Analysis of Dependency Structure of Default Processes Based on Bayesian Copula

Mohammad Seidpisheh\textsuperscript{1,2}, Armin Pourkhanali\textsuperscript{3}, Karim Norouzipour\textsuperscript{3}, Adel Mohammadpour\textsuperscript{1}

\textsuperscript{1}Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran.  
\textsuperscript{2}Department of Statistics, Allameh Tabataba’i University, Tehran, Iran.  
\textsuperscript{3}Department of Mathematics, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran.

Abstract. One of the main problems in credit risk management is the correlated default. In large portfolios, computing the default dependencies among issuers is an essential part in quantifying the portfolio’s credit. The most important problems related to credit risk management are understanding the complex dependence structure of the associated variables and lacking the data. This paper aims at introducing a new methodology for credit risk management based on Bayesian copulas. In this paper, the focus is specifically on a new method of simulating the joint distribution of default risk. This methodology joins the use of copulas and Bayesian models. Using copulas, the joint multivariate probability distribution of a random vector can be separated into individual components characterized by marginal distributions. The model is based on a jump diffusion process for the intensities. Another important problem in credit risk management is the lack of data, which influences the parameter estimation. Considering this drawback, the employment of Bayesian methods and simulation tools could be a natural solution to the problem. This suggests the use of Bayesian models, computed

Mohammad Seidpisheh(m84seidpisheh@yahoo.com), Armin Pourkhanali(a.pourkhanali@gmail.com), Karim Norouzipour(k.norozipour@iasbs.ac.ir), Adel Mohammadpour(adel@aut.ac.ir)

Received: December 2012; Accepted: August 2013
via simulation methods and in particular, Markov chain Monte Carlo. Bayesian methods in Student’s $t$ copula are efficient enough for heavy tail distribution. Moreover, our main outcome is that the application of Bayesian methodology causes a reduction of measure while that copula is Student’s $t$. Finally, the conclusion of Bayesian copulas with classic copulas was compared through a simulation study.

**Keywords.** Bayesian copula, credit risk, jump diffusion process, Markov chain Monte Carlo.

**MSC:** 62C10, 60E05, 91G40, 65C05.

## 1 Introduction

Financial management systems especially credit risk is one of the main issues in today’s world. Credit risk is the risk of changes in value associated with unexpected changes in credit quality (Duffie and Singleton [11]). To measure the credit risk, the probability of default is one of the usual methods. The history of financial institutions has shown that many failures of banking associations were due to dependent defaults. As a result, the analysis of dependency among defaults in risk investigation of loan portfolio will be very important. The complex structure of the credit losses critically depends on the dependencies between default events. To estimate default risk in the portfolio, both the individual default rates of each firm and dependency structure probability of defaults across all firms need to be considered (more about the financial importance of default dependence see, Giesecke [15] and Lucas [19]).

Rating agencies now play a crucial role in determining the return on bonds and the cast for issuers. Issuer ratings (PD (Probability of Default) ratings) focus on the ability of a borrower to honor its obligations promptly. Many popular approaches exist to compute PDs in the marketplace, developed by firms such as KMV Corporation, Moodys Risk Management Systems (MRMS), CreditMetrics™ model etc.

There are three main approaches to simulate the joint distribution of dependent probability of defaults: using credit market historical data (Lucas [19]), using Structural models (Das and Tufano [8]) and using reduced form models (also called Intensity-based models). For more details see Jarrow and Turnbull [17], Madan and Unal [20] and Das and Sundaram [7]. There are several settings for intensity. One simple setting is the Poisson process with constant positive intensity that was extended by Jarrow and Turnbull [17]. These traditional generalizations
allow the default intensities to be time dependent and these became stochastic default intensities. In reduced models, it is assumed that the default intensities are exponential distribution. To generalize the basic models, Duffie and Singleton [10] developed a basic model in that it was assumed default intensities varies randomly during the time that it is called doubly stochastic intensity. There are various stochastic processes for intensities such as CIR processes (Cox, Ingersoll and Ross [2]).

