Loss Distribution Generation in Credit Portfolio Modeling

Igor Jouravlev, MMF, Walden University, USA

Ruth A. Maurer, Ph.D., Professor Emeritus of Mathematical and Computer Sciences, Colorado School of Mines, USA

Key words: Loss distribution, Copula, Credit Portfolio, Large Homogeneous Portfolio, Monte Carlo Simulations, Importance Sampling, Saddle-point Approximation

Introduction

In the current paper we analyze several methods for generation of loss distribution for credit portfolios. Loss distributions play an important role in pricing of credit derivatives and in credit portfolio optimization. A loss distribution is a function of the number of entities in the portfolio, their credit ratings, the notional amount and recovery of each entity, default probabilities, loss given defaults, and the correlation/dependence structure between entities incorporated in the portfolio. Direct generation of loss distribution may require Monte Carlo simulation which is time consuming and is not effective when applied for credit portfolio optimization. To overcome computational complexity a number of approaches were undertaken based on assumptions imposed on the input parameters, goals of loss distributions generation, and the accepted level of tolerance and computational errors.

Literature review

A wide range of literature was dedicated to generation of loss distributions for credit portfolios. Vasicek (1987, 2002) developed a large homogeneous portfolio approximation that played an important role in one of the first synthetic CDOs pricing methods developed by J.P. Morgan. This method was generalized by considering a finite number of obligors in the portfolio. Hull and White (2004) introduced a bucketing approach, and Andersen, Sidenius and Basu

(2003) offered a recursive method valid in more general cases. Glasserman and Li (2005) proposed the two-step-importance sampling method in which they applied Monte Carlo simulation methods with variance reduction techniques. A Fourier analytical approach in loss distribution generation was analyzed by Merino and Nyfeler (2002), Reiss (2003), and Grundke (2007) which was used for pricing of CDO by Greogry and Laurent (2004), Laurent (2004), and Laurent and Gregory (2005). This method requires a good implementation of the fast Fourier transform. Saddle-point approximation was analyzed by Arvanitis and Gregory (2001), Gordy (2002), Martin (2006), Glasserman (2008). We will analyze and compare these methods, their advantages and disadvantages.

Large homogeneous approximations of loss distributions

Large homogeneous portfolio (LHP) approximation was the first method developed by Oldrich Vasicek in 1987 at KMV Corporation. In this model, Vasicek assumed that the portfolio contains an infinite number of entities. Each entity has the same notional amount, default probability, and recovery rate (or loss given default). Loss given default can be calculated as 1 – recovery rate. The correlation structure is presented using one-factor Gaussian copula. Although these assumptions were very restrictive, the closed formula derived by Vasicek was easy to implement and the model was much faster than the one used Monte Carlo simulations. According to this model, the risk neutral portfolio cumulative loss distribution and probability density function for a large portfolio with underlying Gaussian copula can be expressed as follows (Vasicek, 2002):

$$P(L \le x) = \Phi\left(\frac{\sqrt{1-\rho}\,\Phi^{-1}(x) - \Phi^{-1}(p)}{\sqrt{\rho}}\right) \tag{1}$$

$$f(x) = \frac{\sqrt{1-\rho}}{\sqrt{\rho}} \exp\left(-\frac{1}{2\rho}\left(\sqrt{1-\rho}\Phi^{-1}(x) - \Phi^{-1}(p)\right)^2 + \frac{1}{2}\left(\Phi^{-1}(x)\right)^2 \right) \qquad (2)$$

where $L \in [0,1]$ – portfolio loss estimated as a fraction of the whole portfolio, Φ – cumulative normal distribution, Φ^{-1} – is the inverse of cumulative normal distribution, p – the probability of default of a loan in the portfolio; in this case we assume that all credits have the same probability of default, and the number of credits is infinite, ρ - correlation incorporated into the credit portfolio using one-factor Gaussian copula. Although the portfolio contains loans with the same default characteristics; the loans are in fact different, and, moreover, correlated with correlation ρ. The advantage of using this formula is that it can be used for fast approximation of loss distribution for any range $[a,b]$, where a < b, and $a,b \in [0,1]$.

