Chapter 11
Copula-Based Models for Multivariate Discrete Response Data

Aristidis K. Nikoloulopoulos

Abstract In this survey we review copula-based models and methods for multivariate discrete data modeling. Advantages and disadvantages of recent contributions are summarized and a general modeling procedure is suggested in this context.

11.1 Introduction

One goal in the theory of dependence modeling and multivariate copulas is to develop copula-based models and inferential procedures for multivariate discrete responses with covariates. Discrete response types include binary, ordinal categorical, and count data. Examples of data include, among others, familial data (measurements for each member of an extended multi-generation family) in medical genetics applications, repeated measurements in health studies, item response data in psychometrics applications, etc. These multivariate discrete data have different dependence structures including features such as negative dependence. To this end, the desiderata properties of multivariate copula families for modeling multivariate discrete data are given below (see also [19, 45, 48]):

P1: Wide range of dependence, allowing both positive and negative dependence.
P2: Flexible dependence, meaning that the number of bivariate marginals is (approximately) equal to the number of dependence parameters.
P3: Computationally feasible cumulative distribution function (cdf) for likelihood estimation.
P4: Closure property under marginalization, meaning that lower-order marginals belong to the same parametric family.

A.K. Nikoloulopoulos (✉)
School of Computing Sciences, University of East Anglia, Norwich, UK
e-mail: A.Nikoloulopoulos@uea.ac.uk

P. Jaworski et al. (eds.), *Copulae in Mathematical and Quantitative Finance*,
Lecture Notes in Statistics 213, DOI 10.1007/978-3-642-35407-6_11,

P5: No joint constraints for the dependence parameters, meaning that the use of covariate functions for the dependence parameters is straightforward.

In the existing literature, none of the existing parametric families of multivariate copulas satisfy all these conditions; hence there are many challenges for copula-based models for discrete response data.

Multivariate copulas for discrete response data have been around a long time, e.g., in [19], and earlier for some simple copula models. There are also papers with simple bivariate (multivariate) discrete distributions where actually the construction is more or less a copula, but the authors do not refer to copulas, e.g., [6, 27, 34]. Simple parametric families of copulas satisfy P3; hence the joint likelihood is straightforward to derive from the probability mass function (pmf) as a finite difference of the cdf, but they provide limited dependence; see, e.g., the contributions in [8, 29, 32, 44–46, 62].

The multivariate normal (MVN) copula generated by the MVN distribution inherits the useful properties of the latter, thus allowing a wide range for dependence (P1–P2) and overcomes the drawback of limited dependence inherent in simple parametric families of copulas [41]. The MVN copula with discrete margins has been in use for a considerable length of time, e.g. [19], and much earlier in the biostatistics [2], psychometrics [35], econometrics [14], and literature. It is usually known as a multivariate, or multinomial, probit model. The multivariate probit model is a simple example of the MVN copula with univariate probit regressions as the marginals. Implementation of the MVN copula for discrete data (discretized MVN) is possible, but not easy, because the MVN distribution as a latent model for discrete response requires rectangle probabilities based on high-dimensional integrations or their approximations [45].

Similarly, this is the case for other elliptical copulas which have also been applied to discrete data [10] and lead to a model with more probabilities in the joint upper and joint lower tails than expected with discretized MVN. Another interesting contribution and flexible modeling approach are the pair-copula constructions as developed in [48] which can also allow asymmetries, i.e., more probability in joint upper or lower tails.

The remainder of the survey proceeds as follows. Section 11.2 sets the notation and provides background material on copulas for multivariate discrete response data. In Sect. 11.3 the parametric families of copulas used so far in the literature for modeling-dependent discrete data are presented. Their properties, which inherit to the copula-based models advantages and disadvantages, are described. Section 11.4 discusses estimation methods and classifies them depending on the properties of the parametric family of copulas. In Sect. 11.5 the Kendall's tau for discrete response data is presented. We conclude this survey with some discussion.

11.2 Multivariate Discrete Distributions via Copulas

By definition, a d-variate copula $C(u_1, \ldots, u_d)$ is a multivariate cdf with uniform marginals on the unit interval; see, e.g., [19, 37]. From Sklar [51], in order to express a multivariate discrete distribution for the discrete (binary, count, etc.) vector $\mathbf{Y} = (Y_1, \ldots, Y_d)$ given a vector of covariates $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ with $\mathbf{x}_j \in \mathbb{R}^{p_j}$, $j = 1, \ldots, d$, one needs to combine discrete (Bernoulli, Poisson, etc.) marginal distribution functions $F_{Y_1}(y_1; \mathbf{x}_1), \ldots, F_{Y_d}(y_d; \mathbf{x}_d)$ with a d-variate copula such for all $\mathbf{y} = (y_1, \ldots, y_d)$,

$$H_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) = \Pr(Y_1 \le y_1, \ldots, Y_d \le y_d; \mathbf{x}) = C\Big(F_{Y_1}(y_1; \mathbf{x}_1), \ldots, F_{Y_d}(y_d; \mathbf{x}_d)\Big). \tag{11.1}$$

Because the margins are discrete, as emphasized in [9], there are many possible copulas, but all of these coincide on the closure of $\mathrm{Ran}(F_1) \times \cdots \times \mathrm{Ran}(F_d)$, where $\mathrm{Ran}(F_j)$ denotes the range of F_j.

For discrete random vectors, multivariate probabilities of the form $h_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) = \Pr(Y_1 = y_1, \ldots, Y_d = y_d; \mathbf{x})$ involve 2^d finite differences of $H_{\mathbf{Y}}(\mathbf{y}; \mathbf{x})$. Let $\mathbf{s} = (s_1, \ldots, s_d)$ be vertices where each s_j is equal to either y_j or $y_j - 1$, $j = 1, \ldots, d$. Then the joint pmf $h_{\mathbf{Y}}(\cdot)$ is given by,

$$h_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) = \sum \mathrm{sgn}(\mathbf{s}) C\Big(F_{Y_1}(s_1; \mathbf{x}_1), \ldots, F_{Y_d}(s_d; \mathbf{x}_d)\Big), \tag{11.2}$$

where the sum is taken over all vertices $\mathbf{s}$, and $\mathrm{sgn}(\mathbf{s})$ is given by,

$$\mathrm{sgn}(\mathbf{s}) = \begin{cases} 1, & \text{if } s_j = y_j - 1 \text{ for an even number of } j\text{'s.} \\ -1, & \text{if } s_j = y_j - 1 \text{ for an odd number of } j\text{'s.} \end{cases}$$

Therefore likelihood inference for discrete data is simpler for copulas with computationally feasible form of the cdf (P3). Essentially, the specification of the multivariate discrete distribution in (11.1), exploiting the use of copula functions, provides complete inference, i.e., maximum likelihood estimation, calculation of joint and conditional probabilities, and standard goodness of fit procedures.