There are several ways to investigate the intensity approach of credit risk: models of contagion, doubly stochastic correlated default intensity process and copula functions. In the simplest approach it is assumed that there is a dependency among defaults because of dependency among the probability of defaults in various firms. This dependency is the result of common economic factors that affect on these probabilities. One of the usual approaches in the analysis of dependent probability of defaults is the copula function approach.

Das and Geng [6] simulated joint probabilities of ”default for the U.S.” corporations using credit ratings data for copula functions, which has been developed less in this paper. For more details about this approach see Schonbucher and Schubert [24], Yu [25] and Frey et al. [13].

In reduced form models, default probabilities are usually expressed as intensities, which we show as $\lambda_i(t), i = 1, \ldots, N$. The intensities for all $N$ issuers vary over time. The survival probabilities over a horizon $T$ are shown as $s_i(T) = E\left\{\exp \left[ -\int_0^T \lambda_i(t) \, dt \right] \right\}$ and the probability of default is therefore $PD_i(t) = 1 - s_i(t)$. At the rating level, a jump-diffusion model is chosen for the average intensity of the class. This approach has been suggested by Das and Geng [6], that their focus is on the classical method to estimate parameters of copulas. But we extended it to a mixture of Gaussian marginal distributions for the residual of default processes and moreover, to estimate parameters of copulas we used classical and Bayesian methods. In many cases there is a dependency between two stochastic processes $X(t)$ and $Y(t)$. To study these processes, simulation is inevitable. So using some alterations, first we changed the residuals to time independent ones and second using copula we simulated the residuals of these processes and at last using simulated residuals, we studied the behavior of processes.

This paper aims to show the efficiency of the copula method in modeling correlated default. There are some useful features of the analysis for modelers of portfolio credit risk. For instance simulation model, based on estimating the joint system of over 200 issuers is able to replicate the
empirical joint distribution of default. Also the simulation approach is fast, efficient and allows the rapid generation of scenarios to assess risk in credit portfolios. For more details about copulas, see Joe [18], Nelsen [22], Denuit et al. [9] or Mari and Kotz [21], and for more on the use of copulas in credit risk modeling, see Schonbucher and Schubert [24], Frey et al. [13], Cherubini et al. [1] and Durante and Sempi [12].

So it was shown that the application of Bayesian methodology causes a great reduction of measure when copula is Student’s $t$. But the deficiency of the Bayesian method in normal copula compared to the obvious advantage in Student’s $t$ is that the reduction of measure is negligible.

The rest of the paper was organized as follows: Section 2 describes the dataset and Section 3 focuses on copula functions. Section 4 describes the model proposed, Section 5 reports the results of correlated defaults simulation and comparisons, and the last section expresses conclusion.

## 2 Data Description

Our study on default risk is in hundreds active corporations in Tehran Stock Exchange. The Tehran Stock Exchange database was used to develop a parsimonious numerical method of modeling and simulating correlated default processes for hundreds of issuers. The empirical examination of the joint stochastic process of default risk was carried out during the period of 1999-2011, and for each issuer, we had PDs based on their econometric models for every semester. The algorithm of Hartigan and Wong [16] was used for partitioning the data into resembling classes to produce an operational assortment. This method was run using R software (k-means function). Firms were clustered by the k-means method, which aims to partition firms into six groups such that the sum of squares from firms to the assigned cluster centers is minimized. Some issuers fall into rating class 7, which comprises unrated issuers, and the PDs within this class range from high to low. PDs from rating class 7 were not considered. Table 1 reports empirical averages and standard deviations of our data from the first rating class to the sixth rating class. It is obvious from Table 1 that the mean increases from the first rating class to the sixth rating class, as the standard deviation. According to Table 2, Kendall’s $\tau$ of default probability of each rating class shows the dependence between rating classes, therefore it needs to joint distribution function of rating classes.
Table 1: The results of the time series of average PDs for each rating class. Mean and standard deviation (StdDev) of default probability explained by the common component of default probability are reported both for default probability levels and changes (PD in changes(t) = PD in levels(t + 1) - PD in levels(t) = ΔPD in levels(t)).

