Stat Comput (2008) 18: 313–320
DOI 10.1007/s11222-008-9058-y

Copula, marginal distributions and model selection: a Bayesian note

Ralph dos Santos Silva · Hedibert Freitas Lopes

Received: 17 October 2007 / Accepted: 22 February 2008 / Published online: 14 March 2008

Abstract Copula functions and marginal distributions are combined to produce multivariate distributions. We show advantages of estimating all parameters of these models using the Bayesian approach, which can be done with standard Markov chain Monte Carlo algorithms. Deviance-based model selection criteria are also discussed when applied to copula models since they are invariant under monotone increasing transformations of the marginals. We focus on the deviance information criterion. The joint estimation takes into account all dependence structure of the parameters' posterior distributions in our chosen model selection criteria. Two Monte Carlo studies are conducted to show that model identification improves when the model parameters are jointly estimated. We study the Bayesian estimation of all unknown quantities at once considering bivariate copula functions and three known marginal distributions.

Keywords Copula · Deviance information criterion · Marginal distribution · Measure of dependence · Monte Carlo study · Skewness

1 Introduction

There has been a surge of interest in applications where multivariate distributions are determined by combining univariate marginal distributions with copula functions (Genest et al. 2006; Huard et al. 2006; Pitt et al. 2006). Copula functions are multivariate distributions defined on the unit hypercube $[0,1]^d$ while all univariate marginal distributions are uniform on the interval $(0,1)$. More precisely,

$$C(u_1,\ldots,u_d) = Pr(U_1 \leqslant u_1,\ldots,U_d \leqslant u_d), \tag{1.1}$$

where $U_i \sim \text{Uniform}(0,1)$ for $i=1,2,\ldots,d$. Moreover, if $(X_1,\ldots,X_d) \in \mathbb{R}^d$ has a continuous multivariate distribution with $F(x_1,\ldots,x_d) = Pr(X_1 \leqslant x_1,\ldots,X_d \leqslant x_d)$, then by Sklar's theorem (Sklar 1959) there is a unique copula function $C:[0,1]^d \to [0,1]$ of F such that

$$F(x_1,\ldots,x_d) = C(F_1(x_1),\ldots,F_d(x_d)) \tag{1.2}$$

for all $(x_1,\ldots,x_d) \in \mathbb{R}^d$ and continuous marginal distributions $F_i(x_i) = Pr(X_i \leqslant x_i)$, $x_i \in \mathbb{R}$ for $i=1,2,\ldots,d$. In addition, if C is a copula on $[0,1]^d$ and $F_1(x_1),\ldots,F_d(x_d)$ are cumulative distribution functions on $\mathbb{R}$, then $F(x_1,\ldots,x_d) = C(F_1(x_1),\ldots,F_d(x_d))$ is a cumulative distribution function on $\mathbb{R}^d$ with univariate marginal distributions $F_1(x_1),\ldots,F_d(x_d)$. Extensive theoretical discussion on copulas can be found in Joe (1997) and Nelsen (2006).

Therefore, flexible multivariate distributions can be constructed with pre-specified, discrete and/or continuous marginal distributions and copula function that represents the desired dependence structure. The joint is usually estimated by a standard two-step procedure where (i) the marginal distributions are approximated by empirical distribution, while (ii) the parameters in the copula function are estimated by maximum likelihood (Genest et al. 1995). A full two-step maximum likelihood approach is used, for instance, in Hürliman (2004) and Roch and Alegre (2006), where the parameters of the marginal distributions are estimated in the first

R.S. Silva (✉)
Australian School of Business, University of New South Wales, Sydney, 2052, Australia
e-mail: ralph.stat@gmail.com

H.F. Lopes
Graduate School of Business, University of Chicago, Chicago, 60637, USA
e-mail: hlopes@chicagogsb.edu

step and, conditional on those estimates, copula parameters are estimated.

In this paper we discuss how to conduct inference on all unknown parameters in the copulas and marginal distributions using a Bayesian approach. Such a joint estimation procedure takes into account all dependence structure of the parameters' posterior distributions in our chosen model selection criteria. Huard et al. (2006), for instance, only considered copulas selection without counting for marginal modeling. Pitt et al. (2006) proposed a Gaussian copula based regression model in a Bayesian framework with an efficient sampling scheme. We do not usually have such an efficient sampling scheme for our copula-based multivariate distributions but we can easily extend the Gaussian copula regression modeling accounting for joint estimation and model selection.

We use the deviance information criterion (*DIC*) (Spiegelhalter et al. 2002), and other related criteria, in order to select the copula-based model. It is easy to show that these criteria are invariant to monotone increasing transformations of the marginal distributions (Sect. 3.1), making them thereby particularly suitable for copula modeling and selection. We perform several Monte Carlo studies to examine the Bayesian copula selection. We consider a variety of copula functions and marginal distributions to cover a broad spectrum of joint specifications. We compare our findings with those based on the two-step procedures.