11.3 Copula-Based Models for Discrete Response Data

In this section, we review several existed copula-based models for discrete data [8, 10, 29, 32, 44, 46, 48, 56, 57, 62]. The authors assumed that the copula C comes from a specific parametric family or class of copulas; hence, its properties are inherited to the model. Although C is not uniquely defined outside the Cartesian product of the ranges of the marginal distribution functions, there is no harm in assuming that it arises from a parametric class of copulas [10, 48].

If the same copula applies for all clusters and have covariates on board, in particular continuous covariates, the number of potential values is so high and the copula becomes unique in the limit (infinite clusters). However, generally speaking the copula is not unique (identifiable) in the discrete case except on the range of the marginals [9]. The non-identifiability is a separate theoretical issue and does not have any bearing on copula dependence modeling for discrete data [19, 55], which is the main focus of this survey.

11.3.1 Archimedean

Meester and Mackay [32] proposed a parametric model for cluster correlated categorical (binary and ordinal) data, based on the d-variate Frank copula. The Frank copula belongs to the large class of Archimedean copulas. Multivariate Archimedean copulas, see, e.g., [19], have the form,

$$C(u_1, \ldots, u_d\,;\,\theta) = \phi\left(\sum_{j=1}^{d} \phi^{-1}(u_j\,;\,\theta)\,;\,\theta\right), \tag{11.3}$$

where the generator $\phi(u\,;\,\theta)$ is the Laplace transform (LT) of a univariate family of distributions of positive random variables indexed by the parameter θ, such that $\phi(\cdot)$ and its inverse have closed forms. One may refer to [31] for a general definition of an Archimedean copula where the generator is more general than an LT but still needs to satisfy certain regularity conditions. Hence, one can relax the completely monotone condition to d times alternating in sign, then Archimedean copulas based on extensions of LTs are obtained, and some of these might have negative dependence.

The Frank copula in the Archimedean family [19, page 141] has Laplace transform $\phi_F(t) = -\theta^{-1}\log\left[1-(1-e^{-\theta})e^{-t}\right], \theta > 0$. This d-variate copula is permutation-symmetric in the d arguments, thus it is a distribution for exchangeable uniform random variables on the unit interval. The Frank copula interpolates from the independence ($\theta \to 0$) to the Fréchet upper (perfect positive dependence) bound ($\theta \to \infty$). For extension of $\phi_F(t)$ for $\theta < 0$, the Frank family extends to countercomonotonicity ($\theta \to -\infty$) for $d = 2$ and only a little into negative dependence for dimensions $d \geq 3$ [19, 31]. Joe [19, pages 158–159] shows how narrow is the range of negative dependence for trivariate Frank and beyond. Hence, for bivariate discrete data a model based on Frank copula is quite popular [3, 28, 30, 32]. For another application of d-variate Frank copula for familial binary data, see [57], and for applications of various Archimedean copula-based models for multivariate count data, see [46].

To sum up, d-variate ($d > 2$) Archimedean copulas satisfy properties P3, P4, and P5, but not P1 and P2, because they allow only for exchangeable dependence, and its range becomes narrower as the dimension increases.

11.3.2 Partially Symmetric

Zimmer and Trivedi [62] and Nikoloulopoulos and Karlis [46] modeled dependent discrete response data using partially symmetric copulas. Joe [17] extended multivariate Archimedean copulas to a more flexible class of copulas using nested LTs, the so-called partially-symmetric d-variate copulas with $d-1$ dependence parameters. Note in passing that these copulas are also called Hierarchical or nested Archimedean copulas; see, e.g., [4, 15, 16]. The multivariate form has a complex notation, so we present the trivariate extension of (11.3) to help the exposition. The trivariate form is given by,

$$C(u_1,u_2,u_3)=\phi_1\left(\phi_1^{-1}\circ\phi_2\left(\phi_2^{-1}(u_1)+\phi_2^{-1}(u_2)\right)+\phi_1^{-1}(u_3)\right), \qquad (11.4)$$

where ϕ_1,ϕ_2 are LTs and $\phi_1^{-1}\circ\phi_2\in Ł_\infty^\star=\{\omega:[0,\infty)\longrightarrow[0,\infty)|\omega(0)=0,$ $\omega(\infty)=\infty,(-1)^{j-1}\omega^j\geq 0, j=1,\ldots,\infty\}$. From the above formula it is clear that (11.4) has (1,2) bivariate margin of the form (11.3) with LT $\phi_2(\cdot;\theta_2)$, and (1,3), (2,3) bivariate margins of the form (11.3) with LT $\phi_1(\cdot;\theta_1)$. As the dimension increases there are many possible LT nestings. Bivariate margins associated with LTs that are more nested are larger in concordance than those that are less nested. For example, for (11.4) the (1,2) bivariate margins is more dependent (concordant) than the remaining bivariate margins.

Although partially symmetric copulas have a closed form cdf, they do not provide flexible dependence due to moderate number of dependence parameters ($d-1$ distinct parameters) and do not allow for negative dependence by construction. To sum up, partially symmetric copulas satisfy properties P3, P4, and P5, but not P1 and P2.

11.3.3 Farlie–Gumbel–Morgenstern

Gauvreau and Pagano [8] considered a d-variate Farlie–Gumbel–Morgenstern (FGM) copula . The multivariate FGM copula is,

$$\begin{aligned}C(u_1,\ldots,u_d;\boldsymbol{\Theta})=\Bigg(1&+\sum_{1\leq j<k\leq d}^{d}\theta_{jk}(1-u_j)(1-u_k)\\&+\sum_{1\leq j<k<l\leq d}^{d}\theta_{jkl}(1-u_j)(1-u_k)(1-u_l)+\cdots\\&+\quad\theta_{12\cdots d}(1-u_1)(1-u_2)\cdots(1-u_d)\Bigg)\prod_{j=1}^{d}u_j, \qquad (11.5)\end{aligned}$$

where $\boldsymbol{\Theta}=\{\theta_{jk},\theta_{jkl},\cdots,\theta_{12\cdots d}\}$. For more details, see [25, 26].

However, the conditions on the parameters $\boldsymbol{\Theta}$ so that FGM is indeed a copula are not investigated in [8]. The conditions on the parameters so that FGM is indeed a copula can be obtained by considering the 2^d cases for $u_j = 0$ or $1,\ j = 1\ldots,d$, and verifying that the copula density is positive, i.e. $c(u_1,\ldots,u_d) \geq 0$.