The above mentioned formulas (1) and (2) depend on 2 very important parameters – correlation and probability of default of a loan in the portfolio. A credit portfolio with high value of correlation (say 90%) with very small value of probability of default of a loan (say < 10 bps) will have loss distribution concentrated around 0%, and can be approximately considered as a credit portfolio riskless.

One of the extensions of the Gaussian LHP approach is to use double-*t* one factor model proposed by Hull and White (2004). They assumed that common and individual factors are t-distributed and derived a formula that gives fast approximation of the loss cumulative distribution function.

The LHP approximation can be extended to the case of the Student-t copula. This approach also allows one to obtain analytical formulas for density and the cumulative distribution function of the portfolio loss distribution (Schloegl & O'Kane, 2005). Schloegl and

O'Kane (2005) also compared the VaR implied by the Student-t copula to that obtained using the Gaussian, Calyton, and Gumbel copulas. According to Schloegl and O'Kane (2005), the returns ζ_i of each obligor follow a multivariate Student-t distribution, so that (p. 578)

$$\zeta_i = \frac{\beta Z + \sqrt{1-\beta^2}\,\varepsilon_i}{\sqrt{\dfrac{W}{\nu}}} \tag{3}$$

where $Z, \varepsilon_1, ..., \varepsilon_M \sim N(0,1), W \sim \chi^2(\nu), \beta \in]0,1[$. Issuer *i* defaults if and only if $\zeta_i \le D$ where *D* is a certain threshold. This inequality is equivalent to:

$$\sqrt{1-\beta^2}\,\varepsilon_i \le D\sqrt{\frac{W}{\nu}} - \beta Z = \phi\,. \tag{4}$$

Conditionally on ϕ, the default indicator functions $1_{[\zeta_1 \le D]}$ are all independent and the conditional default probability can be written as a function of the standard normal cumulative density function $P[\zeta_i \le D \mid \phi] = \Phi\left(\frac{\phi}{\sqrt{1-\beta^2}}\right)$. By the law of large numbers, the fraction of defaulting issuers will tend to $\Phi\left(\frac{\phi}{\sqrt{1-\beta^2}}\right)$ for each realization of ϕ as $M \to \infty$. Assuming that exactly this fraction of issuer defaults for each realization of ϕ, thereby replacing the random variable *L* by $E[L \mid \phi]$, we have that $L \approx h(\phi)$, where $h(x) = (1-R)\Phi\left(\frac{x}{\sqrt{1-\beta^2}}\right)$, and $\forall(\theta \in [0,1])P[\phi \le \theta] = F\left(h^{-1}(\theta)\right)$, where *F* is the cumulative distribution function of ϕ. Based on this logic, analytic formulas for cumulative distribution function and probability density function can be expressed using the following formulas (Schloegl & O'Kane (2005), pp. 579 – 580):

$$F(t) = P[\phi \le t] = \Phi\left(\frac{t}{\beta}\right) + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{t}{\beta}} \Gamma\left(\frac{\nu}{2}, \frac{\nu(t+\beta u)^2}{2D^2}\right) e^{-\frac{u^2}{2}} du \tag{5}$$

$$f(t) = \frac{1}{\beta\sqrt{\pi}\, 2^{\frac{\nu+1}{2}} \Gamma\left(\frac{\nu}{2}\right)} \int_0^{+\infty} e^{-\frac{1}{2\beta^2}\left(t - D\sqrt{\frac{w}{\nu}}\right)^2} w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}} dw\,. \tag{6}$$

As was noted by O'Kane (2008), there are approaches that approximately estimate loss distributions and can serve as acceptable compromises to the trade-off between speed and accuracy. These approaches are the Gaussian approximation, binomial and adjusted binomial distributions (pp. 354 – 360). The Gaussian approximation uses a Gaussian density which fits the first two moments of the conditional loss distribution. It is then possible to obtain a closed-form expression for the expected tranche loss conditional on the market factor (O'Kane, 2008, p. 354). The idea behind the binomial approximation is to approximate the exact multinomial distribution with a binomial distribution. The reason for this is that the shape of binomial distribution is a better fit to the multinomial distribution than Gaussian. However, in this approach we match the first moment of the exact conditional loss distribution. The further improvement is based in finding a way to fit the variance. For this purpose, an adjusted binomial approximation was proposed by O'Kane (2008) to ensure that we match first two moments of the exact inhomogeneous loss distributions (pp. 358 – 360).