The rest of the paper is organized as follows. In Sect. 2 we outline the copula models, then we address the Bayesian estimation approach as well as the model selection criteria. Section 3 exhibits a series of Monte Carlo studies with concluding remarks given in Sect. 4.

2 Copulas and marginals

In this section we introduced several copula functions, densities and dependence measures, such as the *Kendall* measures, and outline the Bayesian inferential scheme and several model selection criteria with emphasis on the *deviance information criterion*. The most commonly used copula functions (see (1.1)) are

$$\begin{aligned}
\text{Clayton:}\quad & C(u,v|\theta) = (u^{-\theta}+v^{-\theta}-1)^{-1/\theta} \quad \text{for } \theta\in(0,\infty),\\
\text{Frank:}\quad & C(u,v|\theta) = -\frac{1}{\theta}\log\left(1+\frac{(\exp(-\theta u)-1)(\exp(-\theta v)-1)}{\exp(-\theta)-1}\right) \quad \text{for } \theta\in\mathbb{R}\backslash\{0\},\\
\text{Gaussian:}\quad & C(u,v|\theta) = \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}}\exp\left\{\frac{2\theta st-s^2-t^2}{2(1-\theta^2)}\right\} ds dt,\\
\text{Gumbel:}\quad & C(u,v|\theta) = \exp\left\{-[(-\log u)^{\theta}+(-\log v)^{\theta}]^{1/\theta}\right\}, \quad \theta\in[1,\infty),\\
\text{Heavy tail:}\quad & C(u,v|\theta) = u+v-1+[(1-u)^{-1/\theta}+(1-v)^{-1/\theta}-1]^{-\theta} \quad \text{for } \theta\in(0,\infty),\\
\text{Student-}t\text{:}\quad & C(u,v|\theta) = \int_{-\infty}^{T_\nu^{-1}(u)}\int_{-\infty}^{T_\nu^{-1}(v)} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu}{2})\nu\pi\sqrt{1-\theta^2}}\left(1+\frac{s^2+t^2-2\theta st}{\nu(1-\theta^2)}\right) ds dt
\end{aligned}$$

with $\theta\in[-1,1]$ for the Gaussian and Student-t copula functions. More discussion and properties of these copula functions can be found in Clayton (1978), Frank (1979), Gumbel (1960) and Hougaard (1986) (Gumbel and Gaussian) and Demarta and McNeil (2005) (Student-t).

The Monte Carlo study is based on single parameter copula functions and fixed number of degrees of freedom for the Student-t copula, i.e. $\nu = 4$. By doing that, we force the Gaussian and the Student-t copulas to have different properties. Additionally, the heavy tail copula is the survival Clayton copula with a simple change of parameter. More precisely, survival copulas come from the definition of the joint survival function, which in the bivariate case is $\bar{F}(x_1,x_2) = Pr(X_1 > x_1, X_2 > x_2)$, and are given by $\bar{C}(u,v) = u+v-1+C(1-u,1-v)$. Clayton, Gumbel and heavy tail copulas can model independence and positive dependence. However, the Clayton copula has lower tail dependence while the Gumbel and the heavy tail copulas have upper tail dependence. Also, Frank, Gaussian and Student-t copulas can model negative dependence, independence and positive dependence, but none has tail dependence. The bivariate tail dependence indexes are defined through conditional probabilities in the lower-quadrant and upper-quadrant tails, respectively. They capture the amount of dependence in the tails which are related to extreme values. The *lower tail index* is given by $\lambda_L = \lim_{u\to 0} Pr(U < u \mid V < u) = \lim_{u\to 0} C(u,u)/u$ and the *upper tail index* by $\lambda_U = \lim_{u\to 0} Pr(U > u \mid V > u) = \lim_{u\to 0}(1-2u-C(u,u))/(1-u)$.

When modeling joints through copula, correlation coefficient looses its meaning and a commonly used alternative for measuring dependence is the Kendall's τ (Kruskal 1958; Nelsen 2006). Kendall's τ is written in terms of the copula function as

$$\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1, \tag{2.1}$$

which is not analytically available for most copula functions. A few exceptions are, for instance, the Clayton, the Gumbel and the Gaussian copulas where it equals $\tau = \theta/(\theta+2)$, $\tau = 1-\theta^{-1}$ and $\tau = (2/\pi)\arcsin\theta$, respectively. Since Kendall's τ measures the dependence structure of the copula function, it can subsequently be used to elicit or tuning the copula parameter θ. Another important measure of dependence used in copula model is Spearman's ρ (Kruskal 1958; Nelsen 2006). However, for most copula function, Spearman's ρ does not have an analytical tractable form and are not entertained in this paper for simplicity. Other measures of dependence and association can be found in Nelsen (2006).