To simplify the notation a simpler version of a d-variate FGM copula that does not include higher order terms is given below,

$$C(u_1,\ldots,u_d;\theta_{jk} : 1 \leq j < k \leq d) = \Big(1 + \sum_{1\leq j<k\leq d}^{d} \theta_{jk}(1-u_j)(1-u_k)\Big)\prod_{j=1}^{d} u_j. \tag{11.6}$$

It has density function,

$$c(u_1,\ldots,u_d;\theta_{jk} : 1 \leq j < k \leq d) = 1 + \sum_{j<k}^{d} \theta_{jk}(1-2u_j)(1-2u_k).$$

The necessary and sufficient conditions on the parameters θ_{jk} so that (11.6) is a copula are straightforward. For $d = 3$, the conditions can be conveniently summarized as follows: $1 + \theta_{12} + \theta_{13} + \theta_{23} \geq 0, 1 + \theta_{12} \geq \theta_{13} + \theta_{23}, 1 + \theta_{13} \geq \theta_{12} + \theta_{23}, 1 + \theta_{23} \geq \theta_{12} + \theta_{13}$, or more succinctly $-1 + |\theta_{12} + \theta_{23}| \leq \theta_{13} \leq 1 - |\theta_{12} - \theta_{23}|, -1 \leq \theta_{12}, \theta_{13}, \theta_{23} \leq 1$. Similar conditions for higher dimension $d > 3$ can also be obtained by considering the 2^d cases for $u_j = 0$ or $1,\ j = 1\ldots,d$, and verifying that $c(u_1,\ldots,u_d) \geq 0$. For further details see [60].

In addition to the joint constraints limitation, the FGM copula has a limited range of dependence and is inappropriate for general modeling unless the responses are weakly dependent. Even for the bivariate case with no joint constraints between the parameters, it is easy to see that the range of dependence is limited. Gauvreau and Pagano [8] studied the range of the dependence parameter, say θ_{12}, in terms of Pearson's correlation parameter for binary data, say ρ_{12}, through the relation

$$\rho_{12} = \theta_{12}\sqrt{\pi_1\pi_2(1-\pi_1)(1-\pi_2)},$$

where $\pi_j = \Pr(Y_j = 1),\ j = 1, 2$. However, since $-1 \leq \theta_{12} \leq 1$ the bounds of the Pearson's correlation are,

$$\pm\sqrt{\pi_1\pi_2(1-\pi_1)(1-\pi_2)}.$$

Li and Wong [29] used a similar parametric family of copulas with the FGM copula in [8],

$$C(u_1,\ldots,u_d;\theta_{jk} : 1 \leq j < k \leq d) = \prod_{j=1}^{d} u_j \prod_{1\leq j<k\leq d}^{d} \Big(1 + \theta_{jk}(1-u_j)(1-u_k)\Big). \tag{11.7}$$

However, the conditions on the parameters θ_{jk} so that (11.7) is a copula are not investigated by the authors. For $d = 3$, the necessary conditions can be conveniently summarized as follows: $-1 \le \theta_{12}, \theta_{13}, \theta_{23} \le 1$ and $-(1 + \theta_{j\ell}) \le \theta_{jk} + \theta_{k\ell} \le \min(1, 1 + \theta_{j\ell} + \theta_{jk}\theta_{k\ell})$ for all different permutations of (j, k, ℓ) in $(1, 2, 3)$, see [39].

The sufficient conditions (nonnegativity of the entire density function in $[0.1]^d$) are hard to prove for $d > 2$ because the density of (11.7) is a higher order polynomial function (quadratic for $d = 3$, cubic for $d = 4$, etc.) of each u_j taken separately. However, considering the 2^d cases for $u_j = 0$ or 1, $j = 1, \dots, d$, and verifying that the copula density of (11.7) is positive provides the necessary conditions on the parameters for the copula in (11.7); these are also sufficient for $d = 2$ since the bivariate density is a linear function of each u_j taken separately, see [25, Sect. 4, page 419]. In addition to the joint constraints limitation, the copula in (11.7) has a limited range of dependence as the FGM copula in (11.5) or (11.6) and it resembles FGM for the bivariate case.

To sum up, the FGM copulas satisfy properties P3 and P4, but not P1, P2, and P5. Because of the dependence range limitation, the FGM copulas are not very useful for general modeling unless the responses are weakly dependent.

11.3.4 Finite Normal Mixture

Nikoloulopoulos and Karlis [45] modeled multivariate count data proposing a copula generated by a mixture of two independent MVN distributions. The finite normal mixture copula cdf takes the form,

$$C(u_1, \dots, u_d; \pi, \mu_1 = 1, \dots, \mu_d) = \mathscr{F}_{1\dots d}\left[\mathscr{F}_1^{-1}(u_1; \pi, 1), \dots, \mathscr{F}_d^{-1}(u_d; \pi, \mu_d); \pi, \boldsymbol{\mu}\right],$$

where

$$\mathscr{F}_{1\dots d}(\cdot; \pi, \boldsymbol{\mu}) = \pi \Phi_d(\cdot; \boldsymbol{\mu}, \mathbf{I}_d) + (1 - \pi)\Phi_d(\cdot; -\boldsymbol{\mu}, \mathbf{I}_d) \tag{11.8}$$

is the d-variate cdf of a mixture of two d-variate normal cdfs with mixing probability π, $\Phi_d(\cdot; \boldsymbol{\mu}, \mathbf{I}_d)$ denotes the cdf of the d-variate normal distribution function with mean $\boldsymbol{\mu} = (1, \mu_2, \dots, \mu_d)$ and covariance matrix the d-variate diagonal identity matrix $\mathbf{I}_d$, and $\mathscr{F}_j(\cdot; \pi, \mu_j) = \pi\Phi(\cdot; \mu_j, 1) + (1 - \pi)\Phi(\cdot; -\mu_j, 1)$, $j = 1, \dots, d$ is the univariate cdf of a mixture of two univariate normal cdfs. Essentially, since the variables are uncorrelated upon conditioning by the component, the d-variate normal cdfs in (11.8) can be easily calculated as the product of univariate normal cdfs.

In this construction the mixing operation introduces the dependence structure. The covariance matrix of the 2-finite normal mixture distribution is of the form,

$$\boldsymbol{\Delta} = \mathbf{I}_d + \boldsymbol{\mu}\boldsymbol{\mu}^\top \tag{11.9}$$

$$\boldsymbol{\Delta} = \begin{bmatrix} 2 & \mu_2 & \dots & \mu_{d-1} & \mu_d \\ \mu_2 & 1+\mu_2^2 & \dots & \mu_2\mu_{d-1} & \mu_2\mu_d \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{m-1} & \mu_2\mu_{d-1} & \dots & 1+\mu_{d-1}^2 & \mu_{d-1}\mu_d \\ \mu_d & \mu_2\mu_d & \dots & \mu_{d-1}\mu_d & 1+\mu_d^2 \end{bmatrix}.$$

Clearly, the covariance matrix in (11.9) is identifiable and has $d-1$ dependence parameters. This dependence construction is similar to the partially symmetric copula of [17]; in the present case, however, the (j,k) marginal for $j \neq k \neq 1$ has two copula parameters, and thus more flexible association.