The LHP approximation can be extended by using normal inverse Gaussian distribution (NIG). The normal inverse Gaussian distribution is a mixture of normal and inverse Gaussian distribution. According to Kalemanova, Schmid, and Werner (2007), a non-negative random variable *Y* has inverse Gaussian distribution with positive parameters α and β if its density function can be represented using the following form:

$$f_{IG}(y;\alpha,\beta)=\begin{cases}\dfrac{\alpha}{\sqrt{2\pi\beta}}y^{-3/2}e^{\left(-\frac{(\alpha-\beta y)^2}{2\beta y}\right)}, \text{if } y>0\\ 0, \text{otherwise}\end{cases} \tag{7}$$

A random variable X follows the NIG distribution with parameters α,β,μ,δ ($X \sim NIG(\alpha,\beta,\mu,\delta)$) if

$$\begin{aligned} &X \mid Y = y \sim N(\mu+\beta y, y)\\ &Y \sim IG(\delta\gamma,\gamma^2), \gamma=\sqrt{\alpha^2-\beta^2}\end{aligned} \tag{8}$$

The parameters should satisfy the following conditions: $0 \le |\beta| < \alpha$ and $\delta > 0$.

The density of the random variable $X \sim NIG(\alpha,\beta,\mu,\delta)$ is given by the following formula:

$$f_{NIG}(x;\alpha,\beta,\mu,\delta)=\frac{\alpha\delta\exp(\delta\gamma+\beta(x-\mu))}{\pi\sqrt{\delta^2+(x-\mu)^2}}K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right) \tag{9}$$

where

$$K_1(w)=\frac{1}{2}\int_0^\infty \exp\left(-\frac{1}{2}w\left(t+t^{-1}\right)\right)dt \tag{10}$$

is the modified Bessel function of the third kind. Kalemanova et al. (2007) suggested using the following parameters in the LHP model with NIG copula (with the dependency parameter ϕ defined as a square root of correlation assumed in a credit portfolio):

$$\begin{aligned} &M \sim NIG\left(\alpha,\beta,-\frac{\beta\gamma^2}{\alpha^2},\frac{\gamma^3}{\alpha^2}\right)\\ &V_i \sim NIG\left(\frac{\sqrt{1-\phi^2}}{\phi}\alpha,\frac{\sqrt{1-\phi^2}}{\phi}\beta,-\frac{\sqrt{1-\phi^2}}{\phi}\frac{\beta\gamma^2}{\alpha^2},\frac{\sqrt{1-\phi^2}}{\phi}\frac{\gamma^3}{\alpha^2}\right)\end{aligned} \tag{11}$$

where M is a systematic market risk factor, and V_i are idiosyncratic factors.

Then, asset returns will also follow NIG distribution with the following parameters:

$$M \sim NIG\left(\frac{\alpha}{\phi}, \frac{\beta}{\phi}, -\frac{1}{\phi}\frac{\beta\gamma^2}{\alpha^2}, \frac{1}{\phi}\frac{\gamma^3}{\alpha^2}\right) \tag{12}$$