<table>
<thead>
<tr>
<th>Rating class</th>
<th>PD in levels</th>
<th>PD in changes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Dev</td>
</tr>
<tr>
<td>1</td>
<td>0.0034</td>
<td>0.0104</td>
</tr>
<tr>
<td>2</td>
<td>0.0081</td>
<td>0.0609</td>
</tr>
<tr>
<td>3</td>
<td>0.0201</td>
<td>0.0928</td>
</tr>
<tr>
<td>4</td>
<td>0.0410</td>
<td>0.1131</td>
</tr>
<tr>
<td>5</td>
<td>0.0514</td>
<td>0.1975</td>
</tr>
<tr>
<td>6</td>
<td>0.0792</td>
<td>0.2160</td>
</tr>
</tbody>
</table>

Table 2: Computed Kendall’s τ for default probability of each rating class. Note that the upper right triangle shows the dependence for PD levels. Moreover, the lower left triangle expresses the dependence for PD changes.

<table>
<thead>
<tr>
<th>Rating class</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td>-0.2097</td>
<td>-0.1232</td>
<td>-0.0194</td>
<td>-0.0930</td>
<td>-0.0978</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1.0000</td>
<td>-0.0209</td>
<td>0.0109</td>
<td>0.3385</td>
<td>0.1094</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>1.0000</td>
<td>-0.3648</td>
<td>0.4534</td>
<td>0.1931</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>0.4510</td>
<td>0.3893</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
<td>0.4103</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0000</td>
</tr>
</tbody>
</table>

3 Copula Functions

Now we define a statistical tool widely and especially used in the financial field, that allows us to express the dependence structure of a vector of variables: the copula. A copula is a statistical tool which has been recently used in finance and engineering to build flexible joint distributions in order to model a high number of variables. Consider $X_1, \ldots, X_r$. 
to be random variables and $H$ as their joint distribution function. Then we have the following definition.

**Definition 3.1.** A $r$-dimensional copula is a function $C : [0, 1]^r \to [0, 1]$ with the following properties:

1. For all $(u_1, \ldots, u_r) \in [0, 1]^r$, then $C(u_1, \ldots, u_r) = 0$ if at least one coordinate of $(u_1, \ldots, u_r)$ is 0;
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$, for all $u_i \in [0, 1], (i = 1, \ldots, r)$;
3. $C$ is $r$-increasing, (see Nelsen [22], Definition 2.10.2).

Sklar’s theorem clarifies the role that copulae play in the relationship between multivariate distribution functions and their univariate margins.

**Theorem 3.1 (Sklar’s theorem).** Let $H$ be a joint distribution function with margins $F$ and $G$. Then there exists a copula $C$ such that for all $x, y$ in $\mathbb{R}$,

$$H(x, y) = C(F(x), G(y)).$$

(1)

If $F$ and $G$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $\text{Ran} F \times \text{Ran} G$. Conversely, if $C$ is a copula and $F$ and $G$ are distribution functions, then the function $H$ defined by (1) is a joint distribution function with margins $F$ and $G$ (Nelsen [22]).