Mathematically, this family has nice features, a moderate number of parameters to model dependence (including negative dependence), and a rather simple computational form but does not provide such flexible or wide range of dependence. For example, it cannot model multivariate discrete data with strong or with negative dependence among many random variables or at least it cannot capture all the possible structures. To sum up, finite normal mixture copulas satisfy properties P3, P4, and P5, but not P1 and P2.

11.3.5 Mixtures of max-id

Joe [19] and Nikoloulopoulos and Karlis [44, 46] applied mixtures of max-id copulas to model multivariate discrete data. Joe and Hu [22] extended multivariate Archimedean copulas to a more flexible class of copulas using mixture of max-id copulas $C_{jk}^{(m)}$ of the form,

$$C(u_1,\dots,u_d;\theta,\theta_{jk}:1\le j<k\le d) = \tag{11.10}$$

$$\phi\left(-\sum_{1\le j<k\le d}\log C_{jk}^{(m)}\big(e^{-p_j\phi^{-1}(u_j;\theta)}, e^{-p_k\phi^{-1}(u_k;\theta)};\theta_{jk}\big) + \sum_{j=1}^{d}\nu_j p_j\phi^{-1}(u_j;\theta);\theta\right),$$

where $p_j = (\nu_j + d - 1)^{-1}$, $j = 1,\dots,d$. Since the mixing operation introduces dependence, this copula has a dependence structure that comes from the form of $C_{jk}^{(m)}(\cdot;\theta_{jk})$ and the form of the Laplace transform $\phi(\cdot;\theta)$. Another interesting interpretation is that the Laplace transform ϕ introduces the smallest dependence between random variables (exchangeable dependence), while the copulas $C_{jk}^{(m)}$ add some pairwise dependence. The parameters ν_j are included in order that the parametric family of multivariate copulas (11.10) is closed under margins. Regarding ν_j zero or fixed, the copula of the form (11.10) is a family with $1 + d(d-1)/2$ parameters that allows only positive but flexible dependence structure.

One may simplify the form of the copula by assuming $C_{jk}^{(m)}(u_j, u_k) = u_j\, u_k$ (known as independence or product copula) together with $\nu_j = \nu_k = -1$, for some pairs. This implies that for those pairs of variables, the minimum level of dependence is introduced by ϕ.

This construction, on the one hand, does not impose any constraints between the dependence parameters θ_{jk}, but, on the other hand, does not allow for negative dependence [22]. The latter is the only drawback of this class of parametric families of copulas.

To sum up, d-variate mixtures of max-infinitely divisible copulas satisfy all properties except P1, since they don't allow for negative associations. Note in passing that using mixtures of max-id copulas the use of covariate functions for the copula dependence parameters is straightforward since they fulfill P5.

11.3.6 Elliptical

Two well-known members of elliptical copulas [1, 7], the MVN and Student t copulas, have been used in the literature for prediction and modeling of dependent discrete data.

Joe [19] and Song [55] modeled dependent discrete data using the MVN copula,

$$C(u_1, \ldots, u_d; \mathbf{R}) = \Phi_d\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d); \mathbf{R}\right), \tag{11.11}$$

where $\Phi_d(\cdot; \mathbf{R})$ denotes the standard MVN distribution function with correlation matrix $\mathbf{R} = (\rho_{jk} : 1 \leq j < k \leq d)$ and Φ is the cdf of the univariate standard normal. The MVN copula inherits the dependence structure of MVN distribution, and thus admits a wide range of flexible dependence allowing both positive and negative dependence (P1–P2). The drawback of the MVN copula is with relation to the computation of the rectangle probabilities. This computation involves repeated multidimensional integration since MVN lacks a closed form cdf. Consequently, likelihood inference might be difficult; see [45]. Note that for the special case of positive exchangeable correlation structures, the d-dimensional integrals conveniently reduce to 1-dimensional integrals [24, p. 48].

The pmf can be obtained by computing the following rectangle probability [40, 48],

$$\begin{aligned} h_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) &= \Pr(Y_1 = y_1, \ldots, Y_d = y_d; \mathbf{x}) \qquad (11.12)\\ &= \Pr(y_1 - 1 < Y_1 \leq y_1, \ldots, y_d - 1 < Y_d \leq y_d; \mathbf{x}) \\ &= \int_{\Phi^{-1}[F_{Y_1}(y_1-1;\mathbf{x}_1)]}^{\Phi^{-1}[F_{Y_1}(y_1;\mathbf{x}_1)]} \cdots \int_{\Phi^{-1}[F_{Y_d}(y_d-1;\mathbf{x}_d)]}^{\Phi^{-1}[F_{Y_d}(y_d;\mathbf{x}_d)]} \phi_d(z_1, \ldots, z_d; \mathbf{R}) dz_1 \ldots dz_d, \end{aligned}$$

where ϕ_d denotes the standard MVN density with correlation matrix $\mathbf{R}$.

There are several papers in the literature that focus on the computation of the MVN rectangle probabilities for general correlation structures, and, conveniently, the implementation of the proposed algorithms is available in contributed R packages.[1] Schervish [50] proposed a locally adaptive numerical integration method but this method, while more accurate, is time consuming and restricted to a low dimension. Therefore, Genz and Bretz [11] proposed a randomized quasi Monte Carlo method with the use of antithetic variates and Joe [18] proposed two approximations to MVN probabilities. These advances in computation of MVN probabilities can be used to implement MVN copula models with discrete response data.

Genest et al. [10] modeled dependent binary data using the Student t copula,

$$C(u_1,\ldots,u_d;\mathbf{R},\nu)=T_d\left(T^{-1}(u_1;\nu),\ldots,T^{-1}(u_d;\nu);\mathbf{R},\nu\right), \tag{11.13}$$

where $T(\cdot;\nu)$ is the univariate Student t cdf with (non-integer) ν degrees of freedom, and $T_d(\cdot;\mathbf{R},\nu)$ is the cdf of a multivariate Student t distribution with ν degrees of freedom and correlation matrix $\mathbf{R}$. Student t copula share with the MVN copula the ability to accommodate any feasible pattern of association in a set of random variables. However, Student t copula can also account for tail dependence in multivariate continuous data [42], whereas MVN copula cannot. In the context of multivariate discrete data that means that more probabilities can be assigned in the joint upper and joint lower tails than with the MVN copula. Student t copula cannot also be expressed in closed form; however, the rectangle probabilities can also be computed using the methods in [11].

To sum up, elliptical copulas satisfy properties P1, P2, and P4, but not P3, and P5, since they lack a closed form cdf and a positive-definite matrix is required respectively.