For the Monte Carlo study that follows, the exponential, the normal, and the skewed normal (Fernández and Steel 1998) were selected as marginal distributions. Such a choice is just to cover a broad spectrum of behavior and it is not planned to be exhaustive. Other choices like gamma, Weibull and log-normal distributions are more appropriate in certain fields of applications and we discuss this point in Sect. 6. The probability density function of the skewed normal is given by $\frac{2\gamma}{(\gamma^2+1)\sigma}\phi(\gamma(x-\mu)/\sigma)$, for $x<0$ and by $\frac{2\gamma}{(\gamma^2+1)\sigma}\phi(\gamma^{-1}(x-\mu)/\sigma)$, for $x\geq 0$. Here, ϕ is the probability density function of a standard normal distribution and $\gamma>0$.

3 Bayesian inference

Let (X_1, X_2) be a bivariate random variable with joint probability function given by

$$f(x_1,x_2\mid\Psi) = c(F_1(x_1\mid\Psi), F_2(x_2\mid\Psi)\mid\Psi)f_1(x_1\mid\Psi)f_2(x_2\mid\Psi), \tag{3.1}$$

where Ψ is the parameter vector comprising copula and marginal distribution parameters, and f_i and F_j, $j=1,2$, represent the probability and cumulative marginal distribution functions, respectively. C is a cumulative distribution function and c is a copula density, i.e. the probability density function calculated as the mixed derivative in x_1 and x_2.

Now, let $\mathbf{x} = ((x_{11},x_{21}),\dots,(x_{1n},x_{2n}))$ be a sample of size n from independent and identically distributed data from the probability function in (3.1), then the likelihood function is given by $L(\mathbf{x}|\Psi) = \prod_{i=1}^n c(F_1(x_{1i}\mid\Psi), F_2(x_{2i}\mid\Psi)\mid\Psi)f_1(x_{1i}\mid\Psi)f_2(x_{2i}\mid\Psi)$, leading to the posterior distribution $g(\Psi|\mathbf{x}) \propto L(\mathbf{x}|\Psi)g(\Psi)$ for prior distribution $g(\Psi)$.

Model specification is completed by assigning (independent) prior distributions for the components of the parameter vector with known distributions and large enough variances such that we have diffuse prior distributions for all models. By doing that, we do not add much prior information in our copula-based models used in our Monte Carlo study and hence our chosen model selection criteria rely almost entirely on the data. For instance, Gaussian and Student-t copula parameters have uniform distributions while, the Clayton copula parameter has a gamma distribution with mean one and variance 10^6. For the marginal distributions we use standard proper priors such as normal prior for locations parameters, inverted gamma priors for scale parameters, and gamma priors for skewness parameters. These choices do not necessarily suggest any specific agenda, are used for their computational tractability and do not affect our overall conclusions. It is worth mentioning that prior information can also be placed on copula models themselves, but this is not pursued in this paper.

As copula functions and densities have usually nonstandard forms, copula-based bivariate distributions lead in most cases to analytically intractable posterior distributions and customized Markov chain Monte Carlo (MCMC) scheme are necessary (Gamerman and Lopes 2006) in order for posterior inference and copula selection be performed. In other words, most full conditionals draws are obtained from slice sampling (Neal 2003) yielding an MCMC algorithm with good sampling properties, which means good mixing, low autocorrelations and fast convergence rate. We verify these properties empirically by graphical analysis and do not show them here for conciseness.

4 Copula-based model selection

Most of the current literature focuses on the estimation and selection of copula functions conditionally on the first step estimation, i.e. conditionally on the estimated marginal distributions. In this section we summarize alternative measures of model adequacy, determination and/or selection that can be used in copula-based distributions. To begin with, let $L(\mathbf{x}|\Psi_k,\mathcal{M}_k)$ be the likelihood function for model $\mathcal{M}_k$, comprising the univariate marginal distributions and the copula function or density and define the deviance function, $D(\Psi_k) = -2\log L(\mathbf{x}|\Psi_k,\mathcal{M}_k)$. The Akaike information criterion (*AIC*), the Bayesian information criterion (*BIC*), their expected versions, *EAIC* and *EBIC*, and the deviance information criterion (*DIC*) are defined as $AIC(\mathcal{M}_k) = D(E[\Psi_k|\mathbf{x},\mathcal{M}_k]) + 2d_k$, $BIC(\mathcal{M}_k) = D(E[\Psi_k|\mathbf{x},\mathcal{M}_k]) +$

$\log(n)d_k$, $EAIC(\mathcal{M}_k) = E[D(\Psi_k)|\mathbf{x}, \mathcal{M}_k] + 2d_k$, $EBIC(\mathcal{M}_k) = E[D(\Psi_k)|\mathbf{x}, \mathcal{M}_k] + \log(n)d_k$ and

$$DIC(\mathcal{M}_k) = 2E[D(\Psi_k)|\mathbf{x}, \mathcal{M}_k] - D(E[\Psi_k|\mathbf{x}, \mathcal{M}_k]), \tag{4.1}$$

respectively, with d_k representing the number of parameters of the $\mathcal{M}_k$ model. Further details on all these measures can be found in Spiegelhalter et al. (2002). Suppose that $\{\Psi_k^{(1)}, \ldots, \Psi_k^{(L)}\}$ corresponds to a sample from the posterior distribution $g(\Psi_k|\mathbf{x}, \mathcal{M}_k)$. Then, $L^{-1}\sum_{\ell=1}^{L} D(\Psi_k^{\ell})$ and $L^{-1}\sum_{\ell=1}^{L} \Psi_k^{\ell}$, are Monte Carlo approximations to $E[D(\Psi_k)|\mathbf{x}, \mathcal{M}_k]$ and $E[\Psi_k|\mathbf{x}, \mathcal{M}_k]$; and approximations to *DIC*, *AIC*, *BIC*, *EAIC* and *EBIC* can be straightforwardly derived.

