2=1}$ are equal.

Proposition 1.5. *Let the univariate margins of a trivariate Bernoulli distribution* (X_1, X_2, X_3) *are* $\mathbb{P}(X_i = 0) = 0.5$ *for all* $i \in \{1, 2, 3\}$. *The trivariate Normal copula realizes* (X_1, X_2, X_3) *if and only if* $C_{ij|k=0} = C_{ij|k=1}$ *for any combinations of* $i, j, k \in \{1, 2, 3\}$ *where* $C_{ij|k}$ *is a radially symmetric copula.*

Proof. The proof can be found in the Appendix.

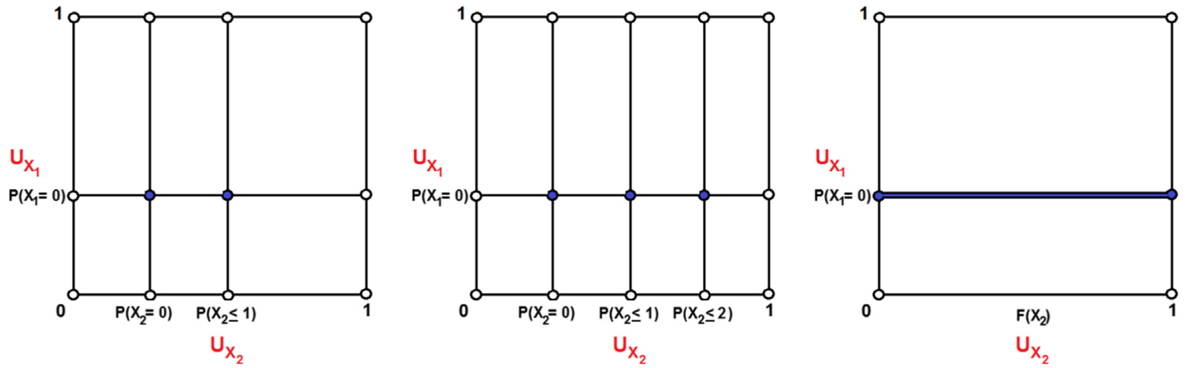
According to Proposition 1.5, if $C_{13|2=0}$ and $C_{13|2=1}$ are the independent copulas and the margins are 0.5, then the distribution of (Y_1, Y_2, Y_3) can be represented by the Normal copula and the variables Y_1 and Y_3 are conditionally independent given variable Y_2. However, the latent variables U_{Y_1} and U_{Y_3} are *not* conditionally independent given U_{Y_2}. This is illustrated in the following example.

Example 1.6. Let (Y_1, Y_2, Y_3) be a trivariate Bernoulli distribution with $\mathbb{P}(Y_i = 0) = 0.5$ for all $i \in \{1, 2, 3\}$ and both $C_{13|2=0}$ and $C_{13|2=1}$ are the independent copulas. Assume that the bivariate margins (Y_1, Y_2) and (Y_2, Y_3) are represented by the bivariate Normal copula with parameters 0.5 and −0.5, respectively. Then Y_1 and Y_3 are conditionally independent given Y_2 and the trivariate Normal copula with parameters $r_{12} = 0.5$, $r_{23} = -0.5$, and $r_{13} = -0.1736$ represents the trivariate Bernoulli distribution (Y_1, Y_2, Y_3). However, $r_{12} \cdot r_{23} = -0.25 \ne -0.1736 = r_{13}$ which means that the latent variables U_{Y_1} and U_{Y_3} are not conditionally independent given U_{Y_2}.

Any multivariate Bernoulli distribution can always be recovered with the non-constant conditional copula-vine approach. This means that it can be constructed using building blocks consisting of bivariate Normal copulas. It is, however, not the case when the conditioning copulas in the copula-vine approach are assumed not to depend on the conditioning variables, and neither in the case of multivariate Normal copula.

2. Mixed discrete–continuous distributions with copulas

In this section, we discuss the extension of the copula modeling to discrete distributions with more than two states and mixed discrete–continuous models. In case when the variables have more states, a copula used in the construction of such a discrete distribution must satisfy more constraints.



(a) X_2 with three states. (b) X_2 with four states. (c) X_2 continuous.

Fig. 5. The unit square corresponding to the latent variable (U_{X_1}, U_{X_2}) for distributions of X_2 with different number of states. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.1. Bivariate case

We start the exposition with the simplest possible case: a bivariate discrete distribution (X_1, X_2) with a margin taking values on $\{0, 1\}$ and another taking values on $\{0, 1, 2\}$. In this case, a copula C that realizes (X_1, X_2) must satisfy the following conditions:

$$\begin{cases} \mathbb{P}(X_1 \le 0, X_2 \le 0) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 0)), \\ \mathbb{P}(X_1 \le 0, X_2 \le 1) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le 1)). \end{cases} \tag{18}$$

The Normal copula cannot always represent (X_1, X_2) anymore because of the over-determined system (18) that needs to be satisfied.

With more states of the variables, the problem deteriorates simply because the number of equations in (18) increases while the number of parameters remains the same. When X_2 is continuous, it has infinitely many states and a copula C that is able to recover the distribution of (X_1, X_2) needs to satisfy the following constraint for all realizations of X_2:

$$\mathbb{P}(X_1 \le 0, X_2 \le x_2) = C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le x_2)). \tag{19}$$

Fig. 5 illustrates the problems of finding a copula graphically for distributions with more than two states. Fig. 5(a) is when X_2 has three states, Fig. 5(b) is when X_2 has four states, and Fig. 5(c) is when X_2 is continuous. The blue dots in Fig. 5(a) and (b) show where in the unit square the copula is constrained. When X_2 is continuous, the copula is constrained at all points on the horizontal blue line in Fig. 5(c).

Using (19), to see whether the copula C can model the mixed discrete–continuous bivariate variable (X_1, X_2), the conditional distribution of X_2 given X_1 should be compared with the conditional distribution of a copula:

$$\mathbb{P}(X_2 \le x_2 | X_1 \le 0) = \frac{C(\mathbb{P}(X_1 \le 0), \mathbb{P}(X_2 \le x_2))}{\mathbb{P}(X_1 \le 0)}. \tag{20}$$

The following example illustrates this.

Example 2.1. Let (X_1, X_2) be a mixed discrete–continuous bivariate random variable with X_1 binary, $\mathbb{P}(X_1 = 0) = 0.75$, and X_2 continuous with marginal distribution $\mathbb{P}(X_2 \le x_2)$. Fig. 6(a) shows the conditional distribution of the latent variable U_{X_2} given $U_{X_1} \le 0.75$ for Normal copulas with different parameters. Fig. 6(b) illustrates the conditional distributions as in (a) in case when Frank and Gumbel copulas are used to model dependence between U_{X_2} and U_{X_1}.

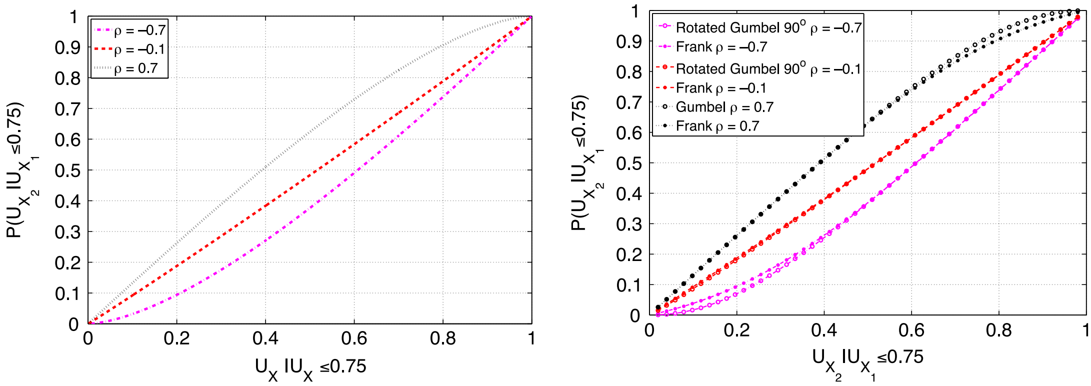
When the variable X_1 is also taken to be continuous, the conditions on the copula become

$$\mathbb{P}(X_1 \le x_1, X_2 \le x_2) = C(\mathbb{P}(X_1 \le x_1), \mathbb{P}(X_2 \le x_2)),$$

which is none other than Sklar's equation (6). In this case, the copula must conform to all points in the unit square of the latent variable (U_{X_1}, U_{X_2}).

2.2. Multivariate case

In the higher-dimensional case, it becomes even more difficult to find the copula that generates the mixed model. The non-constant conditional copula-vine approach is presented in Section 1.3 can be seen as the most flexible model in this case. The conditional probabilities in (16) containing only discrete variables in the conditioning set have to be replaced by the following



(a) Three different $U_{X_2}|U_{X_1} \leq 0.75$ from three Normal copulas.

(b) Six different $U_{X_2}|U_{X_1} \leq 0.75$ from three Frank copulas and three Gumbel copulas.

Fig. 6. The conditional distribution of the latent variable U_{X_2} given variable $U_{X_1} \leq 0.75$ with different copulas.

$$\mathbb{P}(X_j = x_j, X_i \leq x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) = \sum_{s_j=0}^{x_j}(-1)^{s_j}\mathbb{P}(X_j \leq x_j - s_j, X_i \leq x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}) \tag{21}$$

$$= \sum_{s_j=0}^{x_j}(-1)^{s_j}C_{X_j,X_i|\mathbf{V}_{\setminus i}}\left(\mathbb{P}(X_j \leq x_j - s_j|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i}), \mathbb{P}(X_i \leq x_i|\mathbf{V}_{\setminus i} = \mathbf{v}_{\setminus i})\right) \tag{22}$$

where X_j corresponds to the binary variable, X_i corresponds to the continuous variable, and $\mathbf{V}_{\setminus i}$ contains only discrete variables. In this case, the approach as illustrated in Example 2.1 can be used for each combination of conditioning variables in (22).

3. Copula application for the latency time data

In this section, we apply the theory discussed in the first part of the paper to model the disruption length data presented in Zilko et al. (2015). This data has been collected for the purpose of predicting the length of disruptions caused by problems with the train detection system in the Dutch railway network. In this paper, we have chosen five discrete and two continuous variables that are going to be used as predictors of a continuous variable describing the latency time (the time needed for a mechanic to arrive at the problematic site). The full model will contain an additional part that deals with the repair time of the failures.

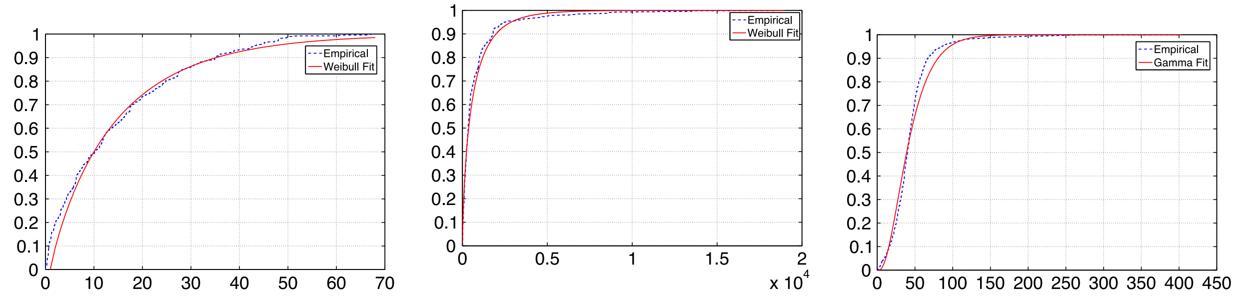
3.1. The data

The data set contains information of the detection system failure in the Dutch railway network from 1 January 2011 until 30 June 2013. During this period, 1932 urgent detection system incidents are recorded. First, we present the variables involved in the latency time model. The latency time model consists of eight variables:

1. The contract type (Contract Type/*CT*): 1 corresponds to the new contract type and 0 corresponds to old contract type. The new contract introduces a penalty factor if the repair work takes too long.
2. The distance to the nearest mechanics workshop (Workshop Distance/*WD*) in kilometers.
3. The distance to the nearest level crossing (Level Cross Distance/*LC*) in meters.
4. Whether the failure occurs during the mechanic's contractual working time or not (Working Time/*WT*): 1 corresponds to the contractual working time and 0 otherwise.
5. Whether the failure occurs during the rush hour period or not (Rush Hour/*RH*): 1 corresponds to the rush hour time and 0 otherwise.
6. Whether the temperature at the time of failure is 25 °C and above or the otherwise (Warm/*WM*): 1 corresponds to temperature of 25 °C and above and 0 otherwise.
7. The existence of another failure that occurs at the same time (Overlap/*OV*): 1 corresponds to the existence of an overlapping failure and 0 otherwise.
8. The latency time (Latency Time/*LAT*) in minutes.