11.3.7 Vine

In the literature, vine copulas are suitable for modeling multivariate continuous data with various features such as tail dependence [23]. Since the densities of multivariate vine copulas can be factorized in terms of bivariate linking copulas and lower-dimensional margins, they are computationally tractable for high-dimensional continuous variables. The cdf of d-dimensional vine copula lacks a closed form and requires $(d-1)$-dimensional integration [19]. Hence, in order to derive the pmf as finite difference of the cdf poses nonnegligible numerical challenges.

---

[1]Both approximations to MVN rectangle in [18], the 1-dimensional integral in the positive exchangeable case, and the method in [50], can be computed with the functions `mvnapp`, `exchmvn`, and `pmnorm`, respectively, in the R package `mprobit` [21]. The methods in [11] can be computed with the function `pmvnorm` in the R package `mvtnorm` [12].

Recently, Panagiotelis et al. [48] decomposed the pmf as follows,

$$\begin{aligned}\Pr(Y_1 = y_1, \ldots, Y_d = y_d) = \Pr(Y_1 = y_1 | Y_2 = y_2, \ldots, Y_d = y_d) \times \\ \Pr(Y_2 = y_2 | Y_3 = y_3, \ldots, Y_d = y_d) \times \cdots \times \Pr(Y_d = y_d).\end{aligned} \tag{11.14}$$

Letting V_h be any scalar element of $\mathbf{V}$ and $\mathbf{V}_{\backslash h}$ its complement, with Y_j not an element of $\mathbf{V}$, each term on the right-hand side of (11.14) has the form $Pr(Y_j = y_j | V = v)$ where y_j is a scalar element of y and $\mathbf{v}$ is a subset of $\mathbf{y}$,

$$Pr(Y_j = y_j | \mathbf{V} = \mathbf{v}) = \frac{\Pr(Y_j = y_j, \mathbf{V}_h = \mathbf{v}_h | V_{\backslash h} = \mathbf{v}_{\backslash h})}{\Pr(\mathbf{V}_h = \mathbf{v}_h, \mathbf{V}_{\backslash h} = \mathbf{v}_{\backslash h})}$$

$$= \frac{\sum_{i_j=0,1} \sum_{i_h=0,1} (-1)^{i_j + i_h} \Pr(Y_j \le y_j - i_j, \mathbf{V}_h \le \mathbf{v}_h - i_h | \mathbf{V}_{\backslash h} = \mathbf{v}_{\backslash h})}{\Pr(\mathbf{V}_h = \mathbf{v}_h, V_{\backslash h} = \mathbf{v}_{\backslash h})}$$

$$= \frac{\sum_{i_j=0,1} \sum_{i_h=0,1} (-1)^{i_j + i_h} C_{Y_j, \mathbf{V}_h | \mathbf{V}_{\backslash h}} \Big(F_{Y_j | \mathbf{V}_{\backslash h}}(y_j - i_j | \mathbf{v}_{\backslash h}), F_{\mathbf{V}_h | V_{\backslash h}}(\mathbf{v}_h - i_h | v_{\backslash h}) \Big)}{\Pr(\mathbf{V}_h = \mathbf{v}_h, \mathbf{V}_{\backslash h} = \mathbf{v}_{\backslash h})}.$$

The above can be applied recursively to (11.14) to decompose a multivariate pmf into bivariate copula families. More details and a three-dimensional illustration can be found in [48].

The computation of the pmf for a discrete vine only requires $2d(d-1)$ bivariate copula function evaluations, compared to 2^d multivariate copula evaluations for the finite difference approach (11.2), and Panagiotelis et al. [48] have developed a fast algorithm for computing the pmf of a vine copula with discrete margins.

A wide variety of dependence structures can be modeled by selecting different copula families as building blocks. Selecting different bivariate copula families in a discrete vine has a substantial impact on the joint probabilities of the multivariate distribution and can provide better fits when we have some discrete multivariate data where asymmetries can easily be seen.

To sum up, discrete vine copulas or pair-copula constructions satisfy all properties except P4. Note that although their cdf is not of closed form the pmf is successively decomposed and likelihood estimation is feasible even for high dimensions.

11.4 Methods of Estimation

For a sample of size n with data $\mathbf{y}_1, \ldots, \mathbf{y}_n$ the joint log-likelihood of the copula-based model is,

$$\ell = \sum_{i=1}^{n} \log h_{\mathbf{Y}}(y_{i1}, \ldots, y_{id}; \mathbf{x}_{i1}, \ldots, \mathbf{x}_{id}). \tag{11.15}$$

Estimation of the model parameters can be approached by the standard maximum likelihood method, by maximizing the joint log-likelihood in (11.15) over the univariate and copula parameters [19] or by a two-step approach called Inference Function of Margins (IFM) method in [19, 20]. In the first step, the univariate parameters are estimated assuming independence, and in the second step the joint log-likelihood in (11.15) is maximized over the copula parameters with the univariate parameters fixed at the estimated values from the first step. When the dependence is not too strong which is a realistic scenario for discrete response data, the IFM method can efficiently (in sense of computing time and asymptotic variance) estimate the model parameters. For parametric families of copulas with a closed form cdf and vine copulas, maximum likelihood or IFM estimation is straightforward.

For the elliptical copulas likelihood inference involves the computation of multidimensional rectangle probabilities of the form (11.12). The advances in computation of rectangle probabilities can be used to implement elliptical copula-based models with discrete response data. Using the the first-order (makes use of all of the univariate and bivariate marginal probabilities) or the second-order approximation (also makes use of trivariate and four-variate marginal probabilities) in [18] to compute the rectangle MVN probabilities in (11.15), the likelihood is successively approximated for weak to moderate correlation parameters. Computing the rectangle MVN/Student t probabilities in (11.15) via simulation based on the methods in [11], a simulated likelihood is implemented; see [40]. Since the estimation of the parameters of the copula-based models is obtained using a quasi-Newton routine [36] applied to the joint log-likelihood in (11.15), the use of quasi Monte Carlo simulation to four decimal place accuracy for evaluations of the rectangles works poorly, because numerical derivatives of the joint log-likelihood with respect to the parameters are not smooth. In order to achieve smoothness, the same set of uniform random variables should be used for every rectangle probability that comes up in the optimization of the simulated likelihood [40]. Asymptotic and small-sample efficiency calculations in [40] show that the simulated likelihood method, which is based on evaluating the multidimensional integrals of the joint likelihood with randomized quasi Monte Carlo methods developed in [11], is as good as maximum likelihood as shown for dimension 10 or lower. These findings are expected to hold in higher dimensions. Although there is an issue of computational burden as the dimension and the sample size increase, this will become marginal, as computing technology is advancing rapidly.