The third and fourth parameters were chosen to get expected value of zero and variance of 1. Using the following notation $F_{NIG(s)}(x) = F_{NIG}\left(x; s\alpha, s\beta, -s\frac{\beta\gamma^2}{\alpha^2}, s\frac{\gamma^3}{\alpha^2}\right)$, the default threshold can be computed as follows:

$$C = F^{-1}_{NIG\left(\frac{1}{\phi}\right)}(p); \tag{13}$$

the default probability of each credit conditional on market factor is given by

$$p(M) = f_{NIG\left(\frac{\sqrt{1-\phi^2}}{\phi}\right)}\left(\frac{C-\phi M}{\sqrt{1-\phi^2}}\right) \tag{14}$$

and the loss distribution of the LHP can be estimated using the following formula (Kalemanova et al., 2007):

$$F(x) = 1 - F_{NIG(1)}\left(\frac{C - \sqrt{1-\phi^2}\,F^{-1}_{NIG\left(\frac{\sqrt{1-\phi^2}}{\phi}\right)}(x)}{\phi}\right) \tag{16}$$

These formulas use special functions; however, the advantage of using them is that computation of loss distributions using Gaussian copula, Student-t, double-t, or NIG copula is much less time consuming than using Monte Carlo simulation for generation loss distributions. One of the advantages of using NIG copula is in the possibility of estimating four parameters of this distribution given observed first four moments.

Schönbucher (2002) used an algorithm from the theory of Archimedean copula functions to estimate limiting loss distributions which are driven by random variable with different dependency structures. This approach allowed presenting simple and realistic formulas for the loan portfolio distribution. The joint distributions in a credit portfolio are modeled different ways than just using a variant of the multivariate normal distribution function, and this approach proved to be feasible. In obtaining loss distributions, it is important to investigate the effect of the implicit assumption of a Gaussian dependency structure on the risk measures and the returns distribution of the portfolio, as well as the consequences of extreme events and lack of available data on credit risk modeling. Schönbucher (2002) showed that in the credit risk case, this effect can be either minor (when the Vasicek model is compared to the Clayton-dependent model) or significant (when one thinks that the Gumbel copula is a realistic alternative).

Finite homogeneous approximation of loss distributions

In finite homogeneous approximation (often called exact computation), the infinity assumption is dropped and the model uses assumptions of a single systematic factor and homogeneity; in this case the numerical procedure is still easy to implement. In most cases, however, the portfolios are not homogeneous, but if we assume that the portfolio is large, granularity adjustment developed by Gordy (2003), Pykhtin and Dev (2003), and Gordy (2004) can be applied. In both large homogeneous and finite homogeneous portfolios the loss distribution can be generated based on assumption of several systematic risk factors using appropriate random number generations for each factors.

For a credit portfolio consisting of *n* entities, in the Gaussian copula case, the probability of *k* defaults (or unconditional loss distribution in discrete case) can be expressed using the following formula (Vasicek, 2002):

$$P_k = \binom{n}{k}\int_{-\infty}^{\infty}\left(\Phi\left(\frac{1}{\sqrt{1-\rho}}\left(\Phi^{-1}(p)-\sqrt{\rho}u\right)\right)\right)^k\left(1-\Phi\left(\frac{1}{\sqrt{1-\rho}}\left(\Phi^{-1}(p)-\sqrt{\rho}u\right)\right)\right)^{n-k} d\Phi(u) \quad (17)$$

where p is the probability of default, ρ is correlation incorporated into the credit portfolio using Gaussian copula, the integrand is the conditional probability of the portfolio loss given the market factor u which is assumed to be normally distributed. If we consider m market factors in this approach, then the integration would be over these m market factors.

Conditional and unconditional loss distributions generation

The previous formula developed by Vasicek (2002) gives an idea on how to obtain the unconditional loss distribution in the general case. First of all, one has to compute a conditional loss distribution conditional on a set of underlying factors in which defaults are independent, and then integrate the conditional loss distribution over the distribution of the underlying factors. In mathematical notation, we need to compute conditional probabilities conditional on market factors first:

$$P[L = L_k \mid M_1, M_2, ..., M_l] = \binom{n}{k}\left(p_t^{i|M_1,M_2,...,M_l}\right)^k\left(1-p_t^{i|M_1,M_2,...,M_l}\right)^{n-k} \quad (18)$$

and then integrate over these factors:

