European Journal of Operational Research 205 (2010) 361–367

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Efficient estimation of large portfolio loss probabilities in *t*-copula models

Joshua C.C. Chan *, Dirk P. Kroese

Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia

ARTICLE INFO

Article history:
Received 10 June 2009
Accepted 4 January 2010
Available online 11 January 2010

Keywords:
Credit risk
Copula models
Rare-event simulation
Cross-entropy method
Conditional Monte Carlo

ABSTRACT

We consider the problem of accurately measuring the credit risk of a portfolio consisting of loans, bonds and other financial assets. One particular performance measure of interest is the probability of large portfolio losses over a fixed time horizon. We revisit the so-called *t*-copula that generalizes the popular normal copula to allow for extremal dependence among defaults. By utilizing the asymptotic description of how the rare event occurs, we derive two simple simulation algorithms based on conditional Monte Carlo to estimate the probability that the portfolio incurs large losses under the *t*-copula. We further show that the less efficient estimator exhibits bounded relative error. An extensive simulation study demonstrates that both estimators outperform existing algorithms. We then discuss a generalization of the *t*-copula model that allows the multivariate defaults to have an asymmetric distribution. Lastly, we show how the estimators proposed for the *t*-copula can be modified to estimate the portfolio risk under the skew *t*-copula model.

1. Introduction

Losses resulting from the failure of an obligor to make a contractual payment, generally referred to as *credit risk*, are one of the major concerns of financial institutions. Consequently, the problem of accurately measuring the credit risk of a portfolio consisting of various financial assets has received considerable attention in the literature. Each obligor in the portfolio is subject to possible default, and such an event is often captured by the so-called *threshold models*, where a default occurs when a latent variable exceeds a given threshold. In order to model the dependence of simultaneous defaults observed empirically, a dependence structure is often imposed on the multivariate default distribution. The most popular choice of such a structure is the multivariate normal distribution. This gives rise to the celebrated *normal copula model*, which is widely used in the financial industry and forms the basis of the Morgan's CreditMetrics and other management systems (Gupton et al., 1997; Li, 2000). See also the monographs by Bluhm et al. (2002) and McNeil et al. (2005). Under the normal copula framework, dependence is often induced via a set of common factors affecting multiple obligors. These factors are typically interpreted as economy-wide risks, to which all the obligors are exposed, though to varying degrees. Conditional on these factors, the obligors then become independent. In terms of performance measures of credit risk, one that is of particular importance is the probability of large portfolio losses over a fixed time horizon. Since this probability is typically not available analytically, Monte Carlo methods are required to estimate this quantity.

To generate more scenarios with large losses in simulation, one common approach is to shift the factor mean via *importance sampling* (IS) (see, e.g., Rubinstein and Kroese, 2007), as suggested in, e.g., Kalkbrener et al. (2004), Joshi (2004) and Egloff et al. (2005). Although this heuristics works well empirically in the context of single-factor normal copula models, there is little theoretical support, and consequently, the procedure might fail for certain sets of parameter values. Glasserman and Li (2005) derive logarithmic limits for the tail of the loss distribution associated with single-factor homogeneous portfolios. In particular, they show that for the regime with moderately high correlation among the obligors, the occurrence of large losses is determined primarily by the common factor, thus justifying the heuristics of shifting the factor mean. Moreover, they propose the following two-step IS procedure: first apply IS to shift the factor mean, then apply IS again conditional on the common factor affecting multiple obligors. They further show that the proposed estimator is logarithmically efficient. Although the utility of this two-step procedure is supported by both theoretical and numerical results, it is difficult to generalize the procedure to the general multi-factor model. In view of this difficulty, Glasserman et al. (2007) analyze the general multi-factor normal copula setting and derive logarithmic asymptotics for the loss distribution. The asymptotic results are later exploited in Glasserman et al. (2008) to develop logarithmically efficient IS techniques to estimate the tail probabilities of large portfolio losses. We refer

* Corresponding author.
E-mail addresses: c.chan5@uq.edu.au (J.C.C. Chan), kroese@maths.uq.edu.au (D.P. Kroese).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2010.01.003

the readers to the recent review in Grundke (2009) for other related approaches.

Despite its popularity, the normal copula model does not capture various stylized facts about financial variables brought forth by recent empirical research. In particular, one of the most prominent features of financial variables is that they exhibit *extremal dependence*, i.e., they are asymptotically dependent. Loosely speaking, the variables take on large values (in absolute terms) simultaneously with non-negligible probability, and it is not captured by the correlation structure implied by the multivariate normal distribution. In view of this inadequacy of the normal copula, Bassamboo et al. (2008) propose the t-copula model, based on the multivariate t-distribution, that attempts to capture the relatively frequent occurrences of extremal comovements of financial variables. They argue that in many instances it is a more adequate way to model dependencies than the normal copula. The authors derive sharp asymptotics for the loss distribution, and show that under the t-copula model, large portfolio losses occur primarily when the so-called *shock variable* takes on small values, while other random variables, including the common factors, are relatively unimportant in determining the occurrence of large losses. In other words, shifting the factor mean alone, as suggested by the aforementioned IS procedures, would not significantly increase the number of scenarios with large losses, and consequently, substantial variance reduction might not be achieved. Therefore, the authors propose two IS algorithms to estimate the probability of large portfolio losses. The first estimator uses IS based on an *exponential change of measure* (ECM) (see, e.g., Asmussen and Glynn, 2007) and has bounded relative error; the second uses a variant of *hazard rate twisting* (HRT) (Juneja and Shahabuddin, 2002), which is shown to be logarithmically efficient. An extensive simulation study shows that while both estimators offer substantial variance reduction, the former provides 6–10 times higher variance reduction than the latter. Nevertheless, the more efficient ECM algorithm involves generating random variables from a nonstandard distribution via rejection sampling, which takes on average three times more time compared to naive Monte Carlo simulation. In addition, the normalizing constant of the proposal density is not known, and has to be computed by numerical routines in order to be used in the likelihood ratio evaluation.