Zhao and Joe [61] proposed composite likelihood estimation methods to overcome the computational issues at the maximization routines for the MVN copula in a high-dimensional context. Composite likelihood is a surrogate likelihood which leads to unbiased estimating equations obtained by the derivatives of the composite log-likelihoods. Estimation of the model parameters can be approached by solving the estimating equations in [61] or equivalently by maximizing the sum of composite likelihoods. First consider the sum of univariate log-likelihoods,

$$\ell_1 = \sum_{i=1}^{n}\sum_{j=1}^{d} \log f_{Y_j}(y_{ij};\mathbf{x}_{ij}),$$

where $f_{Y_j}(y_1;\mathbf{x}_1),\ldots,f_{Y_d}(y_d;\mathbf{x}_d)$ are the univariate marginal pmfs, and then the sum of bivariate log-likelihoods,

$$\ell_2 = \sum_{i=1}^{n}\sum_{j<k} \log h_{\mathbf{Y}_2}(y_{ij},y_{ik};\mathbf{x}_{ij},\mathbf{x}_{ik}),$$

where $\mathbf{Y}_2 = (Y_j, Y_k)$. Composite likelihood estimates can be obtained using a two-stage method (CL1):

1. At the first step the ℓ_1 is maximized over the univariate marginal parameters.
2. At the second step the ℓ_2 is maximized over the copula parameters with univariate marginal parameters fixed as estimated at the first step of the method.

Alternatively, one can use the one stage composite likelihood estimation procedure (CL2), that is maximizing the ℓ_2 over the univariate and copula parameters at one step. The efficiency of composite likelihood estimates has been studied and shown in a series of a papers; see, e.g., [58, 59, 61]. If the interest is both to the univariate and dependence parameters, CL2 method should be performed since CL1 ignores the dependence at the estimation of the univariate marginal parameters.

Bayesian methods have also been used on the estimation of an elliptical-copula-based model. Pit et al. [49] proposed a general Bayesian approach for estimating a MVN copula-based model. Smith, Gan and Kohn [53] extend the work in [49] to other elliptical copula-based models. Very recently, Smith and Khaled [54] suggest efficient Bayesian data augmentation methodology for the estimation of copula-based models for multivariate discrete data. For a detailed exposition of Bayesian approaches on estimation of copula-based models for discrete response data we refer the interested reader to the excellent survey by Smith [52].

11.5 Dependence as Measured by Kendall's Tau

The copula parameters for different parametric families have different range; hence, they are not comparable. To compare strengths of dependence among different copula-based models and ease interpretation, it is useful to convert the estimated parameters to concordance measures such as Kendall's τ's.

For continuous random variables dependence as measured by Kendall's tau $\tau = P_c - P_d$, the difference between the probabilities of concordance (P_c) and discordance (P_d), is associated only with the copula parameters. However for discrete data the marginal distributions also play a role on dependence, and

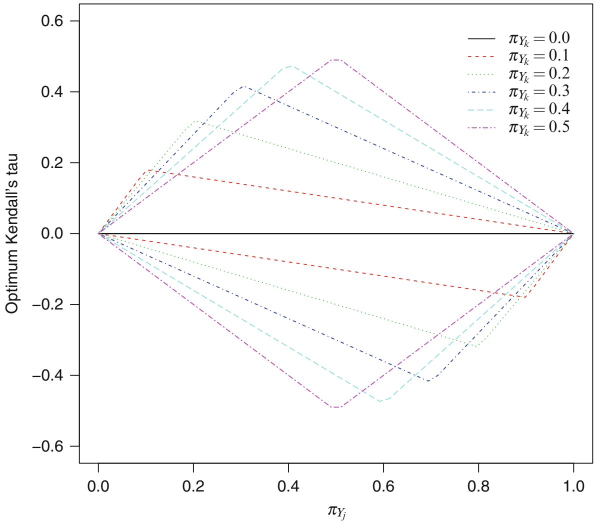

Fig. 11.1 Kendall's tau bounds when $Y_j \sim Bin(1, \pi_{Y_j})$ and $Y_k \sim Bin(1, \pi_{Y_k})$ with $(\pi_{Y_j}, \pi_{Y_k}) \in [0, 1] \times \{0, 0.1, \ldots, 0.5\}$

τ does not attain the boundary values of ± 1, because the probability of ties $P_t = 1 - (P_c + P_d)$ is positive; see [5, 33, 38].

The Kendall's tau for each pair $\mathbf{Y}_2$ is given as below [47],

$$\tau(Y_j, Y_k) = \sum_{y_j=0}^{\infty} \sum_{y_k=0}^{\infty} h_{\mathbf{Y}_2}(y_j, y_k; \mathbf{x}_j, \mathbf{x}_k) \Big\{ 4C(F_{Y_j}(y_j - 1; \mathbf{x}_j), F_{Y_k}(y_k - 1; \mathbf{x}_k)) - h_{\mathbf{Y}_2}(y_j, y_k; \mathbf{x}_j, \mathbf{x}_k) \Big\} + \sum_{y_j=0}^{\infty} f_{Y_j}^2(y_j; \mathbf{x}_j) + \sum_{y_k=0}^{\infty} f_{Y_k}^2(y_k; \mathbf{x}_k) - 1. \tag{11.16}$$

This formula helps to see clearly that in the discrete case the marginals do affect Kendall's tau.

To visualize the effect of the marginal distributions/parameters, we computed the optimum Kendall's tau values using various discrete marginal distributions, i.e., Bernoulli, binomial and Poisson and the Fréchet bound copulas. In Figs. 11.1 and 11.2 optimum Kendall's tau values have been plotted for Bernoulli, i.e., $Y_j \sim Bin(1, \pi_{Y_j})$ and $Y_k \sim Bin(1, \pi_{Y_k})$ and binomial margins, i.e., $Y_j \sim Bin(5, \pi_{Y_j})$ and $Y_k \sim Bin(5, \pi_{Y_k})$ for a grid of (π_{Y_j}, π_{Y_k}) values in $\mathscr{P}_{Y_j} \times \mathscr{P}_{Y_k}$ where $\mathscr{P}_{Y_j} = [0, 1]$ and $\mathscr{P}_{Y_k} = \{0, 0.1, \ldots, 0.5\}$, respectively. In Fig. 11.3 optimum Kendall's tau values have been plotted for Poisson marginal distributions with the same parameter λ up to 50.