$$\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} P[L = L_k \mid M_1 = m_1, ..., M_l = m_l]dF(m_1)...dF(m_l) \quad (19)$$

where $L_k = \frac{k}{n}(1-R)$ is the percentage portfolio loss, $P[L = L_k \mid M_1, M_2, ..., M_l]$ is the probability that exactly k out of n issuers default conditional on market factors $M_1, M_2, ..., M_l$; and $p_t^{i|M_1,...,M_l}$ is the conditional default probability of obligor i at time t. In case of Gaussian copula and in case when we consider only one market factor in our model, we have:

$$p_t^{i|M} = \Phi\left(\frac{\Phi^{-1}(p_i) - \sqrt{\rho_i}M}{\sqrt{1-\rho_i}}\right). \tag{20}$$

Edgeworth expansion can be used as one of the methods for generating of loss distribution (Arvanitis & Gregory, 2001). This expansion uses the higher-order moments of the distributions of the constituent variables (such as number of defaults) and information contained in cumulants. It is well known that for the normal distribution, the first two cumulants are the mean and variance and the others are equal to zero. The higher cumulants give quantitative information about the non-normality of a distribution. Edgeworth expansion states that if the number of defaults are independent random variables with means p_i, standard deviation σ_i and cumulants κ_i, the probability density function of the random variable $Y = \sum_{i=1}^{n} d_i$ which represents number of defaults, can be given by

$$f_Y(y)dy = \frac{e^{-\frac{1}{2}t^2}dt}{\sqrt{2\pi}}\left[1 + \frac{\kappa_3 H_3(t)}{3!\sigma^3}n^{-1/2} + \left(\frac{\kappa_4 H_4(t)}{4!\sigma^4} + \frac{\kappa_3{}^2 H_6(t)}{2!3!^2\sigma^6}\right)n^{-1} + \ldots\right] \tag{21}$$

with

$$t = \frac{y-p}{\sigma\sqrt{n}}, \quad p = \frac{1}{n}\sum_{i=1}^{n} p_i, \; \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2, \; \kappa_r = \frac{1}{n}\sum_{i=1}^{n}\kappa_{i,r} \tag{22}$$

where $H_r(t)$ is the rth Hermite polynomial, which can be obtained by successive differentiation of the function $e^{-\frac{1}{2}t^2}$ using following equation:

$$H_r(t) = (-1)^n e^{\frac{t^2}{2}} \frac{d^r}{dt^r} e^{-\frac{t^2}{2}} \tag{23}$$

Cumulants can be obtained from the power-series expansion of the logarithm of the moment generating function of the random variable. The leading term in this expansion is the normal distribution, and the first and second terms adjust the skewness and kurtosis. These adjustments give better approximation to the loss distribution by incorporating of the skew (which corresponds to the asymmetry of the distribution) and kurtosis (which corresponds to the fat tails of distribution). As was shown by Arvanitis and Gregory (2001), the Edgeworth approximation with four moments is better than the normal approximation, but it becomes less accurate further into the tail. The approximation is poor to the left of the origin, where true probability density function vanishes. For example, the estimation of the unexpected loss at the 99.9th percentile corresponds to a tail probability of 0.001. The normal approximation underestimates the true value by 40% and the Edgeworth approximation (with four terms) underestimates it by 14% (pp. 77-78).

Hull and White (2004) presented two approaches in generating loss distributions. They considered a number of market factors $M_1,...,M_l$ and the conditional default probabilities were considered conditional on these market factors. Defining $\pi_T(k)$ the probability that exactly *k* defaults occur in the portfolio before time *T*, conditional on the default times t_i are independent. The conditional default probability that all the n names will survive beyond time *T* is $\pi_T(0 \mid V_1,...,V_m) = \prod_{i=1}^{n} S_i(T \mid M_1,...,M_l)$ where $S_i(T \mid M_1,...,M_l)$ is the survival probability of the obligor *i*. Hull and White (2004) showed that the conditional probability of exactly *k* defaults by time *T* is