Instead of the two IS algorithms, we propose two novel estimators based on *conditional Monte Carlo* (see, e.g., Asmussen and Glynn, 2007; Rubinstein and Kroese, 2007) to estimate the probability of large portfolio loss under the t-copula model. We prove that the less efficient estimator has bounded relative error. A simulation study similar to that in Bassamboo et al. (2008) further shows that the proposed estimators outperform (in terms of variance reduction) both ECM and HRT algorithms. Moreover, the new algorithms involve only generating random variables from standard distributions, and consequently they are as efficient as naive simulation in terms of random variable generation effort. An additional advantage is that the new algorithms require trivial programming effort and are easier to implement than those proposed in Bassamboo et al. (2008), as the latter require generating random variables from nonstandard distributions. We then consider a generalization of the t-copula model to an asymmetric default distribution, as opposed to the symmetric distribution implied by the normal copula and t-copula models. This generalization is relevant and potentially important as it incorporates the well-documented observation that in practice financial variables are highly skewed (Fernandez and Steel, 1998; Franses and van Dijk, 2000). In a credit risk setting, for instance, there is relatively little potential gain when the underlying economic conditions improve, but there is a substantial downside risk when the market condition worsens. Consequently, the multivariate default distribution is expected to be positively skewed (since a large positive draw of the latent variable represents a default). Failure of taking this asymmetry into account might result in underestimation of the credit risk of the portfolio.

The rest of this article is organized as follows. In Section 2 we formulate the problem of estimating large portfolio losses and introduce the normal copula model. We then discuss the t-copula model in Section 3. Section 4 discusses two estimation methods based on conditional Monte Carlo for estimating the probability of large portfolio loss under the t-copula model. The performance of the these estimators are demonstrated via an extensive simulation study in Section 5. Finally, we consider the skew t-copula model that accommodates an asymmetric default distribution. There we also study how the skewness of the multivariate default distribution affects the probability of large portfolio loss.

2. Problem formulation

Consider a lender owning a portfolio of loans consisting of n obligors, each of whom has a positive, albeit small, probability of defaulting. Let the probability of default for the ith obligor be $p_i \in (0,1)$, which we take as given. In practice, these probabilities can often be estimated by various econometrics models using historical data and other observed characteristics of the current obligors. We further assume that the monetary loss associated with the default of the ith obligor, denoted as c_i, is known. We introduce a vector of underlying latent variables $\mathbf{X} = (X_1, \ldots, X_n)$ so that the ith obligor defaults if X_i exceeds some given threshold level x_i. More specifically, let $f_{X_i}(x)$ denote the (marginal) probability density function (pdf) of X_i. Given the probability of default p_i, the threshold x_i is determined implicitly by

$$\mathbb{P}_{f_{X_i}}(X_i > x_i) = \int_{x_i}^{\infty} f_{X_i}(u)\,du = p_i.$$

One could interpret the latent variable X_i as the underlying financial condition of the ith obligor, which is not directly observable to the lender. However, when the obligor's financial condition become worse than a critical level (x_i), she goes bankrupt and the lender observes a default in the ith loan. Our main interest is to learn about the distribution of the loss incurred from defaults

$$L(\mathbf{X}) = c_1 I_{\{X_1 > x_1\}} + \cdots + c_n I_{\{X_n > x_n\}}, \tag{1}$$

where $I_{\{\cdot\}}$ denotes the indicator function. In particular, we wish to estimate accurately the probability of large losses of the form

$$\ell(\gamma) = \mathbb{P}_f(L(\mathbf{X}) > \gamma), \tag{2}$$

where $\mathbf{X} \sim f(\mathbf{x})$ and $\gamma = bn$ for some $b > 0$. In order to estimate the above probability, one needs to specify the joint distribution of the latent variables $\mathbf{X} = (X_1, \ldots, X_n)$. It is obvious that the usefulness of the model depends critically on the distributional assumptions of the vector $\mathbf{X}$. On the one hand, the researcher wishes to make as few assumptions about the joint distribution as possible, since imposing restrictive but unrealistic assumptions often lead to misleading conclusions. On the other hand, a parameter-rich model often makes the analysis intractable. How this trade-off between flexibility and tractability is made is therefore of vital importance.

One popular model that is widely used in the financial industry is the normal copula model that forms the basis of the CreditMetrics and other related models. The normal copula model attempts to capture the dependence among obligors while maintaining mathematical tractability by assuming the vector of latent variables follows a multivariate normal distribution. More specifically, the underlying correlations are often specified through a linear factor model

$$X_i = w_{i1} Z_1 + \cdots + w_{id} Z_d + w_i \eta_i, \quad i = 1, \ldots, n, \tag{3}$$

where $Z_1, \ldots, Z_d$ are independent and identically distributed (iid) standard normal variables known as *factors* that capture the