Three of the variables, Workshop Distance, Level Cross Distance, and Latency Time, are continuous while the other five are binary.

The margins of the five binary variables are $\mathbb{P}(CT = 0) = 0.4550$, $\mathbb{P}(WT = 0) = 0.6755$, $\mathbb{P}(RH = 0) = 0.7505$, $\mathbb{P}(WM = 0) = 0.9415$, and $\mathbb{P}(OV = 0) = 0.9513$. Zilko et al. (2015) have tested the three continuous variables for



(a) Workshop Distance (*WD*). (b) Level Cross Distance (*LC*). (c) Latency Time (*LAT*).

Fig. 7. The empirical distributions of the three continuous variables: Workshop Distance, Level Cross Distance, and latency time. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the possibility of representing them with parametric distributions. The margins were tested against the exponential, log-normal, Gamma, and Weibull distributions whose parameters were calculated with maximum likelihood. The considered parametric distributions were rejected for all three variables. The best p-values of the Kolmogorov–Smirnov goodness of fit test (KS Test) of these parametric distributions were 3.11e−05 for the Weibull distribution and *WD*, 1.68e−10 for the Weibull distribution and *LC*, and 1.52e−15 for the Gamma distribution and *LAT*. The dashed blue lines in Fig. 7 are the empirical marginal distributions of the three continuous variables while the red lines represent the corresponding fitted parametric distributions. In the following sections where the dependence structure is modeled with copulas, we use the empirical distribution functions to estimate the margins of the continuous variables.

Using the procedure as in Breymann et al. (2003) which is based on the probability integral transform (PIT test) to test whether Normal copula recovers the dependence between the three continuous variables, we obtain a p-value of 0.7936, indicating good fit of the copula.

3.1.1. Normal copula

The correlation matrix R of the multivariate Normal copula representing a mixed discrete and continuous random vector $(X_1, \ldots, X_k, X_{k+1}, \ldots, X_n)$, where the first k variables are binary and the rest are continuous, is estimated via likelihood maximization. We consider M realizations of $(X_1, \ldots, X_n)$ where the ith realization is denoted as $(x_{i1}, \ldots, x_{in}) = (\boldsymbol{x_{id}}, \boldsymbol{x_{ic}})$. We denote $R[k+1, \ldots, n]$ as the part of the correlation matrix corresponding to the purely continuous part of the model and $R[1, \ldots, k|k+1, \ldots, n]$ as the correlation matrix of the Normal copula $C_{R[1,\ldots,k|k+1,\ldots,n]}$ corresponding to the conditional distribution of discrete variables given the continuous variables in the model. The log-likelihood is defined as

$$\ell(R) = \sum_{i=1}^{M} \log \left(c_{R[k+1,\ldots,n]}(\boldsymbol{x_{ic}}) p_{R[1,\ldots,k|k+1,\ldots,n]}(\boldsymbol{x_{id}}|\boldsymbol{x_{ic}}) \right), \tag{23}$$

where the probability mass function is obtained through calculating the finite difference of the multivariate Normal copula of discrete variables conditional on the continuous variables:

$$p_{R[1,\ldots,k|k+1,\ldots,n]}(x_{i1}, \ldots, x_{ik}) = \sum_{s_1=0}^{x_{i1}} \cdots \sum_{s_k=0}^{x_{ik}} (-1)^{\sum_j^k s_j} C_{R[1,\ldots,k|k+1,\ldots,n]} \times (\mathbb{P}(X_1 \leq x_{i1} - s_1), \ldots, \mathbb{P}(X_k \leq x_{ik} - s_k)). \tag{24}$$

The parameters of the Normal copula that maximize (23), computed with the built-in MATLAB command, are shown in Table 1. The computation was heavy where with an Intel(R) Core i5-3470 3.2 GHz processor and 8 GB RAM, the parameter estimation took approximately 18 h with the correlations accuracy set to 10^{-5}.

The log-likelihood of this model is −9475.9. The Normal copula recovers the continuous part of the model well with the PIT Test's p-value of 0.7729. However, comparing the observed and predicted frequencies in the 32 cells contingency table for the five binary variables with the Kullback–Leibler divergence test (KL test), a p-value of 0.0030 is obtained. This indicates that the discrete part is not recovered well.

The Normal copula model with parameters as in Table 1 is a saturated model where the correlations between all pairs are considered. However, it is observed that some of the estimates are small suggesting independencies, for instance between variables *CT* and *WM*. In principal, confidence intervals can be computed via simulation, but this is not feasible due to the very slow parameters estimation. Therefore, we simplify the model by including some conditional independencies implied by the structure of a Bayesian Network (BN) for the latency time model. This structure has been found in Zilko et al. (2015) by performing the hill-climbing greedy search in the space of all possible BN structures on the discretized version of the model where the continuous variables were discretized into four states each based on the algorithm presented in Margaritis (2003). The chosen structure is displayed in Fig. 8.

Table 1
The parameters of the Normal copula.

Parameter	Value	Parameter	Value	Parameter	Value
$\rho_{CT,WD}$	0.1909	$\rho_{WD,RH}$	−0.0717	$\rho_{WT,RH}$	0.5762
$\rho_{CT,LC}$	−0.3009	$\rho_{WD,OV}$	0.2092	$\rho_{WT,OV}$	0.1152
$\rho_{CT,WT}$	−0.1259	$\rho_{WD,LAT}$	0.1282	$\rho_{WT,LAT}$	−0.1404
$\rho_{CT,WM}$	0.0131	$\rho_{LC,WT}$	0.0529	$\rho_{WM,RH}$	0.1349
$\rho_{CT,RH}$	−0.0733	$\rho_{LC,WM}$	0.0209	$\rho_{WM,OV}$	0.3571
$\rho_{CT,OV}$	0.0387	$\rho_{LC,RH}$	0.0240	$\rho_{WM,LAT}$	0.0741
$\rho_{CT,LAT}$	−0.1078	$\rho_{LC,OV}$	0.0183	$\rho_{RH,OV}$	0.1574
$\rho_{WD,LC}$	−0.0215	$\rho_{LC,LAT}$	0.1087	$\rho_{RH,LAT}$	−0.0304
$\rho_{WD,WT}$	−0.0802	$\rho_{WT,WM}$	0.0616	$\rho_{OV,LAT}$	0.1446
$\rho_{WD,WM}$	0.0855				

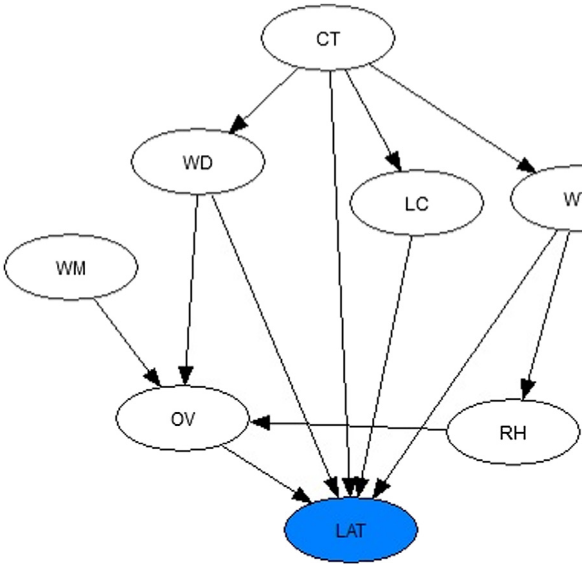


Fig. 8. The Bayesian Network model of Latency Time.

The variable *WM* is independent of *CT*, *WD*, *LC*, *WT*, and *RH*, as indicated by the graph. Moreover, *WD* and *WT* are conditionally independent given *CT*. The log-likelihood of this model is −9482.7. The model recovers the continuous part well with the PIT test's *p*-value of 0.3239. However, the discrete part is not recovered well as indicated by the KL Test's *p*-value of 0.0005. In contrast to the saturated Normal copula model which has 28 parameters, the simplified model has only 18 parameters and the likelihood has not decreased much. Performing the Likelihood Ratio Test applied to the two models yields *p*-value of 0.1920. This indicates that the BN model is a better model for the data.[1]

3.1.2. Copula-vine models

First of all for the copula-vine models, a vine structure needs to be chosen. With 8 nodes, there are $3.426e+16$ possible vines (Morales Napoles, 2009). We choose to keep the purely discrete and purely continuous parts of the model as the sub-vines of the full vine hence clustering the continuous and discrete variables. The discrete and continuous variables are ordered as: *CT*, *WM*, *WT*, *RH*, *OV* and *WD*, *LC*, *LAT*, respectively. Even after clustering the continuous and discrete variables in different sub-vines, there are still 64 different ways to merge these vines (Cooke et al., 2015). We choose the decomposition of the joint that is known as one type of vine structure that is called the *D*-vine (Fig. 9). The choice is motivated by the observed correlations between pairs of variables in the data and graphical simplicity of a D-vine model. Other choices could be considered (some heuristic procedures to choose a vine structure for the data have been investigated in Kurowicka and Joe, 2011).

A regular vine on n variables is a graphical structure. It is a nested set of $n-1$ trees with nodes and edges. Nodes in tree j become edges in tree $j+1$ where $j = 1, \ldots, n-1$, and two edges in tree j can be connected by an edge in tree $j+1$ if they share a common node in tree j. In Fig. 9, the eight variables (numbers have been assigned to the variables to simplify the notation) are represented in the D-vine as nodes. The five discrete variables are shown as white rectangles

[1] Because a semi-parametric model is estimated where the margins are modeled with empirical distribution functions, the asymptotic theory of the likelihood ratio test, in principle, does not apply.

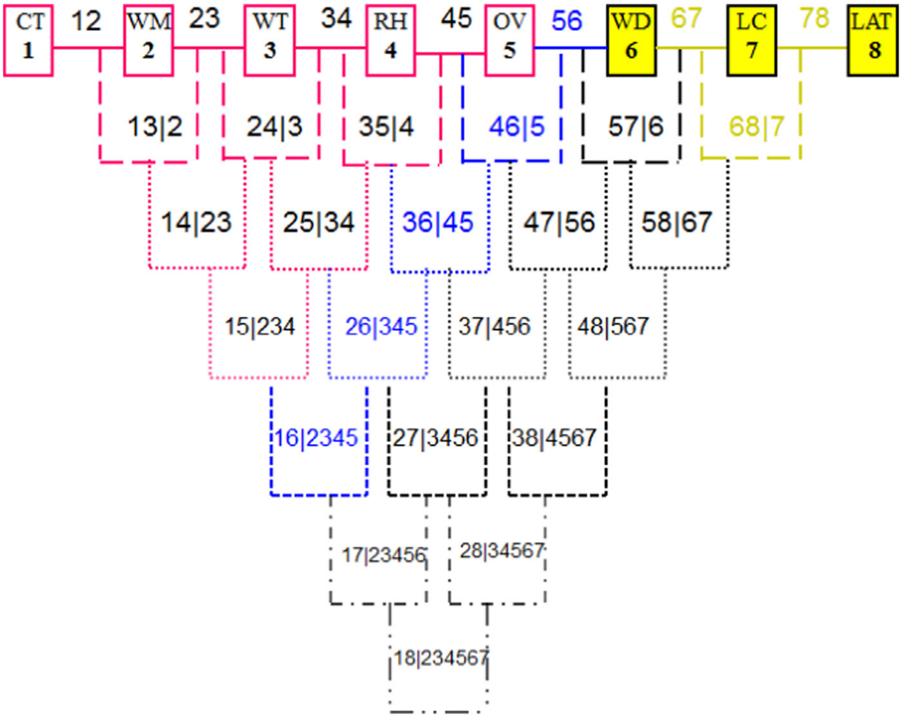


Fig. 9. The chosen D-vine structure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

while the three continuous variables as yellow rectangles. The edges in different trees are represented with different line styles. The red edges of the D-vine correspond to the purely discrete part of the model, the yellow lines correspond to the purely continuous part, and the rest represents the mixed pairs. The edges of the vine indicate the conditioned|conditioning variables and they correspond to the bivariate (conditional) copulas needed in the model.