$$\pi_T(k \mid M_1,...,M_l) = \pi_T(0 \mid M_1,...,M_l) \sum w_{z(1)} w_{z(2)} ... w_{z(k)} \tag{24}$$

where $w_i = \dfrac{1 - S_i\left(T \mid M_1,\ldots,M_l\right)}{S_i\left(T \mid M_1,\ldots,M_l\right)}$ and $z(1),\ldots z(2)$ is the set of *k* different numbers from the finite set$\{1,\ldots,n\}$. Hull and White (2004) provided a fast algorithm for computing the conditional losses. By integrating over the market factors, one can obtain unconditional loss distribution.

Hull and White (2004) proposes also a bucketing approach while Anderson et al. (2003) proposed a recursive method for generating loss distribution. In both approaches, loss given default and notional values can vary between the entities, but they are still assumed to be deterministic.

Hull and White (2004) divided potential losses into the following ranges: $\{0, b_0\}, \{b_0, b_1\}, \ldots, \{b_{K-1}, b_\infty\}$ where $\{b_{k-1}, b_k\}$ is referred as *k*th bucket. The loss distribution is built by including one debt instrument at a time. The procedure keeps track of both the probability of the cumulative loss being in a bucket and the mean cumulative loss conditional that the cumulative loss is in the bucket. The approach offered by Hull and White (2004) does not assume buckets of identical size; it allows the analyst accommodating situations where extra accuracy is needed in some regions of the loss distribution. The latter can be achieved by considering smaller bucket sizes. The loss distribution can be truncated at some level so that the analyst need not spend extra computational time on large losses that have only a very small chance of occurring. Hull and White (2004) reported that their approach is comparable with the Fourier-analytical approach in terms of computational time and accuracy and is numerically stable.

The main idea of the approach suggested by Andersen et al. (2003) was to compute some u as a common divisor of all potential losses, and then consider losses $0, u, 2u, \ldots, n^* u$ rounded to the nearest discrete point as the loss distribution is built up. $n^* u$ here is the maximum possible loss. It is important to note here that the speed of this algorithm depends on the statistical characteristics of the credit spreads, so that in some cases, the value of the common divisor of all potential losses can be very small substantially affecting the speed of the algorithm. The next step was to implement a recursive algorithm to determine the portfolio loss distribution that is used for the conditional probabilities and in case when default events are independent. It should be noted here that the value for *u* can be very small in some credit portfolios and incur computationally extensive loss distribution generation. Suppose we know the loss distribution $P^K(l;t), l = 0,\ldots,l_{\max,K}$ for a reference pool of some size $K \geq 0$, where $l_{\max,K}$ is the sum of all the loss weights such reference pool. Suppose we add another company to the pool with loss weight w_{K+1} and known default probability $p_{K+1}(t)$. Then using independence of defaults we find for the loss distribution of the larger basket (Andersen et al., 2003, p.67):

$$P^{K+1}(l;t) = P^K(l;t)(1 - p_{K+1}(t)) + P^K(l - w_{K+1};t)\, p_{K+!}(t), l = 0,\ldots,l_{\max,K} + w_{k+1} \qquad (25)$$

This recursive relation starts with an empty basket, and increase in basket size leads to the same relative increase in the maximal loss. The cost of building the conditional loss distribution grows as roughly the square of the basket size (Andersen et al., 2003, p.67). The resulting conditional loss distribution is transformed to the unconditional loss distribution by integrating over common factor.