Hence, if $C$ is a copula, then it is the distribution of a multivariate uniform random vector, as it was stated in Sklar’s theorem and in the corollary derived by Nelsen in 1999, see Nelsen [22]. A copula is thus a function that, when applied to marginal distributions, results in a proper multivariate probability distribution function. Since this pdf embodies all the information about the random vector, it contains all the information about the dependence structure of its components. Hence by implementing this technique, we split the distribution of a random vector into individual components (marginal) with a dependence structure (the copula) without losing any information. In this paper, the normal and the Student’s $t$ copula are applied. These two types of copulas belong to the class of elliptical copulas. Elliptical copulas are the copulas of elliptical distributions. Archimedean copulae are an alternative to Elliptical copulae. However, to model only positive dependence (or only partial negative dependence), they present the serious limitations while their multivariate extension involve strict restrictions on bivariate dependence parameters. This is why we do not focus on them here.
4 Model Description

Our data were comprised of firms that were categorized into six classes. After using the clustering method for categorizing the data in six rating classes, an attempt was made to study the dependency structure of default for these classes. We averaged across firms within a rating class to obtain a time series of the average intensity $\lambda_k$ for each $k$ rating class. We assumed that the stochastic processes for the six averages $\lambda_k$s are drawn from a joint distribution characterized by a copula, which establishes the joint dependence between rating classes. For the interpretation of the model and simulation of the joint default process, we also needed the following structure. The rest of the section was organized as follows: Subsection 4.1 introduces a jump diffusion process and Subsection 4.2 explains parameter estimation of marginal distributions and copulas.

4.1 Estimation of the average of each rating class using a Jump Diffusion process

We computed the individual intensity as:

$$\lambda_{kj}(t) = -\log(1 - PD_{kj}(t)), \quad j = 1, \ldots, M_k, \quad k = 1, \ldots, N, \quad t = 1, \ldots, T.$$  

The component $M_k$ describes the total number of issuers within the rating class for which data are available in the $k$th rating class. Moreover, we assume that $N = 6$ and data size of sample is $T$ (41 sample). Let $\lambda_k(t)$ be the average intensity across rating class $k$ at time $t$. Therefore,

$$\lambda_k(t) = \frac{1}{M_k} \sum_{j=1}^{M_k} \lambda_{kj}(t), \quad t = 1, \ldots, T.$$  

We are now prepared to compute the $\lambda_k$. Let $\lambda_k(t)$ follows the stochastic process below:

$$\Delta \lambda_k(t) = \kappa_k(\theta_k - \lambda_k(t)) \Delta t + X_k(t), \quad (2)$$

where

$$X_k(t) = \epsilon_k(t) + J_k(t)L_k(q_k, t),$$

$$\epsilon_k \sim N(0, \sigma_k^2), \quad J_k \sim N(\mu_k, \delta_k^2),$$

and

$$L_k(q_k(t), t) = \begin{cases} 
1 & \text{with probability } q_k \\
0 & \text{with probability } 1-q_k 
\end{cases}$$
Parameters of the jump diffusion process were estimated via Maximum Likelihood Estimation. The jump-sizes were considered to be normal distributed (mean $\mu_k$ and variance $\delta_k^2$). Here, $\kappa_k$ is a parameter controlling the speed of mean reversion of $\lambda_k$. Moreover $\theta_k$ is the level of mean reversion. It was also assumed that $J_k$, $L_k$ and $\epsilon_k$ are independent. Then residuals $X_k(t)$ had a mixture of two normal components. We can depict density function for the residual term $X_k(t)$ into the following:

$$f[x_k(t)] = q_k f[x_k(t)|L_k = 1] + (1 - q_k) f[x_k(t)|L_k = 0],$$

$$f[x_k(t)] = q_k f_{N(\mu_k, \sigma_k^2 + \delta_k^2)} + (1 - q_k) f_{N(0, \sigma_k^2)}.$$

As we mentioned to simulate $(\lambda_1, \ldots, \lambda_6)$ in the first step, we simulated $(x_1, \ldots, x_6)$, and then according to (2) the average of the intensities of each rating class, it can be obtained. When all parameters were estimated, the residuals can be simulated using the copula and marginal distributions of $(X_1, \ldots, X_6)$. For more details about the simulation of intensities see the Appendix.