An important characteristic shared by these measures is invariance under monotone increasing transformations of the marginal distributions; a desirable feature when dealing in copula modeling. For instance, if the *DIC* selects a particular copula function with, say, normal marginal distributions, it will also select the same copula function with log-normal marginal distributions for exponentiated observations. For the Monte Carlo study discussed below, reported values of *AIC* are close to *EAIC* (the same is true for *BIC* and *EBIC*). Also, graphical inspections showed that most of the posterior distributions can be considered close to the multivariate normal distribution, assuring reasonable properties to the *DIC*. Besides being easily to numerically implement, the above criteria provide interesting advantages over commonly used statistical tests since they take into account parameter dependence. Furthermore, they can be applied to any copula family as long as the copula density can be computed. For sake of space, solely results based on *DIC* are reported (Spiegelhalter et al. 2002).

5 Monte Carlo study

This section illustrates the Bayesian approach when jointly estimating and selecting marginal distributions and copula functions as copula-based distributions. Should the copula-based distributions select similar marginal distributions (e.g. normals), then joint estimation can be regarded as a pure copula selection procedure. The studied copula functions and marginal distributions were introduced in Sect. 2, such that all bivariate distributions exhibit the same marginal distributions, i.e. both marginal distributions are normally, exponentially or skewed normally distributed. We entertain several sample sizes as well as several degrees of dependence along with copula functions that have been widely used. The copula functions are chosen to have similar properties, such as the pairs (Gumbel, heavy tail) and (Gaussian, Student-t). Others are chosen to stand out, such as the Clayton copula. One thousand replications of three different artificial data sizes, namely $n = \{100, 200, 500\}$, were performed for two Kendall's τ dependence measure with, $\tau = \{1/3, 2/3\}$.

We use the Kendall τ values to define the respective copula parameter values, e.g. for $\tau = 2/3$, $\theta = 4$ for the Clayton copula, $\theta = 10$ for the Frank copula and $\theta = 0.865$ for the Gaussian copula. For the marginals, we chose standard normal distributions, exponential distributions with mean one, and skew-normal distributions with location parameter zero, scale parameter one and shape parameter equal 3/2.

For the Gaussian copula with normal marginals, draws are directly obtained from the bivariate normal distribution, while an inverse method was used to draw from the Frank copula (Nelsen 2006). Sampling importance resampling techniques (Gamerman and Lopes 2006), were used for all other cases and the resampling rate was kept around five per cent. Posterior inference and model selection are based on a MCMC algorithm that presented fast convergence rate and low autocorrelation for all cases. Only fifteen hundred draws were simulated and inference/selection based on the last thousand draws.

Table 1 summarizes the results of the simulation by presenting the percentage of true copula-based distributions correctly identified by the deviance information criterion.

For $\tau = 1/3$ and $n = 100$, for instance, the Clayton copula and exponential marginals combination is correctly identified 894 times out of the 1 000 replications. Results based on *AIC*, *EAIC*, *BIC* and *EBIC* lead to the roughly similar conclusions and are omitted for conciseness. The correct models are correctly selected for larger sample size or Kendall's τ. Moreover, copulas with similar behavior and small dependence measure τ may lead to poor selection of the respective copula-based distributions, mainly for small and moderate sample sizes. As expected the Clayton copula-based distributions correctly selected most of the time since it induces rather unique bivariate distributions. For instance, the Clayton copula has lower tail dependence while the Gumbel and heavy-tail copulas have upper tail dependence. Finally, regarding model selection as copula selection, results remain quite similar across the three marginal distributions, which corroborates with findings from Huard et al. (2006). Figure 1 complements Table 1 by discriminating amongst alternative fitted copula-based distributions. Even for a sample size of $n = 500$ the Gaussian and Student-t copula-based distributions misclassify one another. Even though the Student-t copula-based distribution with four degrees of freedom is rather different from the Gaussian one, this result shows that it can be difficult to select the best copula-based model when copulas have similar properties, e.g. tail dependence and Kendall's τ.