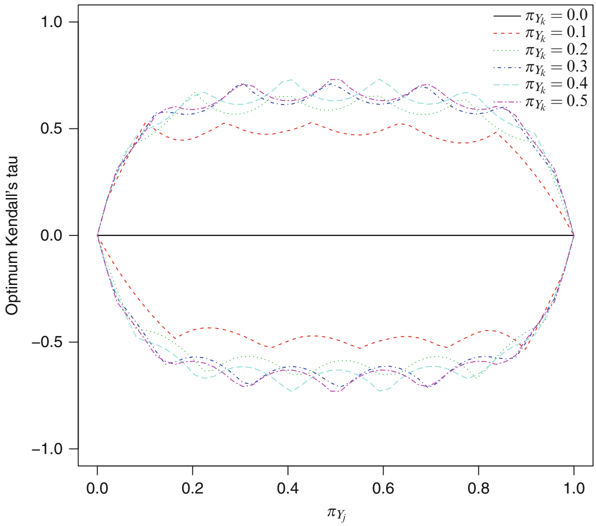

Fig. 11.2 Kendall's tau bounds when $Y_j \sim Bin(5, \pi_{Y_j})$ and $Y_k \sim Bin(5, \pi_{Y_k})$ with $(\pi_{Y_j}, \pi_{Y_k}) \in [0, 1] \times \{0, 0.1, \ldots, 0.5\}$

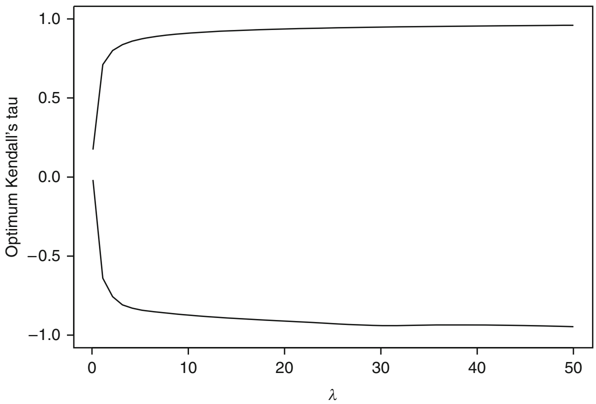

Fig. 11.3 Kendall's tau bounds when both random variables are Poisson with parameter λ

From the figures, one can see that Kendall's tau does not reach the bounds ± 1 for countermonotonic and comonotonic marginals. There is also a clear association between the optimum value of Kendall's tau and the marginal probabilities for binary and binomial data, while this association is negligible for count data with marginal parameters greater than 10. For normalized versions of Kendall's tau one can refer to [13, 38].

11.6 Discussion

This survey summarized copula-based models for discrete response data. We list several desirable properties such a model should have and introduce the models that have been used in copula dependence modeling for discrete data so far. For copula modeling with multivariate discrete data, we suggest models that admit a wide range of dependence, such as the MVN copula. Given the wide range of dependence, MVN copula provides often the best fit or nearly the best fit for discrete data [41]. However MVN copula is inadequate to model multivariate data with refection asymmetry or tail dependence [43]. Although tail dependence degenerate in the discrete case, reflection asymmetry is a realistic scenario. Vine copula constructions are suitable for modeling this kind of data since by using as bivariate blocks asymmetric bivariate copulas tail asymmetry can be accommodated, i.e., more probability in one or both joint tails can be obtained. Essentially, discrete vine copulas are highly flexible since any multivariate discrete distribution can be decomposed as a vine copula, under a set of conditions outlined in [48].

If the discrete responses are positively associated, then parametric families of copulas with a closed form cdf could be also used. Archimedean copulas could be used to model clustered data with exchangeable dependence, while mixtures of max-infinitely divisible copulas could be used for data with a more general positive dependence. Note in passing that, from copulas with positive dependence by construction, one could always get some negative dependence by applying decreasing transformations on some subset of the random variables, but this is restrictive in general, because this construction cannot model negative dependence among many random variables [46].

If the interest is to study the effect of explanatory variables on the dependence structure, Archimedean, partially symmetric, mixtures of max-id, and vine copulas are suitable since allow the use of covariate functions for the copula dependence parameters (see, e.g., [44, 47]); this is not the case for the FGM and elliptical copulas in (11.5)–(11.7), (11.11) and (11.13), because of the joint constraints for the dependence parameters.

Acknowledgments I would like to thank Professor Harry Joe, University of British Columbia, for helpful comments and suggestions on an earlier version of this survey.

References

1. Abdous, B., Genest, C., Rémillard, B.: Dependence properties of meta-elliptical distributions. In: Duchesne, P., Rémillard, B. (eds.) Statistical Modelling and Analysis for Complex Data Problems, pp. 1–15. Kluwer, Dordrecht (2005)
2. Ashford, J.R., Sowden, R.R.: Multivariate probit analysis. Biometrics **26**, 535–546 (1970)
3. Cameron, A.C., Li, T., Trivedi, P.K., Zimmer, D.M.: Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. Econom. J. **7**(2), 566–584 (2004)

4. Choros, B., Härdle, W., Okhrin, O.: CDO pricing with copulae. Discussion paper 2009-013, SFB 649, Humboldt Universität zu Berlin, Berlin (2009)
5. Denuit, M., Lambert, P.: Constraints on concordance measures in bivariate discrete data. J. Multivar. Anal. **93**(1), 40–57 (2005)
6. Famoye, F.: A new bivariate generalized Poisson distribution. Stat. Neerl. **64**, 112–124 (2010)
7. Fang, H.B., Fang, K.T., Kotz, S.: The meta-elliptical distributions with given marginals. J. Multivar. Anal. **82**, 1–16 (2002)
8. Gauvreau, K., Pagano, M.: The analysis of correlated binary outcomes using multivariate logistic regression. Biometrical J. **39**, 309–325 (1997)
9. Genest, C., Nešlehová, J.: A primer on copulas for count data. Astin Bull. **37**, 475–515 (2007)
10. Genest, C., Nikoloulopoulos, A.K., Rivest, L.P., Fortin, M.: Predicting dependent binary outcomes through logistic regressions and meta-elliptical copulas. Braz. J. Probab. Stat. (2013, in press)
11. Genz, A., Bretz, F.: Methods for the computation of the multivariate t-probabilities. J. Comput. Graph. Stat. **11**, 950–971 (2002)
12. Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Hothorn, T.: mvtnorm: Multivariate Normal and t Distributions. R package version 0.9-9992. URL http://CRAN.R-project.org/package=mvtnorm (2012)
13. Goodman, L., Kruskal, W.: Measures of association for cross classifications. J. Am. Stat. Assoc. **49**, 732–764 (1954)
14. Hausman, J., Wise, D.: A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. Econometrica **45**, 319–339 (1978)
15. Hofert, M.: Construction and sampling of nested archimedean copulas. In: Jaworski, P., Durante, F., Härdle, W., Rychlik, T. (eds.) Copula Theory and Its Applications, Proceedings of the Workshop held in Warsaw, 25–26 September, pp. 147–160. Springer, Berlin (2009)
16. Hofert, M., Mächler, M.: Nested archimedean copulas meet R: The `nacopula` package. J. Stat. Softw. **39**, 1–20 (2011)
17. Joe, H.: Parametric families of multivariate distributions with given margins. J. Multivar. Anal. **46**, 262–282 (1993)
18. Joe, H.: Approximations to multivariate normal rectangle probabilities based on conditional expectations. J. Am. Stat. Assoc. **90**(431), 957–964 (1995)
19. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
20. Joe, H.: Asymptotic efficiency of the two-stage estimation method for copula-based models. J. Multivar. Anal. **94**(2), 401–419 (2005)
21. Joe, H., Chou, L.W., Zhang, H.: mprobit: Multivariate probit model for binary/ordinal response. R package version 0.9-3. URL http://CRAN.R-project.org/package=mprobit (2011)
22. Joe, H., Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions. J. Multivar. Anal. **57**(2), 240–265 (1996)
23. Joe, H., Li, H., Nikoloulopoulos, A.K.: Tail dependence functions and vine copulas. J. Multivar. Anal. **101**, 252–270 (2010)
24. Johnson, N.L., Kotz, S.: Continuous Multivariate Distributions. Wiley, New York (1972)
25. Johnson, N.L., Kotz, S.: On some generalized Farlie-Gumbel-Morgenstern distributions. Commun. Stat. **4**, 415–427 (1975)
26. Johnson, N.L., Kotz, S.: On some generalized Farlie-Gumbel-Morgenstern distributions – II Regression, correlation and further generalizations. Commun. Stat. **6**, 485–496 (1977)
27. Lakshminarayana, J., Pandit, S.N.N., Rao, K.S.: On a bivariate Poisson distribution. Commun. Stat.-Theory Method. **28**, 267–276 (1999)
28. Lee, A.: Modelling rugby league data via bivariate negative binomial regression. Aust. N. Z. J. Stat. **41**(2), 141–152 (1999)
29. Li, J., Wong, W.K.: Two-dimensional toxic dose and multivariate logistic regression, with application to decompression sickness. Biostatistics **12**, 143–155 (2011)
30. McHale, I., Scarf, P.: Modelling soccer matches using bivariate discrete distributions with general dependence structure. Stat. Neerl. **61**(4), 432–445 (2007)