The distribution of $(CT, WM, WT, RH, OV, WD, LC, LAT) = (X_1, \dots, X_8)$ is decomposed using factorization (14) and (15) that correspond to the D-vine structure in Fig. 9. In this decomposition, the probabilities in (15) contain two discrete, two continuous, or discrete and continuous variables in the conditioned set. This corresponds to the red, yellow, and blue and black edges in Fig. 9. When both conditioned variables are discrete, their conditional probabilities are calculated as in Section 1.3. When one is continuous and the other discrete, their conditional probability is given by (21). For both continuous variables this probability becomes the conditional copula density.

Estimation.

We consider M realizations of the eight variables in the D-vine model where the ith realization is denoted as $(x_{i1}, \dots, x_{i8})$. As in the Normal copula model, the continuous variables are represented with their empirical distribution functions. The likelihood is

$$\ell(R) = \sum_{i=1}^{M} \log\left(\mathbb{P}(X_1 = x_{i1}, \dots, X_5 = x_{i5})\mathbb{P}(X_6 = x_{i6}|X_1, \dots, X_5)\mathbb{P}(X_7 = x_{i7}|X_1, \dots, X_6)\right.$$
$$\left. \times\, \mathbb{P}(X_8 = x_{i8}|X_1, \dots, X_7)\right), \quad (25)$$

where R is the set of parameters of the vine model containing (conditional) correlations of Normal copula assigned to the edges of the vine in Fig. 9. $\mathbb{P}(X_1 = x_{i1}, \dots, X_5 = x_{i5})$ is calculated with Eqs. (15), (16), and (17) as in Section 1.3, $\mathbb{P}(X_6 = x_{i6}|X_1, \dots, X_5)$ contains the product of factors as in Eq. (22), and the last two parts of the decomposition contain factors as in Eq. (22) and the copula densities when both conditioned variables are continuous.

All copulas are assumed to be Normal. Estimation of the copulas' parameters is performed sequentially starting from the unconditional copulas of (CT, WM), (WM, WT), (WT, RH), (RH, OV), (OV, WD), (WD, LC), and (LC, LAT) in the first tree at the top of the D-vine structure in Fig. 9. Then, we estimate the copulas' parameters in the second tree which contains pairs with one variable in the conditioning set. After that, we estimate the third tree and so on.

To check the significance of the parameters, for each parameter we find its 95% confidence interval using simulation with 1000 samples. If zero is contained in the confidence interval, the parameter value of the copula is set to be equal to 0. This procedure is applied in both the constant and non-constant copula-vine models which follow shortly. In Appendix C, we test the procedure on two artificial data sets whose true parameters values are known. It is shown that the procedure works well as the true values of the parameters are captured in the corresponding confidence intervals.

Table 2
The parameters of the bivariate constant (conditional) copulas. If the corresponding confidence bound includes zero, the parameter is taken to be zero (shown in bold in a bracket).

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
$\rho_{CT,WM}$	0.0158 (**0**) (−0.2957, 0.3138)	$\rho_{WM,OV\mid WT,RH}$	0.3851 (0.1739, 0.4952)
$\rho_{WM,WT}$	0.0540 (**0**) (−0.2446, 0.3794)	$\rho_{WT,WD\mid RH,OV}$	−0.0683 (−0.1309, −0.0102)
$\rho_{WT,RH}$	0.5769 (0.3932, 0.7187)	$\rho_{RH,LC\mid OV,WD}$	0.0203 (**0**) (−0.0465, 0.0635)
$\rho_{RH,OV}$	0.1288 (**0**) (−0.2880, 0.3702)	$\rho_{OV,LAT\mid WD,LC}$	0.1141 (0.0553, 0.2242)
$\rho_{OV,WD}$	0.2035 (0.0143, 0.3571)	$\rho_{CT,OV\mid WM,WT,RH}$	−0.4344 (**0**) (−0.9887, 0.0259)
$\rho_{WD,LC}$	−0.0080 (**0**) (−0.0604, 0.0572)	$\rho_{WM,WD\mid WT,RH,OV}$	0.0152 (**0**) (−0.0633, 0.1143)
$\rho_{LC,LAT}$	0.0876 (0.0179, 0.1554)	$\rho_{WT,LC\mid RH,OV,WD}$	0.0515 (**0**) (−0.0329, 0.1273)
$\rho_{CT,WT\mid WM}$	−0.1213 (−0.1724, −0.0692)	$\rho_{RH,LAT\mid OV,WD,LC}$	−0.0279 (**0**) (−0.1147, 0.0391)
$\rho_{WM,RH\mid WT}$	−0.1268 (−0.2912, −0.0067)	$\rho_{CT,WD\mid WM,WT,RH,OV}$	0.1761 (0.0891, 0.2374)
$\rho_{WT,OV\mid RH}$	−0.1880 (−0.2847, −0.0716)	$\rho_{WM,LC\mid WT,RH,OV,WD}$	0.0120 (**0**) (−0.0643, 0.0926)
$\rho_{RH,WD\mid OV}$	−0.1002 (−0.1957, −0.0314)	$\rho_{WT,LAT\mid RH,OV,WD,LC}$	−0.1523 (−0.2187, −0.0899)
$\rho_{OV,LC\mid WD}$	0.0045 (**0**) (−0.0680, 0.0756)	$\rho_{CT,LC\mid WM,WT,RH,OV,WD}$	−0.2830 (−0.3374, −0.2199)
$\rho_{WD,LAT\mid LC}$	0.1154 (0.0440, 0.1837)	$\rho_{WM,LAT\mid WT,RH,OV,WD,LC}$	0.0413 (**0**) (−0.0255, 0.1145)
$\rho_{CT,RH\mid WM,WT}$	0.2050 (0.0074, 0.3635)	$\rho_{CT,LAT\mid WM,WT,RH,OV,WD,LC}$	−0.1279 (−0.1874, −0.0393)

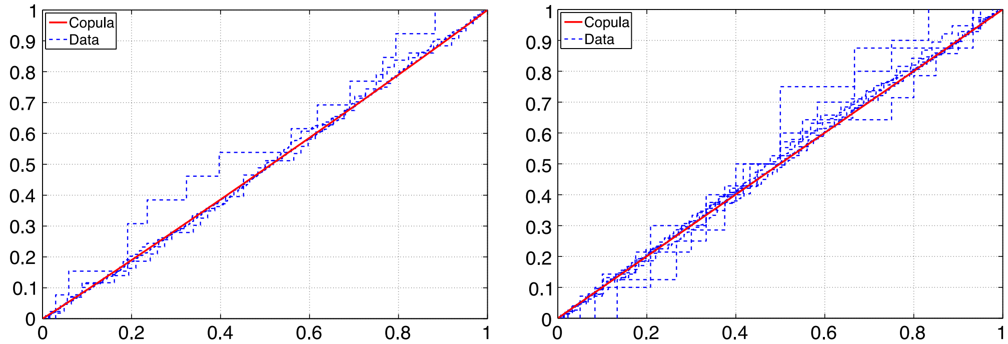
Constant copula-vine model.

In the constant copula-vine model, each pair in the second tree and higher is modeled with one conditional copula. Applying the above procedure, we get a constant conditional Normal copula vine with parameters presented in Table 2.

The procedure simplifies the model by introducing some (conditional) independencies between the variables. We see the variables WM and WT can be modeled independently and the variables OV and LC given WD can be modeled with Normal copula with parameter 0. The log-likelihood of this model is −7692.4. With this model, the discrete part is also not well recovered with the KL test's p-value of 0.0245. However, the continuous part is recovered well with the PIT test's p-value of 0.5582.

In this model, one Normal copula is used to represent a pair of variables regardless the values of their conditioning variables. Fig. 10 shows the conditional distribution of $WD|WT = 0, RH, OV$ and $WD|WM = 0, WT, RH, OV$ for different combinations of variables WT, RH, and OV. The dashed blue lines represent the empirical conditional distributions given different values in each corresponding conditioning set and the solid red lines correspond to the fitted constant Normal copula-vine model. It appears that the solid red lines do not deviate a lot from the dashed blue lines, indicating that the constant Normal copula actually represents the data quite well.

So far the parameters of the D-vine model have been estimated sequentially. This estimation is fast and takes a few minutes to complete. The next step is to perform a full optimization of all parameters in the model together. Only parameters that are not zero after the sequential fitting are considered in the full optimization. This approach is certainly computationally expensive. It took about 12 h to find the parameters of the model with a computer on an Intel(R) Core i5-3470 3.2 GHz processor and 8 GB RAM. The log-likelihood improves to −7476.5. The discrete part is still not well recovered with the KL test's p-value of 0.0329. However, the continuous part is recovered well with PIT test's p-value of 0.8224.



(a) The conditional distribution of $(WD|WT = 0, RH, OV)$.

(b) The conditional distribution of $(WD|WM = 0, WT, RH, OV)$.

Fig. 10. Empirical and constant copula based conditional distributions of variable *WD* given two conditioning sets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Non-constant copula-vine model.

In the non-constant copula-vine model, all pairs where the conditioning sets contain only discrete variables are modeled with non-constant conditional copulas (which correspond to the red and blue edges of the D-vine in Fig. 9). This means that additionally to the non-constant copulas for the purely discrete part of the model, the copulas of the following variables are assumed to depend on the values of the conditioning variables as well: $(RH, WD|OV)$, $(WT, WD|RH, OV)$, $(WM, WD|WT, RH, OV)$, and $(CT, WD|WM, WT, RH, OV)$. The other eleven conditional copulas with at least one continuous variable in the conditioning set are assumed to be constant. Other orderings of variables could be considered in which more conditional copulas would be non-constant. One such structure is presented in Appendix D.

As before, the significance of the parameters is tested by means of simulation. For the pairs where constant conditional copulas are assumed, the procedure is as in the constant copula-vine model. For the other pairs that are modeled with non-constant conditional copulas, the parameters' estimates are calculated for each value of the conditioning variables and the 95% confidence bounds are computed via simulation each with 1000 samples. If the confidence bounds of all the estimates within a pair overlap, the conditional copula of the pair is set to be constant. If additionally the overlapping part contains zero, the parameter is set to 0.

Table 3 shows the Normal copulas' parameters of $(WT, RH|OV)$ and $(CT, WD|WM, WT, RH, OV)$ where two and sixteen, respectively, conditional copulas are estimated. Moreover the *p*-value of the KS test of the conditional distributions of *WD* for different combinations of conditioning variables are presented as well. We observe that the conditional distributions of the Normal copulas fit the data well.

The conditional probability $\mathbb{P}(RH, WD|OV)$ can be modeled with constant conditional copula with parameter −0.1002 because the confidence bounds of $\rho_{RH,WD|OV=0}$ and $\rho_{RH,WD|OV=1}$ overlap. This indicates weak dependence between $RH|OV$ and $WD|OV$. Conditioning on $WM = 0, WT = 1, RH = 1, OV = 1$, there are only 7 samples of *WD* in the data set. With very few data, unsurprisingly the parameter's confidence bound is very wide and contains zero so that *CT* and *WT* given $WM = 0, WT = 1, RH = 1, OV = 1$ can be modeled with the independent copula.

Fig. 11 presents an illustration of the pair $(WT, RH|OV)$ where the dashed blue lines correspond to the empirical conditional distributions and the red solid lines correspond to the fitted copulas for the variables $WD|RH = 0, OV = 0$ and $WD|RH = 0, OV = 1$. The conditional distribution of $WD|RH = 0, OV = 0$ shown in Fig. 11(a) with the solid red line is obtained from the copula model and the dashed blue line is the empirical one obtained from the 1389 samples. The red and blue lines appear to be very close to each other and close to the diagonal, indicating the good fit of Normal copula (confirmed with the KS-test *p*-value of 0.4210) and their closeness to independence.