Table 1 Percentage of correct copula-based model choice based on *DIC* over one thousand replications, three sample sizes (n), six copulas functions C (Clayton, Frank, Gaussian, Gumbel, heavy-tail and Student-t), three marginal densities f (exponential (E), normal (N) and skewed-normal (SN)) and two degrees of dependence (τ)

C	f	$\tau = 1/3$			$\tau = 2/3$		
		$n = 100$	$n = 200$	$n = 500$	$n = 100$	$n = 200$	$n = 500$
Clayton	E	89.4	98.1	99.8	99.9	100.0	100.0
	N	91.4	97.3	99.5	100.0	100.0	100.0
	SN	76.6	91.4	96.1	99.8	100.0	100.0
Frank	E	64.2	80.0	94.1	95.8	99.4	100.0
	N	64.8	79.2	94.1	94.5	99.2	100.0
	SN	59.5	80.5	87.2	94.6	98.6	100.0
Gaussian	E	59.3	75.3	84.4	93.0	93.2	99.6
	N	55.7	75.1	83.9	93.6	94.3	99.6
	SN	47.9	69.5	73.2	93.6	96.1	98.7
Gumbel	E	39.2	63.2	84.2	88.8	97.0	100.0
	N	40.0	63.1	84.3	88.3	96.6	100.0
	SN	34.1	59.0	62.1	87.1	90.2	99.5
Heavy tail	E	78.5	85.5	96.9	93.6	99.8	100.0
	N	77.7	82.3	94.1	94.4	99.5	100.0
	SN	73.9	83.0	94.8	92.3	97.8	100.0
Student-t	E	63.6	73.8	84.1	89.7	98.2	99.4
	N	62.6	76.1	83.1	91.0	98.3	99.4
	SN	56.8	65.4	83.2	88.1	97.5	98.9

5.1 One or two-step?

In a second simulation study we compared the fully Bayesian and the two-step estimation approaches. We argue that posterior dependence between marginal parameters is essentially controlled by the copula parameters.

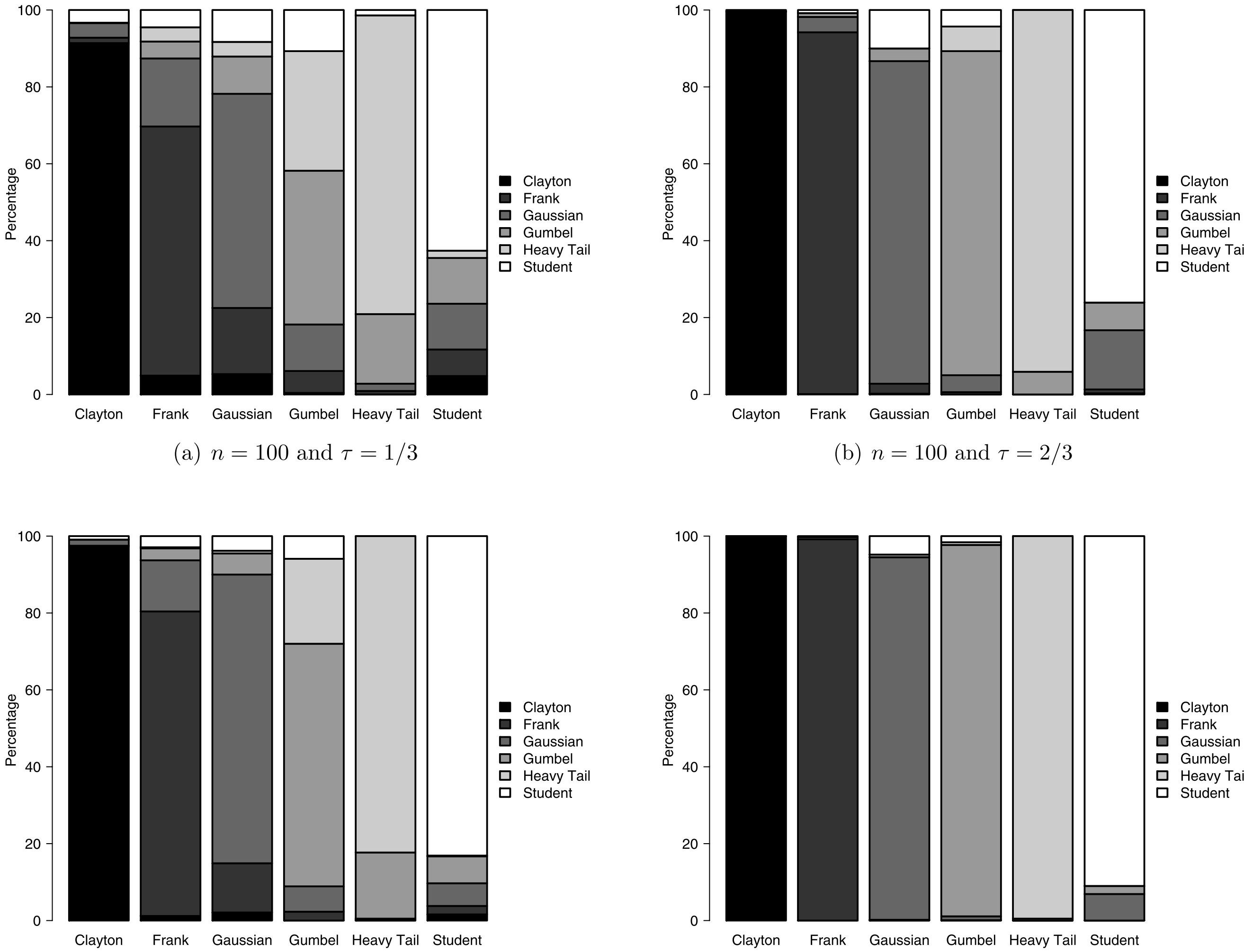
Let us consider a bivariate distribution given by the Gumbel copula with parameter θ and $\mathbf{x}_j = (x_{j1}, \ldots, x_{jn})$ ($j = 1, 2$) random samples from exponential marginal distributions with parameters λ_1 and λ_2, respectively. The prior distribution of $(\theta, \lambda_1, \lambda_2)$ is $g(\theta, \lambda_1, \lambda_2) = g(\theta)g(\lambda_1, \lambda_2)$, where $\theta \sim \text{Gamma}(\epsilon, \epsilon)$ and $g(\lambda_1, \lambda_2) \propto 1$, with $\epsilon = 10^{-10}$, for example. In this case the posterior distribution is given by