### 4.2 Parameter estimation of marginal distributions and copulas

In this step, parameters of the marginal distributions of $X_k$ (residuals) were estimated by the maximum-likelihood method for each rating class. The results of parameter estimation were shown in Table 3. Note that the standard deviation (i.e. $\delta_k$) increases with declining credit quality. $\theta_k$ and $\kappa_k$ were estimated using data (Das and Geng [6]). After estimating the parameters, residuals for each rating class were computed.

**Table 3:** Estimated parameters by the maximum likelihood method for the average of the intensities of each rating class.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_k$</td>
<td>0.0134</td>
<td>0.1577</td>
<td>0.1899</td>
<td>0.1731</td>
<td>0.2163</td>
<td>0.3455</td>
</tr>
<tr>
<td>$\sigma_k$</td>
<td>0.0021</td>
<td>0.0029</td>
<td>0.0751</td>
<td>0.0190</td>
<td>0.0899</td>
<td>1.3921</td>
</tr>
<tr>
<td>$\mu_k$</td>
<td>0.0023</td>
<td>-0.0019</td>
<td>-0.0201</td>
<td>-0.0045</td>
<td>-0.0036</td>
<td>0.0360</td>
</tr>
<tr>
<td>$\delta_k$</td>
<td>0.0289</td>
<td>0.0312</td>
<td>0.1901</td>
<td>0.2011</td>
<td>0.7039</td>
<td>0.9187</td>
</tr>
</tbody>
</table>

This approach has been suggested by Das and Geng [6] that their focus is on the classical method, but in this paper two methods were used.
to estimate parameters of copulas: the classical and Bayesian method, in addition we have extended it to a mixture of Gaussian distributions for marginal distributions of the residuals of default processes in the classical and Bayesian methods. In order to apply the Bayesian methods for estimation of parameters of copula, we need the posterior distribution, calculated by multiplying the likelihood function to the prior distribution (see the B).

In order to generate a random vector from the copula, first, the distributions of marginal are used, deriving from the residual distributions (a mixture of Gaussian). Second, a lot of correlation matrices of copulas were used (i.e. 100000), simulated from the posterior distribution. These matrices were obtained from the MCMC (Markov chain Monte Carlo) method instead of using a fixed parameter estimate of correlation matrices of copulas (see Dalla Valle [4]).

5 Simulating Correlated Defaults

In this section the results of the measure calculated with the classical and Bayesian methods would be shown. Parameters of Student’s $t$ and normal copula according to both the classical and the Bayesian methods are estimated then Student’s $t$ and normal copula are simulated (see the A) for comparing them with measures in Table 4 in which $d_{k,t}$ is defined as an element of matrix $d$ as follows:

$$d_{k,t} = |R_{k,t+1} - S_{k,t+1}|, \quad k = 1, \ldots, 6, \quad t = 1, \ldots, 40,$$

and measure $\bar{d}$

$$\bar{d} = \frac{\sum_{k=1}^{6} \sum_{t=1}^{40} d_{k,t}}{240},$$

where $R_{k,t+1}$ is denoted as the element in the $k$th rating class and time $t + 1$ of our real data and $S_{k,t+1}$ is the element in the $k$th rating class and the time $t + 1$ of matrix $S$ that is obtained from simulated data. Since four methods were applied for these simulations, there are four matrix $S$. It is obvious that if the measure $\bar{d}$ is small and it tend to zero, the related simulation of the copula is effective because of two reasons: first, parameter estimation is efficient. Second, in comparison with other copulas, the related copula fits well.
6 Conclusions

As it is shown in Table 4, Bayesian method in Student’s $t$ copula leads to an improvement of simulation, but it does not in normal copula. This verifies the fact that Bayesian method is not always more efficient than classic one. In this case it is efficient enough for heavy tail distribution. Moreover, our main outcome is that the application of Bayesian methodology causes a great reduction of measure when copula is Student’s $t$, but the drawback of the Bayesian method in normal copula compared to the obvious advantage in Student’s $t$ is that the reduction is negligible.