The entire parameters of the non-constant copula-vine model along with their confidence bounds are presented in Table C.1. We notice that many parameters are set to zero. We observe that the parameters of the Normal copulas for some pairs conditioned on different values of conditioning variables can differ quite significantly as well. For instance, the parameters of the four Normal copulas fitted to the pair $(CT, RH|WM, WT)$ are $\rho_{CT,RH|WM=0,WT=0} = -0.1860$, $\rho_{CT,RH|WM=0,WT=1} = 0$, $\rho_{CT,RH|WM=1,WT=0} = 0.4710$ and $\rho_{CT,RH|WM=1,WT=1} = 0$.

The log-likelihood of this model is −7644.1. The discrete part is recovered well with *p*-value of the KL test to be 0.2155. The continuous part of this model is the same as in the constant copula-vine model where it is already observed to be well-recovered.

Similarly as in the constant copula-vine approach, full optimization via maximum likelihood approach can also be performed to estimate the parameters. With this approach, the log-likelihood improves to −7344.7. The discrete part is recovered with the KL test's *p*-value of 0.2237 and the continuous part is recovered well with the PIT test's *p*-value of 0.6390.

Performing the Likelihood Ratio Test, the *p*-values are very close to 0 for both the sequential and full-optimized approaches between the constant and non-constant copula-vine models. This indicates that the non-constant copula-vine model provides better fit for the data.

Table 3
Parameters of the bivariate (conditional) normal copulas of $(RH, WD|OV)$, $(WM, WD|WT, RH, OV)$ along with the p-values of the conditional distribution of WD for different combinations of the conditioning variables for the copula model and the data.

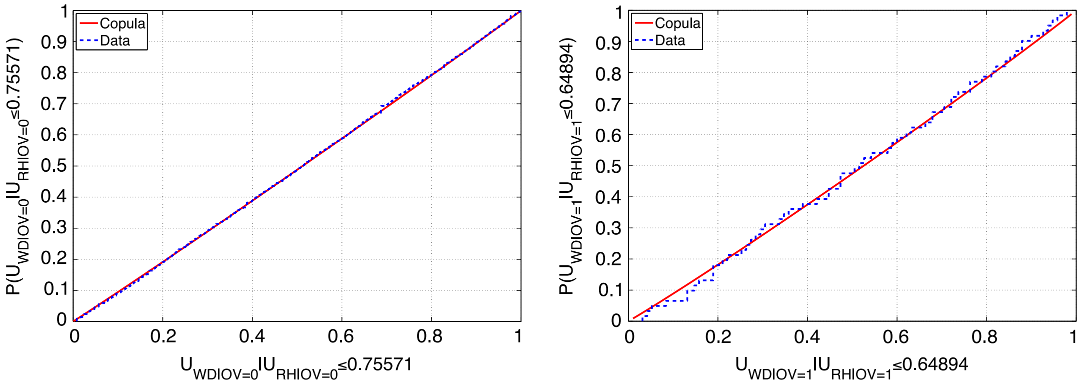
Parameter	Value (Conf. bound)	# Sam	Conditional distribution	p-value
$\rho_{RH,WD\|OV=0}$	−0.0738 (−**0.1002**) (−0.1300, −0.0209)	1389	$F_{WD\|RH=0,OV=0}$	0.4210
$\rho_{RH,WD\|OV=1}$	−0.1213 (−**0.1002**) (−0.1567, −0.0656)	61	$F_{WD\|RH=0,OV=1}$	0.9964
$\rho_{CT,WD\|WM=0,WT=0,RH=0,OV=0}$	0.2444 (0.1635, 0.3111)	431	$F_{WD\|CT=0,WM=0,WT=0,RH=0,OV=0}$	0.2051
$\rho_{CT,WD\|WM=0,WT=0,RH=0,OV=1}$	0.1546 (**0**) (−0.2278, 0.5959)	14	$F_{WD\|CT=0,WM=0,WT=0,RH=0,OV=1}$	0.5824
$\rho_{CT,WD\|WM=0,WT=0,RH=1,OV=0}$	0.2118 (0.0462, 0.4135)	75	$F_{WD\|CT=0,WM=0,WT=0,RH=1,OV=0}$	0.7046
$\rho_{CT,WD\|WM=0,WT=1,RH=0,OV=0}$	0.0965 (**0**) (−0.0491, 0.2421)	152	$F_{WD\|CT=0,WM=0,WT=1,RH=0,OV=0}$	0.6232
$\rho_{CT,WD\|WM=1,WT=0,RH=0,OV=0}$	0.1094 (**0**) (−0.3646, 0.5012)	25	$F_{WD\|CT=0,WM=1,WT=0,RH=0,OV=0}$	0.8916
$\rho_{CT,WD\|WM=1,WT=1,RH=0,OV=0}$	−0.1013 (**0**) (−0.6242, 0.4463)	12	$F_{WD\|CT=0,WM=1,WT=1,RH=0,OV=0}$	0.8323
$\rho_{CT,WD\|WM=1,WT=0,RH=1,OV=0}$	0.2608 (**0**) (−0.3939, 0.8685)	6	$F_{WD\|CT=0,WM=1,WT=0,RH=1,OV=0}$	0.9989
$\rho_{CT,WD\|WM=1,WT=0,RH=0,OV=1}$	0.0012 (**0**) (−0.0720, 0.0995)	3	$F_{WD\|CT=0,WM=1,WT=0,RH=0,OV=1}$	0.8833
$\rho_{CT,WD\|WM=0,WT=1,RH=1,OV=0}$	0.0747 (**0**) (−0.0569, 0.2192)	137	$F_{WD\|CT=0,WM=0,WT=1,RH=1,OV=0}$	0.7577
$\rho_{CT,WD\|WM=0,WT=1,RH=0,OV=1}$	0.2528 (**0**) (−0.5006, 0.7496)	6	$F_{WD\|CT=0,WM=0,WT=1,RH=0,OV=1}$	0.8740
$\rho_{CT,WD\|WM=0,WT=0,RH=1,OV=1}$	0.0021 (**0**) (−0.7106, 0.3661)	7	$F_{WD\|CT=0,WM=0,WT=0,RH=1,OV=1}$	0.9971
$\rho_{CT,WD\|WM=1,WT=1,RH=1,OV=0}$	−0.9891 (**0**) (−0.9991, 0.9921)	1	$F_{WD\|CT=0,WM=1,WT=1,RH=1,OV=0}$	0.3489
$\rho_{CT,WD\|WM=1,WT=1,RH=0,OV=1}$	−0.8765 (**0**) (−0.9451, 0.9871)	1	$F_{WD\|CT=0,WM=1,WT=1,RH=0,OV=1}$	0.3353
$\rho_{CT,WD\|WM=1,WT=0,RH=1,OV=1}$	0.7518 (**0**) (−0.7714, 0.8989)	2	$F_{WD\|CT=0,WM=1,WT=0,RH=1,OV=1}$	0.7754
$\rho_{CT,WD\|WM=0,WT=1,RH=1,OV=1}$	−0.5228 (**0**) (−0.9708, 0.3070)	7	$F_{WD\|CT=0,WM=0,WT=1,RH=1,OV=1}$	0.4689
$\rho_{CT,WD\|WM=1,WT=1,RH=1,OV=1}$	0.0039 (**0**) (−0.9990, 0.9819)	3	$F_{WD\|CT=0,WM=1,WT=1,RH=1,OV=1}$	0.7055

3.2. Models validation

In this section, we compare the performance of different Latency Time models that we have constructed:

1. Eight dimensional Normal copula.
2. Eight dimensional Normal copula with conditional independence statements represented by the BN structure in Fig. 8.
3. The copula-vine approach with constant conditional Normal copula.
4. The copula-vine approach with non-constant conditional Normal copula.

The Akaike Information Criterion (AIC) is used as a criteria for model validation. Moreover, a validation test is performed where for each data point, the quantile of the conditional distribution of latency time corresponding to the observed latency



(a) The conditional distribution of $(WD|RH = 0, OV = 0)$ with 1389 samples from data.

(b) The conditional distribution of $(WD|RH = 0, OV = 1)$ with 61 samples from data.

Fig. 11. Empirical and non-constant copula based conditional distributions of variable *WD* for different combinations of conditioning variables. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
The log-likelihood, the number of parameters, the AIC score, and the *p*-value of the KS-test of each of the four models with the training data.

Model	Log-likelihood	# parameters	AIC score	*p*-value
Normal copula	−9475.9	28	19 007.8	0.8886
Normal copula with conditional independence	−9482.7	18	19 001.4	0.8808
Constant copula-vine (sequential)	−7692.4	16	15 416.8	0.2152
Constant copula-vine (full optimization)	−7476.5	16	14 985.0	0.2247
Non-constant copula-vine (sequential)	−7644.1	24	15 336.2	0.2372
Non-constant copula-vine (full optimization)	−7344.7	24	14 737.4	0.2508

Table 5
The RMSE of the four models when the mean of the conditioned latency time is used as the prediction.

Model	RMSE
Normal copula	29.0023
Normal copula with conditional independence	29.1309
Constant copula-vine (sequential)	29.5089
Constant copula-vine (full optimization)	29.5975
Non-constant copula-vine (sequential)	29.1893
Non-constant copula-vine (full optimization)	29.3628

time is computed. The model represents the data well if these quantiles are distributed uniformly on (0, 1). We use the KS Test to test the closeness of the quantiles of each model to uniform.

The results of both validation tests are summarized in Table 4.

The second and third columns of Table 4 present the log-likelihood and the number of parameters in each model. The AIC score of each model is shown in the fourth column. The quantiles are tested against the uniform distribution in (0, 1) with the KS-Test and the *p*-values are presented in the fifth column.

The results show that all models perform well in the prediction of the latency time's conditional distribution. The non-constant copula-vine model has the highest log-likelihood value as was to be expected. It also has the lowest AIC score which means it is the best model to represent the data.

The aim of the latency time model is to provide a prediction of the latency time when an incident occurs, given the information of the other seven variables. One candidate for the prediction is the expected value of the conditional latency time. For each data point, the conditional distribution of latency time given the realization of the other variables is simulated using 500 samples.

To obtain samples of the conditional latency time distributions of all models, samples are generated from the intervals of the latent variables corresponding to the discrete observations. For the multivariate Normal copula model, the copula is conditioned on these samples and the observed continuous variables. For the copula-vine models, the vine structure conditioned on the latent samples and the observed continuous variables is sampled as in Kurowicka and Cooke (2006) to obtain the conditional latency time distribution.

For each model, the expected value of this sampled distribution is compared to the observed latency time in the data to obtain the root mean squared error (RMSE) for each model. Table 5 summarizes the result.

Table 6
The log-likelihood, the AIC score, the p-value of the KS-test, and the RMSE of each of the four models with the test data.

Model	Log-likelihood	AIC score	p-value	RMSE
Normal copula	−1852.6	3761.2	0.1755	19.8789
Normal copula with conditional independence	−1849.5	3735.1	0.1038	19.9068
Constant copula-vine (sequential)	−1615.2	3262.4	0.1444	19.9165
Constant copula-vine (full optimization)	−1613.8	3259.6	0.1196	20.0192
Non-constant copula-vine (sequential)	−1589.4	3226.8	0.1755	19.7433
Non-constant copula-vine (full optimization)	−1529.4	3106.8	0.1202	19.6875

The result shows that the performance of all four models is very similar, not surprising considering the result shown in Table 4. However, the standard deviation of the latency time in the data is observed at 30.2128. This means that the coefficient of determination value R^2 is approximately only 5% for all models. This illustrates the highly uncertain and complex nature of railway disruption length, in particular the latency time, in the Netherlands. Additional influencing factors could be considered when new, better data is collected.

While Tables 4 and 5 present the results of models validation with the training set, we are also interested to observe the model performance against a set of test data. The test data comes from the train detection problem in the Dutch railway network from 1 May 2014 up to 31 October 2014. A total of 339 urgent incidents were recorded within these six months period. Table 6 summarizes the result.

The result shows that all models perform well with similar performance. The non-constant copula vine model yields the highest log-likelihood and the lowest AIC score hence it is the best model to represent the data. Similarly as in the training set case, with the latency time's standard deviation of 20.3546, the coefficient of determination is approximately only 5% for all models.