Fourier analytical approach is another approach of conditional and unconditional loss distribution generation. This approach is alternative to recursion techniques and considers a map of the original problem into another space where the problem is more analytically tractable. Once the problem in this space is solved, we need to map the solution back to the original space. This approach depends on the successful implementation of fast Fourier techniques. This approach was used, for example, by Gregory and Laurent (2003, 2004) for pricing CDOs. Reiβ (2003) provided detailed information on how loss distribution and its first moments can be obtained using Fourier analytical approach, described the CreditRisk+ model in terms of characteristic functions instead of probability functions. Since construction of loss distribution very much depends on the basic loss unit, Reiβ presented an alternative approach where no basic loss unit had to be introduced. Merino and Nyfeler (2002) described an algorithm that combines techniques from numerical mathematics and actuarial science. Their approach in generation of loss distributions was in grouping all potential losses into exposure buckets. Approximating the Bernoulli default indicators by Poisson random variables allowed reducing the number of random variables which correspond to the number of credits in the portfolio to the number of buckets. Applying the fast Fourier transform, numerical quasi Monte Carlo methods allowed generating loss distributions of credit portfolios containing 500,000 counterparties within four hours with adequate accuracy. It was shown that it was not necessary to simplify the credit risk model or portfolio structure to calculate the body and the tail of the portfolio loss distribution and that the algorithm was useful for analyzing and designing CDO structures.

Grundke (2007) analyzed whether a Fourier-based approach could be an efficient for calculation of risk measures in the context of a credit portfolio model with integrated market risk factors. He applied this approach to CreditMetrics credit portfolio model extended by correlated interest rate and credit spread risk. He showed that Fourier-base methods being superior to Monte-Carlo simulations couldn’t be superior in case of the integrated market and credit portfolio model even after applying standard importance sampling techniques for improving the performance of Fourier-based approach. This is because the higher the confidence level of the VaR, the larger the asset return correlation, or the larger the number of systematic risk factors. For the integrated market and credit portfolio methods, one should combine the Monte Carlo simulation with an importance sampling technique. Combination of Monte-Carlo and importance sampling techniques would be appropriate for those cases where the Fourier-based approach performs badly, for example, for the estimation of small percentiles which are needed in credit risk management.

The FFT approach was the first used loss distribution construction methodology. However, it is slower approach compared to the recursive approach (but still faster than the Monte Carlo approach) of generating loss distributions due to the following reasons (O'Kane, 2008):

- recursions are faster than Fourier methods,
- recursions are easier to implement and don't require access to any specialized numerical libraries;
- using recursions the researcher can build the loss distribution for a specific tranche; this is not possible to do when using FFT.

The saddle-point approximation for generating loss distribution proved to be very accurate in practice. When considering sums of independent random variables, it is convenient to consider the moment generating function (MGF). The MGF of a random variable can be represented as follows (Arvanitis & Gregory, 2001):

$$M_V(s) = \int_{-\infty}^{\infty} e^{st} p_v(t) dt \tag{26}$$

The well known property of the MGF of a random variables is that when independent variables are added, their distributions are convolved, but MGFs are multiplied. Multiplication operation is easy to perform than convolution. On the other side, one has to obtain the loss distribution of from MGF using inversion integration. By suitably approximating the shape of the integrand one obtains an analytical approximation of the probability density function. This technique is known as the method of steepest descents or saddle-point method. The method also allows obtaining analytical approximations to the probability without having to integrate the density function. The saddle-point approximation does not make any prior assumptions about the shape of the loss distribution. As was suggested by Arvanitis and Gregory (2001), saddle-point approximation method is a fast method in obtaining approximate loss distribution and tail probabilities, can be used for approximation to derive expressions for loss distributions that include variable exposures and default probabilities, and allows incorporating correlation into the loss distribution (p. 285). If we denote individual losses $V_1,...,V_n$, the distribution of each loss can be characterized through its cumulant generating function (Glasserman, 2008):

$$\forall \theta \,\&\, \forall (i = 1,...,n)\ \Lambda_i(\theta) = \log E(\exp(\theta V_i)) \tag{27}$$

Each V_i is a random fraction of a largest possible loss upon default of obligor *i*, all obligors considered here are independent. As a consequence of independence, the moment generating function can be represented as

$$\phi_L(\theta) = E\left[\exp\left(\theta \sum_{k=1}^{n} Y_k V_k\right)\right] = \prod_{k=1}^{n} E[\exp(\theta Y_k V_k)] = \prod (1 + p_k(\exp(\Lambda_k(\theta)) - 1)) \tag{28}$$

and the cumulant generating function of *L* as (Glasserman, 2008)

$$\psi_L(\theta) = \log[\phi_L(\theta)] = \log\left[\prod_{k=1}^{n} (1 + p_k(\exp(\Lambda_k(\theta)) - 1))\right] = \sum_{k=1}^{n} (\log(1 + p_k(\exp(\Lambda_k(\theta)) - 1))) \tag{29}$$