31. McNeil, A.J., Nešlehová, J.: Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. Ann. Stat. **37**, 3059–3097 (2009)
32. Meester, S., MacKay, J.: A parametric model for cluster correlated categorical data. Biometrics **50**, 954–963 (1994)
33. Mesfioui, M., Tajar, A.: On the properties of some nonparametric concordance measures in the discrete case. J. Nonparametr. Stat. **17**(5), 541–554 (2005)
34. Miravete, E.J.: Multivariate Sarmanov count data models. CEPR Discussion Paper No. DP7463, University of Texas at Austin; Centre for Economic Policy Research (2009)
35. Muthén, B.: Contributions to factor analysis of dichotomous variables. Psychometrika **43**, 551–560 (1978)
36. Nash, J.: Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation, 2nd edn. Hilger, New York (1990)
37. Nelsen, R.B.: An Introduction to Copulas. Springer, New York (2006)
38. Nešlehová, J.: On rank correlation measures for non-continuous random variables. J. Multivar. Anal. **98**, 544–567 (2007)
39. Nikoloulopoulos, A.K.: Letter to the editor. Biostatistics **13**, 1–3 (2012)
40. Nikoloulopoulos, A.K.: On the estimation of normal copula discrete regression models using the continuous extension and simulated likelihood. ArXiv e-prints, arXiv:1304.0905 (2013)
41. Nikoloulopoulos, A.K., Joe, H., Chaganty, N.R.: Weighted scores method for regression models with dependent data. Biostatistics **12**, 653–665 (2011)
42. Nikoloulopoulos, A.K., Joe, H., Li, H.: Extreme value properties of multivariate t copulas. Extremes **12**, 129–148 (2009)
43. Nikoloulopoulos, A.K., Joe, H., Li, H.: Vine copulas with asymmetric tail dependence and applications to financial return data. Comput. Stat. Data Anal. **56**, 659–3673 (2012)
44. Nikoloulopoulos, A.K., Karlis, D.: Multivariate logit copula model with an application to dental data. Stat. Med. **27**, 6393–6406 (2008)
45. Nikoloulopoulos, A.K., Karlis, D.: Finite normal mixture copulas for multivariate discrete data modeling. J. Stat. Plan. Inference **139**, 3878–3890 (2009)
46. Nikoloulopoulos, A.K., Karlis, D.: Modeling multivariate count data using copulas. Commun. Stat. Simul. Comput. **39**, 172–187 (2010)
47. Nikoloulopoulos, A.K., Karlis, D.: Regression in a copula model for bivariate count data. J. Appl. Stat. **37**, 1555–1568 (2010)
48. Panagiotelis, A., Czado, C., Joe, H.: Pair copula constructions for multivariate discrete data. J. Am. Stat. Assoc. **107**, 1063–1072 (2012)
49. Pitt, M., Chan, D., Kohn, R.: Efficient Bayesian inference for Gaussian copula regression models. Biometrika **93**, 537–554 (2006)
50. Schervish, M.: Algorithm AS 195. Multivariate normal probabilities with error bound. Appl. Stat. **33**, 81–94 (1984)
51. Sklar, M.: Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris **8**, 229–231 (1959)
52. Smith, M.S.: Bayesian approaches to copula modelling. In: Damien, P., Dellaportas, P., Polson, N., Stephens, D. (eds.) Bayesian Theory and Applications. Oxford University Press, USA (2013)
53. Smith, M.S., Gan, Q., Kohn, R.: Modeling dependence using skew t copulas: Bayesian inference and applications. J. Appl. Econom. **27**, 500–522 (2012)
54. Smith, M.S., Khaled, M.A.: Estimation of copula models with discrete margins via Bayesian data augmentation. J. Am. Stat. Assoc. **107**, 290–303 (2012)
55. Song, P.X.K.: Correlated Data Analysis: Modeling, Analytics, and Application. Springer, New Yory (2007)
56. Song, P.X.K., Li, M., Yuan, Y.: Joint regression analysis of correlated data using Gaussian copulas. Biometrics **65**, 60–68 (2009)
57. Trégouët, D.A., Ducimetière, P., Bocquet, V., Visvikis, S., Soubrier, F., Tiret, L.: A parametric copula model for analysis of familial binary data. Am. J. Hum. Genet. **64**(3), 886–93 (1999)
58. Varin, C.: On composite marginal likelihoods. Adv. Stat. Anal. **92**, 1–28 (2008)

59. Varin, C., Reid, N., Firth, D.: An overview of composite likelihood methods. Stat. Sin. **21**, 5–42 (2011)
60. Xu, J.: Statistical modelling and inference for multivariate and longitudinal discrete response. Ph.D. thesis, The University of British Columbia (1996)
61. Zhao, Y., Joe, H.: Composite likelihood estimation in multivariate data analysis. Can. J. Stat. **33**(3), 335–356 (2005)
62. Zimmer, D., Trivedi, P.: Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand. J. Bus. Econ. Stat. **24**(1), 63–76 (2006)