4. Discussions and conclusions

Four copula-based models are constructed in this paper to represent the dependence between an eight-dimensional mixed discrete and continuous variable. These models are based on two construction techniques: through the multivariate Normal copula and through the copula-vine approach. When the multivariate Normal copula is used, it is not easy to give conditions which assert the existence or non-existence of the Normal copula for a specified Bernoulli distribution. With the copula-vine approach, it is shown that the multivariate Bernoulli part of the model can be recovered with a set of non-constant (conditional) bivariate Normal copulas.

The performance of the four models is similar. In practice, the chosen model is aimed to produce the conditional distribution of latency time given information about the other influencing variables. The conditional distribution is obtained through sampling for all models. Generating 500 samples of this conditional distribution with a computer on an Intel(R) Core i5-3470 3.2 GHz processor and 8 GB RAM, the multivariate Normal copula models require, on average, 3.0128 s while the copula-vine models require 3.6684 s on average. This means that the non-constant copula-vine model is the most attractive one because, while the time needed to obtain the conditional latency time distribution is not very different between the models, the parameters estimation can be done much faster when they are performed sequentially.

The latency time model constitutes only one part of the disruption length model. The full disruption length model is going to be constructed by considering two more variables: the cause of failures and the repair time.

In this paper, we take the expected value of the conditioned latency time as a prediction for the length of latency time needed when a disruption occurs. However, even though this choice leads to the minimum RMSE, it might not be the most desired in practice. The aim of the disruption length model is to provide a prediction for the train traffic controllers to deal with the disrupted train traffic caused by an incident. When the traffic controllers have information about possible disruption length, they will take certain steps to reschedule the train traffic. The goal is to return to normal train operation as soon as possible. Thus, the choice of prediction other than the expected disruption length could be of larger benefit to achieve the final goal. We plan to investigate this issue in more detail when we combine our model with a parallel project dealing with effective and efficient train dispatching.

Acknowledgments

This work is funded by ExploRail, a partnership program of the Dutch Technology Foundation STW and ProRail, project no. 12257: "Smart information and decision support for railway operation control centres (SmartOCCR)". The authors would like to acknowledge Dirk Kes, André Duinmeijer, and Erwin van Wonderen from ProRail for the supervision and access to the data set.

Appendix A. Proof of Propositions 1.3 and 1.5

Proof of Proposition 1.3. The zero three-way interaction means:

$$p(1,1,1)p(1,0,0)p(0,1,0)p(0,0,1) = p(0,0,0)p(1,1,0)p(1,0,1)p(0,1,1). \tag{A.1}$$

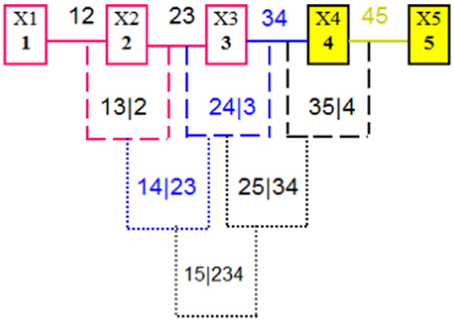


Fig. B.12. The D-vine structure.

Since the margins are equal to 0.5, we obtain:

$$p(0,1,0)+p(0,0,1)+p(0,1,1)+p(0,0,0)=p(1,0,1)+p(1,1,0)+p(1,0,0)+p(1,1,1) \tag{A.2}$$
$$p(1,0,0)+p(0,0,1)+p(1,0,1)+p(0,0,0)=p(0,1,1)+p(1,1,0)+p(0,1,0)+p(1,1,1) \tag{A.3}$$
$$p(1,0,0)+p(0,1,0)+p(1,1,0)+p(0,0,0)=p(0,1,1)+p(1,0,1)+p(0,0,1)+p(1,1,1). \tag{A.4}$$

Subtracting (A.3) from (A.2), (A.4) from (A.3), and (A.2) from (A.4) yields

$$p(0,1,0)+p(0,1,1)=p(1,0,1)+p(1,0,0) \tag{A.5}$$
$$p(0,0,1)+p(1,0,1)=p(1,1,0)+p(0,1,0) \tag{A.6}$$
$$p(1,1,0)+p(1,0,0)=p(0,0,1)+p(0,1,1). \tag{A.7}$$

Substituting (A.5) to (A.2) yields

$$p(0,0,0)+p(0,0,1)=p(1,1,1)+p(1,1,0). \tag{A.8}$$

It can be shown that Eqs. (A.1), (A.5), (A.6), and (A.7) are satisfied if and only if $p(x_1,x_2,x_3)=p(1-x_1,1-x_2,1-x_3)$ for all $x_1,x_2,x_3\in\{0,1\}$. We see immediately that the symmetric distribution satisfies the above equations. To see that the symmetry is also necessary, let us assume that e.g. $p(1,0,0)>p(0,1,1)$. Then, from (A.5), $p(0,1,0)>p(1,0,1)$ which leads to $p(0,0,1)>p(1,1,0)$ from (A.6). Further, this means $p(0,0,1)>p(1,1,0)$ from (A.7) which leads to $p(1,1,1)>p(0,0,0)$ from (A.8). Combining this information together yields

$$p(1,1,1)p(1,0,0)p(0,1,0)p(0,0,1)>p(0,0,0)p(1,1,0)p(1,0,1)p(0,1,1)$$

which cannot be true because of (A.1). Therefore, the proof is complete. □

Proof of Proposition 1.5. Without loss of generality, let X_2 be the conditioning variable. Because $\mathbb{P}(X_i=0)=0.5=P(X_i=1)$ for all $i\in\{1,2,3\}$,

$$\mathbb{P}(X_1\le 0|X_2=1)=1-\mathbb{P}(X_1\le 0|X_2=0) \tag{A.9}$$

and

$$\mathbb{P}(X_1\le 0,X_3\le 0|X_2=1)=\frac{\mathbb{P}(X_1\le 0,X_3\le 0)}{0.5}-\mathbb{P}(X_1\le 0,X_3\le 0|X_2=0)$$
$$C_{13|2=1}(\mathbb{P}(X_1\le 0|X_2=1),\mathbb{P}(X_3\le 0|X_2=1))=\frac{\mathbb{P}(X_1\le 0,X_3\le 0)}{0.5}-C_{13|2=0}(\mathbb{P}(X_1\le 0|X_2=0),\mathbb{P}(X_3\le 0|X_2=0)). \tag{A.10}$$

Using (A.9), the symmetricity of $C_{13|2}$, and the Total Law of Probability, the left hand side of (A.10) becomes:

$$C_{13|2=1}(\mathbb{P}(X_1\le 0|X_2=1),\mathbb{P}(X_3\le 0|X_2=1))=\frac{\mathbb{P}(X_1\le 0,X_2=0,X_3\le 0)+\mathbb{P}(X_1>0,X_2=0,X_3>0)}{0.5}-C_{13|2=1}(\mathbb{P}(X_1\le 0|X_2=0),\mathbb{P}(X_3\le 0|X_2=0)). \tag{A.11}$$

⇒ With the radial symmetry of a trivariate Normal copula, (A.11) becomes:

$$=\frac{\mathbb{P}(X_1\le 0,X_2=0,X_3\le 0)+\mathbb{P}(X_1\le 0,X_2=1,X_3\le 0)}{0.5}-C_{13|2=1}(\mathbb{P}(X_1\le 0|X_2=0),\mathbb{P}(X_3\le 0|X_2=0))$$
$$=\frac{\mathbb{P}(X_1\le 0,X_3\le 0)}{0.5}-C_{13|2=1}(\mathbb{P}(X_1\le 0|X_2=0),\mathbb{P}(X_3\le 0|X_2=0)).$$

Table B.7
The estimated parameters values of the constant copula-vine data set with the constant copula-vine approach.

Parameter	True value	Estimate (Conf. bound)	Parameter	True value	Estimate (Conf. bound)
ρ_{12}	0.3	0.3032 (0.1267, 0.4346)	$\rho_{24\|3}$	−0.3	−0.3033 (−0.3361, −0.2740)
ρ_{23}	0	0.0542 (**0**) (−0.1256, 0.2177)	$\rho_{35\|4}$	0.2	0.2067 (0.1364, 0.2587)
ρ_{34}	0.2	0.1880 (0.0700, 0.2975)	$\rho_{14\|23}$	0	0.0001 (**0**) (−0.0453, 0.0813)
ρ_{45}	−0.1	−0.1088 (−0.1904, −0.0336)	$\rho_{25\|34}$	−0.3	−0.2462 (−0.3237, −0.1768)
$\rho_{13\|2}$	0.11	0.0983 (0.0405, 0.1497)	$\rho_{15\|234}$	0.5	0.5174 (0.4626, 0.5740)

Table B.8
The estimated parameters values of the constant copula-vine data set with the non-constant copula-vine approach.

Parameter	True value	Estimate (Conf. bound)	Parameter	True value	Estimate (Conf. bound)
ρ_{12}	0.3	0.3032 (0.1267, 0.4346)	$\rho_{35\|4}$	0.2	0.2067 (0.1419, 0.2802)
ρ_{23}	0	0.0542 (**0**) (−0.1256, 0.2177)	$\rho_{14\|2=0,3=0}$		−0.0543 (**0**) (−0.1904, 0.0503)
ρ_{34}	0.2	0.1880 (0.0700, 0.2975)	$\rho_{14\|2=0,3=1}$	0	0.1344 (**0**) (−0.0241, 0.2990)
ρ_{45}	−0.1	−0.1088 (−0.1904, −0.0336)	$\rho_{14\|2=1,3=0}$		0.0083 (**0**) (−0.0711, 0.0928)
$\rho_{13\|2=0}$	0.11	0.1488 (**0.0983**) (0.0425, 0.2665)	$\rho_{14\|2=1,3=1}$		−0.1035 (**0**) (−0.2228, 0.0513)
$\rho_{13\|2=1}$		0.0693 (**0.0983**) (−0.0248, 0.1592)	$\rho_{25\|34}$	−0.3	−0.2462 (−0.3403, −0.1786)
$\rho_{24\|3=0}$	−0.3	−0.3177 (−**0.3033**) (−0.3714, −0.2602)	$\rho_{15\|234}$	0.5	0.5174 (0.4702, 0.5731)
$\rho_{24\|3=1}$		−0.2430 (−**0.3033**) (−0.3122, −0.1871)			

Substituting this result back to (A.10) yields $C_{13|2=0}(\mathbb{P}(Y_1 \leq 0|Y_2 = 0), \mathbb{P}(Y_3 \leq 0|Y_2 = 0)) = C_{13|2=1}(\mathbb{P}(Y_1 \leq 0|Y_2 = 0), \mathbb{P}(Y_3 \leq 0|Y_2 = 0))$.

$\Leftarrow$ Because $C_{13|2=0} = C_{13|2=1}$, substituting (A.11) back to (A.10) yields

$$\frac{\mathbb{P}(X_1 \leq 0, X_2 = 0, X_3 \leq 0) + \mathbb{P}(X_1 > 0, X_2 = 0, X_3 > 0)}{0.5} = \frac{\mathbb{P}(X_1 \leq 0, X_3 \leq 0)}{0.5}$$

which leads to

$$\mathbb{P}(X_1 > 0, X_2 = 0, X_3 > 0) = \mathbb{P}(X_1 \leq 0, X_2 = 1, X_3 \leq 0).$$

Substituting this result into Eqs. (A.5), (A.6), (A.7), and (A.8) yields

$$p(x_1, x_2, x_3) = p(1 - x_1, 1 - x_2, 1 - x_3)$$

for all $x_i \in \{0, 1\}$ for all $i \in \{1, 2, 3\}$. This means the trivariate Bernoulli distribution has a radial symmetry. Therefore, the trivariate Normal copula is able to realize (X_1, X_2, X_3). $\square$

Table B.9
The estimated parameters values of the non-constant copula-vine data set with the constant copula-vine approach.