$$
\begin{aligned}
&g(\theta, \lambda_1, \lambda_2|\mathbf{x})\\
&\quad\propto (\lambda_1\lambda_2)^{(n+1)-1} \exp\{-\lambda_1 n\bar{x}_1 - \lambda_2 n\bar{x}_2\}\\
&\qquad \times \prod_{i=1}^{n} c_\theta(1 - \exp\{-\lambda_1 n x_{1i}\}, 1 - \exp\{-\lambda_2 n x_{2i}\}),
\end{aligned}
\tag{5.1}
$$

where c_θ is the Gumbel copula density (Sect. 2). Under a two-step estimation procedure, the posterior distributions of λ_j is $\text{Gamma}(n+1, n\bar{x}_j)$ with mode $\tilde{\lambda}_j = 1/\bar{x}_j$, for $n\bar{x}_j = \sum_{i=1}^{n} x_{ji}$ and $j = 1, 2$.

Another Monte Carlo study is presented in Table 2 to show the parameter dependence in the posterior distribution when two- and one-step point estimation approaches are employed. Once more, one thousand replications of three different artificial and relatively small data sizes, namely $n = \{20, 30, 50\}$, were performed for three Kendall's τ dependence measure with, $\tau = \{1/3, 2/3, 0.9\}$. Table 2 shows the mean, Kendall's τ and the square root of the mean square error (in parentheses) of the parameters' posterior modes for the Gumbel copula and exponential marginals with different two- and one-step estimations. The two-step estimation is based on the posterior mode of the exponential marginal models in the first step, and then on the posterior mode of the copula parameter conditionally on the first step. We also estimate the copula parameter in another two-step procedure by fitting empirical distribution functions to the marginals and maximum likelihood to the copula model, and the copula parameter is also estimated using the sample version of Kendall's τ. Finally, the one-step estimation is based on numerical methods to find the joint posterior mode.

The results shown in Table 2 indicate that two-step and one-step point estimations lead to the same results on average. However, the measure Kendall's $\tilde{\tau}$ over all 1 000 replication estimations of λ_1 and λ_2 shows the existence of the parameter dependence, as it is the case of the posterior distribution of a given data set. Such dependence is determined mainly by the copula parameter. Values of Kendall's $\tilde{\tau}$ are bigger than those used in the data generating process. Moreover, the results based on posterior modes in two-step are



(a) $n = 100$ and $\tau = 1/3$

(b) $n = 100$ and $\tau = 2/3$

(c) $n = 200$ and $\tau = 1/3$

(d) $n = 200$ and $\tau = 2/3$

Fig. 1 Percentage of correctly identified copula-based distributions based on the *DIC* criterion, over 1000 replications, six copula functions, two degrees of dependence, τ, and two sample sizes, n. Both marginal distributions are standard normals. The abscissaes correspond to the data sets and the legends correspond to the best-fitting copula-based distribution

better than those using the empirical distribution function or the sample version of Kendall's τ. This result is expected since both are non parametric approaches and convergence to the true model is slower than the parametric approach. We stress here that the marginal distributions in the joint modeling given by (5.1) are no longer exponential even though numerical and graphical analysis indicates that they are close to exponential distributions. Moreover, these results are similar on those based on maximum likelihood estimation due to our prior distribution choice.

In Table 3 we present some results based on a two-step estimation procedure for the same data sets of the six copulas and standard normal marginals presented in Table 1. First, the empirical distribution functions are fitted to all marginals, then the copula parameters are estimated by the posterior mean with the same vague prior distributions as before. Once again, *DIC* and all related criteria are calculated and comparisons of successful copula-based model identifications are performed. Conditioning on the first step of the two-step estimation procedure, then model selection is primarily copula selection. It can be seen that in most cases the percentages of this two-step estimation procedure are smaller when compared with the joint one showed in Table 1, which in turn has the advantage of parameter dependencies and covariance estimations. However, in some cases identification rates are bigger for this two-step procedure than the one step when $\tau = 1/3$. This fact is due to the non parametric estimation, low dependence structure and probably on the behavior of some copula functions. But the results on the previous example indicates to use at least two-step estimation procedure based on the posterior distributions or equivalently on two-step maximum likelihood estimation.