Table 4: Results of calculating measures with the classical and Bayesian copulas.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Classic Normal Copula</th>
<th>Student’s $t$ Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>0.1541</td>
<td>0.1972</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
<th>Bayes Normal Copula</th>
<th>Student’s $t$ Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>0.1587</td>
<td>0.1136</td>
</tr>
</tbody>
</table>

As in Table 5 is illustrated, the Student’s $t$ copula is applied with the same correlation matrix and two different degrees of freedom in order to calculate measures. Although calculating measures show that the small degree of freedom improves the simulation, it’s not tangible. The Student’s $t$ copula with high degrees of freedom approximates to the normal copula, therefore according to Table 5, heavy tail distribution fits well on our data.

Table 5: Results of calculating measures using Student’s $t$ copula with the same correlation matrix and two differences degrees of freedom.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Student’s $t$ Copula($\nu = 13, \Sigma_0$)</th>
<th>Student’s $t$ Copula($\nu = 43, \Sigma_0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>0.1934</td>
<td>0.1972</td>
</tr>
</tbody>
</table>
Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the manuscript.

References


A

We report below the main steps in order to simulate from a multivariate default process with a given copula. For a detailed review of copula simulation see Dalla Valle et al. [5]. To generate random variables from multivariate default processes, the following algorithm is suggested.

1. Estimate parameters ($\kappa_k$ and $\theta_k$) using real intensities;

2. Determine residuals using equation (2) and step 1;

3. Estimate parameters of the marginal distributions of residuals with numerical method (maximum likelihood method for a mixture of Gaussian);

4. Estimate the parameters of copulas by calculating residuals from step 2;

5. Simulate ($x_1, \ldots, x_6$) using copula and marginal distributions residuals (for more details see Dalla Valle et al. [5]);

6. Finally determine ($\lambda_1, \ldots, \lambda_6$) using simulated ($x_1, \ldots, x_6$) and equation (2).

B

The copula of the multivariate normal distribution is the normal copula:

$$C(u_1, \ldots, u_r) = \Phi_r\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_r)\right)$$

where $C$ shows the normal copula, $\Phi_r$ implies the joint distribution function of the $r$-variate standard normal distribution and $\Phi^{-1}$ shows the inverse of the distribution function of the univariate standard normal distribution.

Let $x = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_r))'$ named the vector of univariate normal inverse distribution functions, where $u_i = \Phi(x_i)$ for $i = 1, \ldots, r$, and
let $\Sigma$ named the correlation matrix, then the normal copula probability density function is presented in the following form

$$c(\Phi(x_1), \ldots, \Phi(x_r)) = \frac{1}{|\Sigma|^{1/2}} \exp \left( - \frac{1}{2} x' (\Sigma^{-1} - I_r) x \right)$$

(3)

using equation (3), the normal copula probability density, we calculate the product over $t$ to get the likelihood function using residuals of intensity processes, which has the following form:

$$f(x|\Sigma) = \frac{1}{|\Sigma|^{T/2}} \exp \left( - \sum_{t=1}^{T} \frac{1}{2} x'_t (\Sigma^{-1} - I_r) x_t \right).$$

The parameter to be estimated is $\Sigma$ and we selected the Inverse Wishart distribution as a conjugate prior:

$$\Sigma \sim \text{InverseWishart}(\alpha, B).$$

Then the posterior distribution is computed using Bayes’ theorem:

$$\pi(\Sigma|x) \sim \text{InverseWishart} \left( \frac{T}{2} + \alpha; B + \frac{1}{2} \sum_{t=1}^{T} x_t x'_t \right).$$

We applied the Gibbs sampler algorithm to simulate the correlation matrix’s estimate. We used the 100,000 matrices simulated by the posterior distribution obtained with the MCMC method. For more details about parameter estimation of copulas see Dalla Valle [4], Dalla Valle [3], Dalla Valle et al. [5], Genest and Favre [14], Bayesian method see Robert and Casella [23].