Parameter	Estimate (Conf. bound)	Parameter	Estimate (Conf. bound)
ρ_{12}	0.3489 (0.1577, 0.4978)	$\rho_{24\|3}$	−0.3028 (−0.3421, −0.2673)
ρ_{23}	−0.0042 (**0**) (−0.1489, 0.1807)	$\rho_{35\|4}$	0.2023 (0.1413, 0.2607)
ρ_{34}	0.1906 (0.0674, 0.3085)	$\rho_{14\|23}$	0.0312 (**0**) (−0.0491, 0.1126)
ρ_{45}	−0.1147 (−0.1921, −0.0389)	$\rho_{25\|34}$	−0.2954 (−0.3724, −0.2186)
$\rho_{13\|2}$	0.1027 (0.0395, 0.1534)	$\rho_{15\|234}$	0.5234 (0.4678, 0.5722)

Table B.10
The estimated parameters values of the non-constant copula-vine data set with the non-constant copula-vine approach.

Parameter	True value	Estimate (Conf. bound)	Parameter	True value	Estimate (Conf. bound)
ρ_{12}	0.3	0.3489 (0.1577, 0.4978)	$\rho_{35\|4}$	0.2	0.2023 (0.1413, 0.2607)
ρ_{23}	0	−0.0042 (**0**) (−0.1489, 0.1807)	$\rho_{14\|2=0,3=0}$	0.1	0.0405 (0.0254, 0.1540)
ρ_{34}	0.2	0.1906 (0.0674, 0.3085)	$\rho_{14\|2=0,3=1}$	0	−0.0572 (**0**) (−0.2293, 0.1153)
ρ_{45}	−0.1	−0.1147 (−0.1921, −0.0389)	$\rho_{14\|2=1,3=0}$	0.4	0.4399 (0.3722, 0.5204)
$\rho_{13\|2=0}$	0.2	0.1917 (0.0899, 0.3107)	$\rho_{14\|2=1,3=1}$	0	0.0004 (**0**) (−0.1484, 0.1174)
$\rho_{13\|2=1}$	0	−0.0096 (**0**) (−0.1180, 0.0742)	$\rho_{25\|34}$	−0.3	−0.2954 (−0.3724, −0.2186)
$\rho_{24\|3=0}$	−0.3	−0.2751 (−**0.3028**) (−0.3342, −0.2183)	$\rho_{15\|234}$	0.5	0.5422 (0.4879, 0.5982)
$\rho_{24\|3=1}$	−0.3	−0.2506 (−**0.3028**) (−0.3267, −0.1863)			

Appendix B. Simulation study

In this section, a small simulation study is performed to test the performance of the (non)-constant copula-vine models. Three binary discrete variables X_1, X_2, and X_3 with margins $\mathbb{P}(X_1 = 0) = 0.6$, $\mathbb{P}(X_2 = 0) = 0.4$, and $\mathbb{P}(X_3 = 0) = 0.7$, respectively, and two uniform continuous variables X_4 and X_5 are involved. The five variables are joint together with a D-vine with structure depicted in Fig. B.12.

Two sets of parameters values are chosen (indicated in the column True Value in the corresponding tables below) and are used to sample the vine structure to generate two data sets each containing 2000 samples. Then, the parameters are estimated with constant and non-constant copula vines as in the main body of this paper for each data set. The following are the results.

1. Constant copula-vine data set.
   The first simulated data set contains data generated with the constant copula-vine approach. Tables B.7 and B.8 showcase the results when the constant and non-constant copula-vine approaches are used to model the data's dependence.
   The true parameter values are recovered with the constant copula-vine approach. Moreover, the discrete part of the model is recovered well with the KL divergence test yields p-value of 0.2048. The continuous part is also well-recovered with p-value of 0.9104.
   The true parameter value is recovered with the non-constant conditional copula vine approach where the procedure notices that the three conditional distributions with discrete only conditioning variables can be modeled with constant

Table C.1
The parameters of the non-constant conditional copula-vine model with the first vine structure. The bolded brackets indicate if the parameter value is taken to be zero or constant.

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
$\rho_{CT,WM}$	0.0158 (**0**) (−0.2957, 0.3138)	$\rho_{CT,OV\|WM=0,WT=1,RH=1}$	−0.5118 (−0.8400, −0.1854)
$\rho_{WM,WT}$	0.0540 (**0**) (−0.2446, 0.3794)	$\rho_{CT,OV\|WM=1,WT=0,RH=1}$	0.5648 (**0**) (−0.1247, 0.9496)
$\rho_{WT,RH}$	0.5769 (0.3932, 0.7187)	$\rho_{CT,OV\|WM=1,WT=1,RH=0}$	0.7001 (0.2727, 0.9900)
$\rho_{RH,OV}$	0.1288 (**0**) (−0.2880, 0.3702)	$\rho_{CT,OV\|WM=1,WT=1,RH=1}$	−0.0105 (**0**) (−0.8197, 0.7740)
$\rho_{OV,WD}$	0.2035 (0.0143, 0.3571)	$\rho_{WM,WD\|WT=0,RH=0,OV=0}$	0.0424 (**0**) (−0.1039, 0.1650)
$\rho_{WD,LC}$	−0.0080 (**0**) (−0.0604, 0.0572)	$\rho_{WM,WD\|WT=0,RH=0,OV=1}$	0.3599 (**0**) (−0.3366, 0.8918)
$\rho_{LC,LAT}$	0.0876 (0.0179, 0.1554)	$\rho_{WM,WD\|WT=0,RH=1,OV=0}$	0.2210 (**0**) (−0.0286, 0.4438)
$\rho_{CT,WT\|WM=0}$	−0.1299 (**−0.1213**) (−0.1910, −0.0649)	$\rho_{WM,WD\|WT=1,RH=0,OV=0}$	−0.0972 (**0**) (−0.3071, 0.1023)
$\rho_{CT,WT\|WM=1}$	−0.1137 (**−0.1213**) (−0.3764, 0.1722)	$\rho_{WM,WD\|WT=0,RH=1,OV=1}$	0.2894 (**0**) (−0.6480, 0.9041)
$\rho_{WM,RH\|WT=0}$	0.3705 (0.2488, 0.4954)	$\rho_{WM,WD\|WT=1,RH=0,OV=1}$	0.8917 (**0**) (−0.0575, 0.9900)
$\rho_{WM,RH\|WT=1}$	−0.2464 (−0.4333, −0.0605)	$\rho_{WM,WD\|WT=1,RH=1,OV=0}$	0.2312 (**0**) (−0.0554, 0.5121)
$\rho_{WT,OV\|RH=0}$	0.2558 (0.0739, 0.3680)	$\rho_{WM,WD\|WT=1,RH=1,OV=1}$	−0.0916 (**0**) (−0.7500, 0.4982)
$\rho_{WT,OV\|RH=1}$	−0.2800 (−0.5151, −0.0807)	$\rho_{WT,LC\|RH,OV,WD}$	0.0514 (**0**) (−0.0331, 0.1290)
$\rho_{RH,WD\|OV=0}$	−0.0738 (**−0.1002**) (−0.1300, −0.0209)	$\rho_{RH,LAT\|OV,WD,LC}$	−0.0261 (**0**) (−0.1047, 0.0441)
$\rho_{RH,WD\|OV=1}$	−0.1213 (**−0.1002**) (−0.1567, −0.0656)	$\rho_{CT,WD\|WM=0,WT=0,RH=0,OV=0}$	0.2444 (0.1635, 0.3111)
$\rho_{OV,LC\|WD}$	0.0045 (**0**) (−0.0680, 0.0756)	$\rho_{CT,WD\|WM=0,WT=0,RH=0,OV=1}$	0.1546 (**0**) (−0.2278, 0.5959)
$\rho_{WD,LAT\|LC}$	0.1154 (−0.0440, 0.1837)	$\rho_{CT,WD\|WM=0,WT=0,RH=1,OV=0}$	0.2118 (0.0462, 0.4135)
$\rho_{CT,RH\|WM=0,WT=0}$	−0.1860 (−0.2762, −0.0559)	$\rho_{CT,WD\|WM=0,WT=1,RH=0,OV=0}$	0.0965 (**0**) (−0.0491, 0.2421)
$\rho_{CT,RH\|WM=0,WT=1}$	0.0850 (**0**) (−0.0304, 0.2271)	$\rho_{CT,WD\|WM=1,WT=0,RH=0,OV=0}$	0.1094 (**0**) (−0.3646, 0.5012)
$\rho_{CT,RH\|WM=1,WT=0}$	0.4710 (0.1671, 0.7451)	$\rho_{CT,WD\|WM=0,WT=0,RH=1,OV=1}$	0.0021 (**0**) (−0.7106, 0.3661)
$\rho_{CT,RH\|WM=1,WT=1}$	−0.3282 (**0**) (−0.6950, 0.2561)	$\rho_{CT,WD\|WM=0,WT=1,RH=0,OV=1}$	0.2528 (**0**) (−0.5006, 0.7496)
$\rho_{WM,OV\|WT=0,RH=0}$	0.1057 (**0**)	$\rho_{CT,WD\|WM=1,WT=0,RH=0,OV=1}$	0.0012 (**0**)

Table C.1 (*continued*)

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
	(−0.0002, 0.3236)		(−0.0720, 0.0995)
$\rho_{WM,OV\|WT=0,RH=1}$	0.7767 (0.4771, 0.9207)	$\rho_{CT,WD\|WM=0,WT=1,RH=1,OV=0}$	0.0747 (**0**) (−0.0569, 0.2192)
$\rho_{WM,OV\|WT=1,RH=0}$	0.7417 (0.6117, 0.8770)	$\rho_{CT,WD\|WM=1,WT=0,RH=1,OV=0}$	0.2608 (**0**) (−0.3939, 0.8685)
$\rho_{WM,OV\|WT=1,RH=1}$	0.5844 (**0**) (−0.0023, 0.8859)	$\rho_{CT,WD\|WM=1,WT=1,RH=0,OV=0}$	−0.1013 (**0**) (−0.6242, 0.4463)
$\rho_{WT,WD\|RH=0,OV=0}$	−0.1116 (**−0.0683**) (−0.1896, −0.0442)	$\rho_{CT,WD\|WM=0,WT=1,RH=1,OV=1}$	−0.5228 (**0**) (−0.9708, 0.3070)
$\rho_{WT,WD\|RH=0,OV=1}$	−0.1125 (**−0.0683**) (−0.4796, 0.2657)	$\rho_{CT,WD\|WM=1,WT=0,RH=1,OV=1}$	0.7518 (**0**) (−0.7714, 0.8989)
$\rho_{WT,WD\|RH=1,OV=0}$	−0.0174 (**−0.0683**) (−0.1251, 0.0887)	$\rho_{CT,WD\|WM=1,WT=1,RH=0,OV=1}$	0.8765 (**0**) (−0.9451, 0.9871)
$\rho_{WT,WD\|RH=1,OV=1}$	0.2840 (**−0.0683**) (−0.1238, 0.6281)	$\rho_{CT,WD\|WM=1,WT=1,RH=1,OV=0}$	−0.9891 (**0**) (−0.9991, 0.9921)
$\rho_{RH,LC\|OV,WD}$	0.0206 (**0**) (−0.0719, 0.0727)	$\rho_{CT,WD\|WM=1,WT=1,RH=1,OV=1}$	0.0039 (**0**) (−0.9990, 0.9819)
$\rho_{OV,LAT\|WD,LC}$	0.1141 (0.0582, 0.2014)	$\rho_{WM,LC\|WT,RH,OV,WD}$	0.0118 (**0**) (−0.0643, 0.0926)
$\rho_{CT,OV\|WM=0,WT=0,RH=0}$	0.1087 (**0**) (−0.0691, 0.3166)	$\rho_{WT,LAT\|RH,OV,WD,LC}$	−0.1522 (−0.2280, −0.0878)
$\rho_{CT,OV\|WM=0,WT=0,RH=1}$	−0.2696 (**0**) (−0.6787, 0.1655)	$\rho_{CT,LC\|WM,WT,RH,OV,WD}$	−0.2829 (−0.3484, −0.2172)
$\rho_{CT,OV\|WM=0,WT=1,RH=0}$	0.6394 (0.4190, 0.7696)	$\rho_{WM,LAT\|WT,RH,OV,WD,LC}$	0.0413 (**0**) (−0.0255, 0.1145)
$\rho_{CT,OV\|WM=1,WT=0,RH=0}$	−0.6379 (**0**) (−0.9036, 0.0017)	$\rho_{CT,LAT\|WM,WT,RH,OV,WD,LC}$	−0.1227 (−0.1967, −0.0373)

Table D.2
The log-likelihood, the number of parameters, the AIC score, and the *p*-value of the KS-test of the two vine structures. The result of the first structure is presented in brackets.