Table 2 Mean, Kendall's τ and square root of the mean square error (in parentheses) of the parameters' posterior modes for Gumbel copula and exponential marginals over 1 000 replications. Values in the row Υ are those used in the data generating processes. Three degrees of dependence were chosen by $\tau = \{1/3, 2/3, 0.9\}$. Here, *PM* stands for posterior mode, *EDF* for empirical distribution function and *KD* for Kendall's rank correlation

n	Two-step						One-step			
	$\tilde{\lambda}_1^{PM}$	$\tilde{\lambda}_2^{PM}$	$\tilde{\tau}$	$\tilde{\theta}^{PM}$	$\tilde{\theta}^{EDF}$	$\tilde{\theta}^{KD}$	$\tilde{\lambda}_1^{PM}$	$\tilde{\lambda}_2^{PM}$	$\tilde{\tau}$	$\tilde{\theta}^{PM}$
Υ	0.5	2.0	1/3	1.5	1.5	1.5	0.5	2.0	1/3	1.5
20	0.525	2.103	0.410	1.573	1.684	1.582	0.526	2.112	0.412	1.579
	(0.127)	(0.509)	–	(0.333)	(0.415)	(0.400)	(0.126)	(0.513)	–	(0.334)
30	0.518	2.083	0.366	1.548	1.622	1.542	0.519	2.089	0.386	1.552
	(0.098)	(0.380)	–	(0.246)	0.288	(0.279)	(0.097)	(0.381)	–	0.247
50	0.504	2.026	0.435	1.536	1.584	1.534	0.504	2.027	0.422	1.539
	(0.075)	(0.292)	–	(0.197)	(0.219)	(0.217)	(0.073)	(0.286)	–	(0.197)
Υ	0.5	2.0	2/3	3.0	3.0	3.0	0.5	2.0	2/3	3.0
20	0.522	2.101	0.723	3.207	3.386	3.349	0.525	2.114	0.737	3.225
	(0.123)	(0.499)	–	(0.734)	(1.040)	(1.215)	(0.121)	(0.492)	–	(0.733)
30	0.522	2.101	0.719	3.207	3.386	3.349	0.525	2.114	0.726	3.225
	(0.123)	(0.499)	–	(0.734)	(1.040)	(1.215)	(0.121)	(0.492)	–	(0.733)
50	0.507	2.032	0.712	3.102	3.142	3.131	0.507	2.033	0.731	3.112
	(0.073)	(0.300)	–	(0.440)	(0.518)	(0.572)	(0.073)	(0.296)	–	(0.443)
Υ	0.5	2.0	0.9	10.0	10.0	10.0	0.5	2.0	0.9	10.0
20	0.529	2.113	0.921	10.538	9.760	12.221	0.531	2.122	0.927	10.611
	(0.130)	(0.519)	–	(2.565)	(5.372)	(8.408)	(0.129)	(0.514)	–	(2.556)
30	0.517	2.067	0.920	10.339	9.345	11.044	0.517	2.070	0.922	10.394
	(0.096)	(0.388)	–	(1.950)	(2.809)	(4.013)	(0.095)	(0.383)	–	(1.931)
50	0.511	2.044	0.917	10.238	9.331	10.485	0.512	2.048	0.924	10.271
	(0.075)	(0.299)	–	(1.516)	(1.903)	(2.299)	(0.073)	(0.292)	–	(1.511)

Table 3 Percentage of correctly identified copula-based distributions based on the *DIC* criterion for the two-step estimation, over 1 000 replications, six copula functions, two degrees of dependence, τ, and two sample sizes, n. In the two-step estimation procedure, both marginal distributions were empirically estimated, followed by Bayesian copula function parameters estimation

Copula	$\tau = 1/3$			$\tau = 2/3$		
	$n = 100$	$n = 200$	$n = 500$	$n = 100$	$n = 200$	$n = 500$
Clayton	92.2	97.8	100.0	98.4	100.0	100.0
Frank	70.0	81.3	94.4	92.7	98.2	100.0
Gaussian	51.1	69.8	93.0	63.4	86.3	99.2
Gumbel	30.9	55.4	85.5	52.3	76.7	97.4
Heavy tail	59.1	71.6	89.0	70.6	90.6	99.7
Student-*t*	54.0	80.0	97.9	62.5	86.5	99.0

6 Concluding remarks

Copula functions are powerful tools for constructing multivariate distributions and studying dependency in general. Two-step point estimation procedures are widely used in the literature, leading to the same results as a one-step procedures based on posterior mode or maximum likelihood estimation. However, from a Bayesian perspective it is desirable to jointly estimate all parameters so that a complete characterization of the posterior distribution and hence the dependence among the parameters can be constructed. Moreover, this dependence structure should be taken into account when model selection criteria are employed. We conjecture that copulas will be widely used in a few years in more structured models like hierarchical ones, where dependence among parameters can give different results in two- or one-step estimations. The results of our analysis show that the *DIC* is a good model selection criterion. We also point out that such results are corroborated by *AIC*, *EAIC*, *BIC* and *EBIC*, but not presented here for conciseness. These criteria are invariant to monotone increasing transformations of the marginal distributions and easy to include in MCMC routines.