Model	Log-likelihood	# parameters	AIC score	*p*-value
Constant, sequential	−7617.2 (−7692.4)	20 (16)	15 274.4 (15 416.8)	0.2384 (0.2152)
Constant, full optimization	−7489.9 (−7476.5)	20 (16)	15 019.8 (14 985.0)	0.2501 (0.2247)
Non-constant, sequential	−7593.1 (−7644.1)	28 (24)	15 242.2 (15 336.2)	0.2417 (0.2372)
Non-constant, full optimization	−7309.7 (−7344.7)	28 (24)	14 675.4 (14 737.4)	0.2717 (0.2508)

conditional copula. Also, the discrete part of the model is recovered well with this approach with *p*-value of 0.2048. As in the constant case, the continuous part is also well-recovered with the same *p*-value.

2. Non-constant copula-vine data set.

   The second simulated data set contains data generated with the non-constant copula-vine approach. Tables B.9 and B.10 showcase the results when the constant and non-constant copula-vine approaches are used to model the data's dependence.

   With the constant copula-vine approach, the discrete part of the model is not recovered because the data is generated with non-constant copula-vines. The KL divergence test yields *p*-value of 0.0156. The parameter $\rho_{13|2}$ is taken to be constant while the data is generated with non-constant conditional copula of $\rho_{13|2=0} = 0.2$ and $\rho_{13|2=1} = 0$. However, the continuous part is recovered well with the *p*-value of 0.8673.

   The true parameter value is recovered with the non-constant conditional copula vine approach. Moreover, the discrete part of the model is recovered well with this approach with *p*-value of 0.9497. The continuous part is recovered well with *p*-value as in the constant case for this data set.

Table D.3
The parameters of the non-constant conditional copula-vine model with the second vine structure. The bolded brackets indicate if the parameter value is taken to be zero or constant.

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
$\rho_{WD,CT}$	0.1917 (0.1056, 0.2773)	$\rho_{WD,OV\|CT=0,WM=1,WT=0,RH=0}$	0.0021 (**0**) (−0.0535, 0.6521)
$\rho_{CT,WM}$	0.0158 (**0**) (−0.3057, 0.2988)	$\rho_{WD,OV\|CT=0,WM=1,WT=0,RH=1}$	0.0347 (**0**) (−0.9900, 0.9286)
$\rho_{WM,WT}$	0.0540 (**0**) (−0.2347, 0.3612)	$\rho_{WD,OV\|CT=0,WM=1,WT=1,RH=0}$	0.7979 (**0**) (−0.5865, 0.8125)
$\rho_{WT,RH}$	0.5769 (0.4001, 0.7189)	$\rho_{WD,OV\|CT=0,WM=1,WT=1,RH=1}$	−0.6522 (**0**) (−0.8752, 0.2321)
$\rho_{RH,OV}$	0.1288 (**0**) (−0.2879, 0.3812)	$\rho_{WD,OV\|CT=1,WM=0,WT=0,RH=0}$	0.1171 (**0**) (−0.0601, 0.3007)
$\rho_{OV,LC}$	0.0048 (**0**) (−0.1550, 0.1846)	$\rho_{WD,OV\|CT=1,WM=0,WT=0,RH=1}$	−0.0474 (**0**) (−0.8984, 0.6787)
$\rho_{LC,LAT}$	0.0876 (0.0255, 0.1643)	$\rho_{WD,OV\|CT=1,WM=0,WT=1,RH=0}$	0.2015 (**0**) (−0.1376, 0.5172)
$\rho_{WD,WM\|CT=0}$	0.1164 (**0.0799**) (0.0110, 0.2311)	$\rho_{WD,OV\|CT=1,WM=0,WT=1,RH=1}$	0.2260 (**0**) (−0.2172, 0.7088)
$\rho_{WD,WM\|CT=1}$	0.0525 (**0.0799**) (−0.0592, 0.2003)	$\rho_{WD,OV\|CT=1,WM=1,WT=0,RH=0}$	0.4038 (**0**) (−0.2102, 0.8559)
$\rho_{CT,WT\|WM=0}$	−0.1299 (−**0.1002**) (−0.1879, −0.0449)	$\rho_{WD,OV\|CT=1,WM=1,WT=0,RH=1}$	0.3689 (**0**) (−0.8062, 0.9023)
$\rho_{CT,WT\|WM=1}$	−0.1137 (−**0.1002**) (−0.3664, 0.1822)	$\rho_{WD,OV\|CT=1,WM=1,WT=1,RH=0}$	0.6910 (**0**) (−0.4852, 0.7102)
$\rho_{WM,RH\|WT=0}$	0.3705 (0.2578, 0.4877)	$\rho_{WD,OV\|CT=1,WM=1,WT=1,RH=1}$	0.0450 (**0**) (−0.7890, 0.89991)
$\rho_{WM,RH\|WT=1}$	−0.2464 (−0.4201, −0.0591)	$\rho_{CT,LC\|WM=0,WT=0,RH=0,OV=0}$	−0.3077 (−0.3869, −0.2382)
$\rho_{WT,OV\|RH=0}$	0.2558 (0.0799, 0.3770)	$\rho_{CT,LC\|WM=0,WT=0,RH=0,OV=1}$	−0.2434 (**0**) (−0.5945, 0.2589)
$\rho_{WT,OV\|RH=1}$	−0.2800 (−0.5201, −0.0937)	$\rho_{CT,LC\|WM=0,WT=0,RH=1,OV=0}$	−0.1166 (**0**) (−0.3313, 0.1014)
$\rho_{RH,LC\|OV=0}$	0.0221 (**0**) (−0.0536, 0.0676)	$\rho_{CT,LC\|WM=0,WT=0,RH=1,OV=1}$	0.1810 (**0**) (−0.9859, 0.9105)
$\rho_{RH,LC\|OV=1}$	0.0475 (**0**) (−0.0356, 0.1213)	$\rho_{CT,LC\|WM=0,WT=1,RH=0,OV=0}$	−0.4050 (−0.5365, −0.2839)
$\rho_{OV,LAT\|LC}$	0.1345 (**0**) (−0.0505, 0.1792)	$\rho_{CT,LC\|WM=0,WT=1,RH=0,OV=1}$	0.2889 (**0**) (−0.7398, 0.9262)
$\rho_{WD,WT\|CT=0,WM=0}$	0.0039 (−**0.0705**) (−0.0718, 0.0921)	$\rho_{CT,LC\|WM=0,WT=1,RH=1,OV=0}$	−0.1858 (**0**) (−0.3254, 0.0279)
$\rho_{WD,WT\|CT=0,WM=1}$	−0.0131 (−**0.0705**) (−0.4530, 0.4603)	$\rho_{CT,LC\|WM=0,WT=1,RH=1,OV=1}$	0.1443 (**0**) (−0.6481, 0.7972)
$\rho_{WD,WT\|CT=1,WM=0}$	−0.1262 (−**0.0705**) (−0.2042, −0.0441)	$\rho_{CT,LC\|WM=1,WT=0,RH=0,OV=0}$	−0.3548 (−0.7212, −0.0048)
$\rho_{WD,WT\|CT=1,WM=1}$	−0.2133 (−**0.0705**)	$\rho_{CT,LC\|WM=1,WT=0,RH=0,OV=1}$	−0.0132 (**0**)

Table D.3 (*continued*)

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
	(−0.4898, −0.0596)		(−0.3285, 0.4215)
$\rho_{CT,RH\|WM=0,WT=0}$	−0.1860 (−0.2962, −0.0669)	$\rho_{CT,LC\|WM=1,WT=0,RH=1,OV=0}$	−0.3088 (**0**) (−0.8474, 0.2670)
$\rho_{CT,RH\|WM=0,WT=1}$	0.0850 (**0**) (−0.0294, 0.2311)	$\rho_{CT,LC\|WM=1,WT=0,RH=1,OV=1}$	−0.8929 (**0**) (−0.9892, 0.2102)
$\rho_{CT,RH\|WM=1,WT=0}$	0.4710 (0.1661, 0.7513)	$\rho_{CT,LC\|WM=1,WT=1,RH=0,OV=0}$	−0.4044 (**0**) (−0.7651, 0.3944)
$\rho_{CT,RH\|WM=1,WT=1}$	−0.3282 (**0**) (−0.6871, 0.2444)	$\rho_{CT,LC\|WM=1,WT=0,RH=1,OV=1}$	−0.8541 (**0**) (−0.9120, 0.1714)
$\rho_{WM,OV\|WT=0,RH=0}$	0.1057 (**0**) (−0.0044, 0.3536)	$\rho_{CT,LC\|WM=1,WT=1,RH=1,OV=0}$	0.8752 (**0**) (−0.2158, 0.8952)
$\rho_{WM,OV\|WT=0,RH=1}$	0.7767 (0.4769, 0.9217)	$\rho_{CT,LC\|WM=1,WT=1,RH=1,OV=1}$	0.1058 (**0**) (−0.9912, 0.9137)
$\rho_{WM,OV\|WT=1,RH=0}$	0.7417 (0.6154, 0.8840)	$\rho_{WM,LAT\|WM,WT,RH,OV,LC}$	0.0482 (**0**) (−0.0282, 0.1168)
$\rho_{WM,OV\|WT=1,RH=1}$	0.5844 (**0**) (−0.0413, 0.8359)	$\rho_{WD,LC\|CT=0,WM=0,WT=0,RH=0,OV=0}$	0.1327 (0.0325, 0.2354)
$\rho_{WT,LC\|RH=0,OV=0}$	0.0989 (**0.0518**) (0.0249, 0.1775)	$\rho_{WD,LC\|CT=0,WM=0,WT=0,RH=0,OV=1}$	−0.1660 (**0**) (−0.6103, 0.3359)
$\rho_{WT,LC\|RH=0,OV=1}$	−0.0254 (**0.0518**) (−0.3784, 0.3627)	$\rho_{WD,LC\|CT=0,WM=0,WT=0,RH=1,OV=0}$	−0.1555 (**0**) (−0.3646, 0.0429)
$\rho_{WT,LC\|RH=1,OV=0}$	−0.0499 (**0.0518**) (−0.1707, 0.0583)	$\rho_{WD,LC\|CT=0,WM=0,WT=0,RH=1,OV=1}$	0.0270 (**0**) (−0.5905, 0.6237)
$\rho_{WT,LC\|RH=1,OV=1}$	1248 (**0.0518**) (−0.3387, 0.6159)	$\rho_{WD,LC\|CT=0,WM=0,WT=1,RH=0,OV=0}$	0.0412 (**0**) (−0.1530, 0.2015)
$\rho_{RH,LAT\|OV,LC}$	−0.0378 (**0**) (−0.1009, 0.0294)	$\rho_{WD,LC\|CT=0,WM=0,WT=1,RH=0,OV=1}$	−0.1182 (**0**) (−0.6320, 0.5787)
$\rho_{WD,RH\|CT=0,WM=0,WT=0}$	0.0 − 0.0885 (**0**) (−0.2119, 0.0557)	$\rho_{WD,LC\|CT=0,WM=0,WT=1,RH=1,OV=0}$	0.0088 (**0**) (−0.1520, 0.1420)
$\rho_{WD,RH\|CT=0,WM=0,WT=1}$	0.0331 (**0**) (−0.0953, 0.1447)	$\rho_{WD,LC\|CT=0,WM=0,WT=1,RH=1,OV=1}$	−0.5884 (**0**) (−0.7108, 0.5534)
$\rho_{WD,RH\|CT=0,WM=1,WT=0}$	−0.0110 (**0**) (−0.6160, 0.5105)	$\rho_{WD,LC\|CT=0,WM=1,WT=0,RH=0,OV=0}$	0.3723 (**0**) (−0.0190, 0.6955)
$\rho_{WD,RH\|CT=0,WM=1,WT=1}$	0.6528 (0.1277, 0.9858)	$\rho_{WD,LC\|CT=0,WM=1,WT=0,RH=0,OV=1}$	0.0025 (**0**) (−0.6094, 0.5787)
$\rho_{WD,RH\|CT=1,WM=0,WT=0}$	−0.0964 (**0**) (−0.2475, 0.0042)	$\rho_{WD,LC\|CT=0,WM=1,WT=0,RH=1,OV=0}$	0.4145 (**0**) (−0.5369, 0.6384)
$\rho_{WD,RH\|CT=1,WM=0,WT=1}$	−0.0357 (**0**) (−0.2025, 0.1097)	$\rho_{WD,LC\|CT=0,WM=1,WT=0,RH=1,OV=1}$	0.4555 (**0**) (−0.65 335, 0.5923)
$\rho_{WD,RH\|CT=1,WM=1,WT=0}$	0.0501 (**0**) (−0.4512, 0.4445)	$\rho_{WD,LC\|CT=0,WM=1,WT=1,RH=0,OV=0}$	−0.0666 (**0**) (−0.6293, 0.5766)
$\rho_{WD,RH\|CT=1,WM=1,WT=1}$	0.2930 (**0**) (−0.2297, 0.7713)	$\rho_{WD,LC\|CT=0,WM=1,WT=1,RH=0,OV=1}$	−0.0032 (**0**) (−0.9355, 0.6207)
$\rho_{CT,OV\|WM=0,WT=0,RH=0}$	0.1087 (**0**)	$\rho_{WD,LC\|CT=0,WM=1,WT=1,RH=1,OV=0}$	0.0003 (**0**)

(*continued on next page*)

Table D.3 (*continued*)

Parameter	Value (Conf. bound)	Parameter	Value (Conf. bound)
	(−0.0721, 0.3212)		(−0.6187, 0.6770)
$\rho_{CT,OV\|WM=0,WT=0,RH=1}$	−0.2696 (**0**) (−0.6687, 0.1721)	$\rho_{WD,LC\|CT=0,WM=1,WT=1,RH=1,OV=1}$	0.5232 (**0**) (−0.6394, 0.6745)
$\rho_{CT,OV\|WM=0,WT=1,RH=0}$	0.6394 (0.4190, 0.7696)	$\rho_{WD,LC\|CT=1,WM=0,WT=0,RH=0,OV=0}$	−0.0422 (**0**) (−0.1206, 0.0312)
$\rho_{CT,OV\|WM=0,WT=1,RH=1}$	−0.5118 (−0.8215, −0.1795)	$\rho_{WD,LC\|CT=1,WM=0,WT=0,RH=0,OV=1}$	−0.0009 (**0**) (−0.3849, 0.4802)
$\rho_{CT,OV\|WM=1,WT=0,RH=0}$	−0.6379 (**0**) (−0.8902, 0.0120)	$\rho_{WD,LC\|CT=1,WM=0,WT=0,RH=1,OV=0}$	0.2766 (0.0699, 0.5152)
$\rho_{CT,OV\|WM=1,WT=0,RH=1}$	0.5648 (**0**) (−0.1297, 0.9321)	$\rho_{WD,LC\|CT=1,WM=0,WT=0,RH=1,OV=1}$	0.4752 (**0**) (−0.7275, 0.5070)
$\rho_{CT,OV\|WM=1,WT=1,RH=0}$	0.7001 (0.2802, 0.9902)	$\rho_{WD,LC\|CT=1,WM=0,WT=1,RH=0,OV=0}$	−0.0838 (**0**) (−0.2379, 0.0686)
$\rho_{CT,OV\|WM=1,WT=1,RH=1}$	−0.0105 (**0**) (−0.8297, 0.7251)	$\rho_{WD,LC\|CT=1,WM=0,WT=1,RH=0,OV=1}$	−0.0317 (**0**) (−0.6738, 0.5990)
$\rho_{WM,LC\|WT=0,RH=0,OV=0}$	0.1031 (**0**) (−0.0104, 0.2448)	$\rho_{WD,LC\|CT=1,WM=0,WT=1,RH=1,OV=0}$	0.0539 (**0**) (−0.0930, 0.2170)
$\rho_{WM,LC\|WT=0,RH=0,OV=1}$	0.4134 (**0**) (−0.0495, 0.8440)	$\rho_{WD,LC\|CT=1,WM=0,WT=1,RH=1,OV=1}$	0.3084 (**0**) (−0.5442, 0.6368)
$\rho_{WM,LC\|WT=0,RH=1,OV=0}$	−0.2080 (**0**) (−0.4908, 0.0423)	$\rho_{WD,LC\|CT=1,WM=1,WT=0,RH=0,OV=0}$	0.2437 (**0**) (−0.1330, 0.6317)
$\rho_{WM,LC\|WT=0,RH=1,OV=1}$	−0.3524 (**0**) (−0.9641, 0.5087)	$\rho_{WD,LC\|CT=1,WM=1,WT=0,RH=0,OV=1}$	−0.3865 (**0**) (−0.6357, 0.5463)
$\rho_{WM,LC\|WT=1,RH=0,OV=0}$	−0.1405 (**0**) (−0.4141, 0.0432)	$\rho_{WD,LC\|CT=1,WM=1,WT=0,RH=1,OV=0}$	0.1461 (**0**) (−0.5465, 0.6277)
$\rho_{WM,LC\|WT=1,RH=0,OV=1}$	0.2428 (**0**) (−0.2875, 0.3106)	$\rho_{WD,LC\|CT=1,WM=1,WT=0,RH=1,OV=1}$	−0.4752 (**0**) (−0.5262, 0.6218)
$\rho_{WM,LC\|WT=1,RH=1,OV=0}$	−0.0932 (**0**) (−0.3622, 0.2421)	$\rho_{WD,LC\|CT=1,WM=1,WT=1,RH=0,OV=0}$	0.2478 (**0**) (−0.2941, 0.7060)
$\rho_{WM,LC\|WT=1,RH=1,OV=1}$	−0.0365 (**0**) (−0.7570, 0.5941)	$\rho_{WD,LC\|CT=1,WM=1,WT=1,RH=0,OV=1}$	0.0001 (**0**) (−0.5741, 0.7188)
$\rho_{WT,LAT\|WH,OV,LC}$	−0.1573 (−0.2402, −0.0745)	$\rho_{WD,LC\|CT=1,WM=1,WT=1,RH=1,OV=0}$	0.5495 (**0**) (−0.5960, 0.5883)
$\rho_{WD,OV\|CT=0,WM=0,WT=0,RH=0}$	0.3105 (0.0854, 0.5598)	$\rho_{WD,LC\|CT=1,WM=1,WT=1,RH=1,OV=1}$	0.5520 (**0**) (−0.5217, 0.5804)
$\rho_{WD,OV\|CT=0,WM=0,WT=0,RH=1}$	0.0768 (**0**) (−0.3225, 0.5967)	$\rho_{CT,LAT\|WM,WT,RH,OV,LC}$	−0.1110 (−0.1976, −0.0368)
$\rho_{WD,OV\|CT=0,WM=0,WT=1,RH=0}$	0.1082 (**0**) (−0.4586, 0.5742)	$\rho_{WD,LAT\|CT,WM,WT,RH,OV,LC}$	0.0997 (0.0115, 0.1690)
$\rho_{WD,OV\|CT=0,WM=0,WT=1,RH=1}$	0.4862 (0.1721, 0.8688)		

We observe that the parameters estimation procedure of the (non)-constant copula-vine models works for the two artificially generated data sets as the true parameters values are captured in the corresponding confidence intervals. The constant copula-vine model does not recover the discrete part of the data well in the non-constant data set.

Appendix C. Table of parameters

See Table C.1.

Appendix D. A second vine structure

In the main body of this paper, the ordering of *CT*, *WM*, *WT*, *RH*, *OV*, *WD*, *LC*, *LAT* is chosen in a D-vine. This is one of the many ordering possibilities which cluster the discrete and continuous variables together. Another ordering with a D-vine can be chosen such that more parameters are involved in the model. The ordering of *WD*, *CT*, *WM*, *WT*, *RH*, *OV*, *LC*, *LAT* increases the parameters number to 28 as more mixed discrete–continuous pairings are introduced in the left part of the structure. Table D.3 presents the parameters values of the second structure with the non-constant conditional copula-vine model.

As with the first structure, constant and non-constant conditional copula-vine model each with sequential and full optimization approach can be constructed. The log-likelihoods, numbers of parameters, AIC scores, and p-values of the latency time predictions of the four models are presented in Table D.2. Moreover, the corresponding results from the first structure as in Table 4 are presented as well in brackets as comparisons. We observe that the log-likelihoods are slightly improved with the second structure. Moreover, both structures predict the latency time distribution well.

References

Breymann, W., Dias, A., Embrechts, P., 2003. Dependence structures for multivariate high-frequency data in finance. Quant. Finance 3, 1–14.
Carley, H., 2002. Maximum and minimum extensions of finite subcopulas. Comm. Statist. Theory Methods 31, 2151–2166.
Cooke, R.M., Kurowicka, D., Wilson, K., 2015. Sampling, conditionalizing, counting, merging, searching regular vines. J. Multivariate Anal. 138, 4–18.
De Leon, A.R., Wu, B., 2010. Copula-based regression models for a bivariate mixed discrete and continuous outcome. Stat. Med. 30, 175–185.
Denuit, M., Lambert, P., 2005. Constraints on concordance measures in bivariate discrete data. J. Multivariate Anal. 93 (1), 40–57.
Genest, C., Nešlehová, J., 2007. A primer on copulas for count data. Astin Bull. 37 (2), 475–514.
He, J., Li, H., Edmonson, A.C., Rader, D.J., Li, M., 2012. A Gaussian copula approach for the analysis of secondary phenotypes in case-control genetic association studies. Biostatistics 13, 497–508.
Kurowicka, D., Cooke, R., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. In: Wiley Series in Probability and Statistics, Wiley.
Kurowicka, D., Joe, H., 2011. Dependence Modeling. Vine Copula Handbook. World Scientific.
Margaritis, D., 2003. Learning Bayesian network model structure from data (Ph.D. thesis), Carnegie-Mellon University, Pittsburgh, PA.
Mesfioui, M., Tajar, A., 2005. On the properties of some nonparametric concordance measures in the discrete case. J. Nonparametr. Stat. 17 (5), 541–554.
Morales Napoles, O., 2009. Bayesian belief nets and vines in aviation safety and other applications (Ph.D. thesis), Delft University of Technology.
Nelsen, R.B., 2006. An Introduction to Copula. Springer.
Nešlehová, J., 2007. On rank correlation measures for non-continuous random variables. J. Multivariate Anal. 98 (3), 554–567.
Nikoloulopoulos, A.K., 2013. On the estimation of normal copula discrete regression models using the continuous extension and simulated likelihood. J. Statist. Plann. Inference 143, 1923–1937.
Nikoloulopoulos, A.K., 2015. Efficient estimation of high-dimensional multivariate normal copula models with discrete spatial responsesd. Stoch. Environ. Res. Risk Assess. 1–13.
Nikoloulopoulos, A.K., Karlis, D., 2008. Multivariate logit copula model with an application to dental data. Stat. Med. 27, 6393–6406.
Panagiotelis, A., Czado, C., Joe, H., 2012. Pair copula constructions for multivariate discrete data. Am. Stat. Assoc. 107 (499), 1063–1072.
Sklar, A., 1959. Fonctions de répartition à *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231.
Smith, M.S., Khaled, M.A., 2012. Estimation of copula models with discrete margins via Bayesian data augmentation. Am. Stat. Assoc. 107 (497), 290–303.
Song, P.X.K., Li, M., Yuan, Y., 2009. Joint regression analysis of correlated data using Gaussian copula. Biometrics 65, 60–68.
Stöber, J., Hong, H.G., Czado, C., Ghosh, P., 2015. Comorbidity of chronic diseases in the elderly: Patterns identified by a copula design for mixed responses. Comput. Statist. Data Anal. 88, 28–39.
Whittaker, J., 1990. Graphical Models in Applied Multivariate Statistics. John Wiley & Sons.
Zilko, A.A., Kurowicka, D., Hanea, A.M., Goverde, R.M., 2015. The copula Bayesian network with mixed discrete and continuous nodes to forecast railway disruption lengths. In: Proc. 6th International Conference on Railway Operations Modelling and Analysis, RailTokyo2015, Tokyo, Japan, March 2015.