Our results are limited as we only use a small set of one-parameter bivariate copula functions, a selection of sample sizes and degrees of dependence. But our results indicate that model selection based on the deviance function are well suited for a range applications in copula modeling. Our experience shows that for low degrees of dependence copula-based distributions can be hard to identify even from relatively big samples. Moreover, estimation methods based on the likelihood function give better results than some non parametric ones. The latter has the advantage of being parameter free but has a low convergence rate to the true model when compared to the former. Moreover, model choices would not change as long as the same prior information is introduced in all models, for instance using the same measure of dependence for copula parameters. However, prior distribution plays an important role in small sample sizes and in this case a careful study in necessary. Our experience also shows that model choice does not change as long as the same marginal distributions are specified for all copula-based distributions even if these marginals are misspecified.

Finally, if the reader is only interested in point estimation of simple copula models like those shown in this paper, we would recommend a two-step estimation approach such as a Bayesian one or maximum likelihood estimation. This is especially useful for high dimensions since a complete search over the parameter space can be time consuming. On the other hand, from a Bayesian point of view, we would recommend to jointly estimate all unknown parameters, thereby taking into account parameter dependence. Such an approach is also beneficial in the model selection stage.

Acknowledgements We are grateful to both referees and associate editor for invaluable comments which led to significant improvement of the paper. The first author is also grateful to the Brazilian CAPES foundation for financial support.

References

Akaike, H.: Information theory and an extension of the maximum likelihood principle. In: Second Int. Symp. on Inf. Theory (Tsahkadsor, 1971), pp. 267–281. Akadémia Kiadó, Budapest (1973)

Clayton, D.: A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika **65**, 141–151 (1978). doi:10.1093/biomet/65.1.141

Demarta, S., McNeil, A.J.: The *t* copula and related copulas. Int. Stat. Rev. **73**(1), 111–129 (2005)

Fernández, C., Steel, M.: On Bayesian modelling of fat tails and skewness. J. Am. Stat. Assoc. **93**, 359–371 (1998). doi:10.2307/2669632

Frank, M.J.: On the simultaneous associative of $f(x,y)$ and $x+y-f(x,y)$. Aequ. Math. **19**, 194–226 (1979). doi:10.1007/BF02189866

Gamerman, D., Lopes, H.F.: Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2nd edn. Chapman & Hall, London (2006)

Genest, C., Quessy, J.-F., Rémillard, B.: Goodness-of-fit procedures for copula models based on the probability integral transformation. Scand. J. Stat. **32**(2), 337–366 (2006). doi:10.1111/j.1467-9469.2006.00470

Genest, C., Ghoudi, K., Rivest, L.-P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika **82**(3), 543–552 (1995). doi:10.1093/biomet/82.3.543

Gumbel, E.J.: Distributions des valeurs extrême en plusiers dimensions. Publ. Inst. Stat. Univ. Paris **9**, 171–173 (1960)

Hougaard, P.: A class of multivariate failure time distributions. Biometrika **73**, 671–678 (1986). doi:10.1093/biomet/73.3.671

Huard, D., Évin, G., Favre, A.-C.: Bayesian copulas selection. Comput. Stat. Data Anal. **51**(2), 809–822 (2006). doi:10.1016/j.csda.2005.08.010

Hürliman, W.: Fitting bivariate cumulative returns with copulas. Comput. Stat. Data Anal. **45**(2), 355–372 (2004). doi:10.1016/S0167-9473(02)00346-8

Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)

Kruskal, W.H.: Ordinal measures of association. J. Am. Stat. Assoc. **53**, 814–861 (1958). doi:10.2307/2281954

Neal, R.: Slice sampling (with discussion). Ann. Stat. **31**, 705–767 (2003). doi:10.1214/aos/1056562461

Nelsen, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)

Pitt, M., Chan, D., Kohn, R.: Efficient Bayesian inference for Gaussian copula regression models. Biometrika **93**, 537–554 (2006). doi:10.1093/biomet/93.3.537

Roch, O., Alegre, A.: Testing the bivariate distribution of daily equity returns using copulas. An application to the Spanish stock market. Comput. Stat. Data Anal. **51**(2), 1312–1329 (2006). doi:10.1016/j.csda.2005.11.007

Schwarz, G.: Estimating the dimension of a model. Ann. Stat. **6**, 461–464 (1978). doi:10.1214/aos/1176344136

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

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.: Bayesian measures of model complexity and fit (with discussion). J.R. Stat. Soc. Ser. B **64**, 583–639 (2002). doi:10.1111/1467-9868.00353