Analysis of this function shows that it is increasing, convex, infinitely differentiable and takes zero when $\theta = 0$. For a loss level $x > 0$, the saddle-point is the root θ_x of the equation $\psi_L'(\theta_x) = x$ which is unique.

Consider the cumulant generating function $\psi_L(\theta)$. The derivatives of $\psi_L(\theta)$ give cumulants of *L*. The first cumulant when $\theta = 0$ is the mean

$$\psi_L^{'}(0) = \sum_{k=1}^{n} p_k \Lambda_k^{'}(0) = \sum_{k=1}^{n} p_k E[V_k] = E[L]. \tag{30}$$

When losses are assumed to be constant, $V_k = v_k$, we have

$$\psi_L^{'}(\theta) = \sum_{k=1}^{n} \overline{p}_k v_k. \tag{31}$$

In other words, the derivative of the cumulant generating function is the expected loss where original default probabilities are replaced with $\overline{p}_k = \frac{p_k \exp(v)}{1 + p_k(\exp(v) - 1)}$.

In general, the formula for derivative is

$$\psi_k^{'}(\theta) = \sum_{k=1}^{n} \overline{p}_k(\theta) \Lambda_k^{'}(\theta) \tag{32}$$

which can be interpreted as the expected loss when the original default probabilities are replaced by $\overline{p}_k(\theta) = \frac{p_k \exp(\Lambda_k(\theta))}{1 + p_k(\exp(\Lambda_k(\theta)) - 1)}$ and the expected losses $E[V_k]$ are replaced with $\Lambda_k^{'}(\theta)$ where original expected loss $E[V_k]$ coincides with $\Lambda_k^{'}(0)$ because Λ_k is the cumulant generating function of V_k. Thus, each value of θ determines a modified set of default probabilities and a modified loss given default for each obligor (Glasserman, 2008, pp. 455 – 456).

Due to complexity of the formulas provided, it is important to either develop a fast algorithm or approximation formulas. A saddle-point approximation can be represented the following way (Glasserman, 2008):

$$P(L > x) \approx \exp\left(-\theta_x x + \psi_L(\theta_x) + 0.5\psi_L^{''}(\theta_x)\right)\Phi\left(-\theta_x \sqrt{\psi_L^{''}(\theta_x)}\right) \tag{33}$$

and the closely related Lugannani-Rice approximation is

$$P(L > x) \approx 1 - \Phi(r(x)) + \phi(r(x))\left(\frac{1}{\lambda(x)} - \frac{1}{r(x)}\right) \tag{34}$$

where $r(x) = \sqrt{2(\theta_x x - \psi_L(\theta_x))}$ and $\lambda(x) = \theta_x \sqrt{\psi_L''(\theta_x)}$.

A modification of the previous formula is

$$P(L > x) \approx 1 - \Phi\left(r(x) + \frac{1}{r(x)} \log \frac{\lambda}{r(x)}\right). \tag{35}$$

This modification has the advantage that it always produces a value between 0 and 1 (Glasserman, 2008, p. 456).

To generate the unconditional loss distribution, as it was shown before, one has to compute a conditional loss distribution based on the assumption of obligor independence and then integrate over the possible values of market factors. Factor models are popular since obligors are conditionally independent in such models. When integrating over the possible values of market factors, fast integration procedures should be used; these procedures depend on the probability distribution of the chosen market factors and the number of market factors.

Glasserman (2008) proposed and algorithm for generation of unconditional loss distribution. The model specifies two sets of parameters – corresponding to a high-default regime and a low-default regime, with independent obligors in each regime. Then the model uses a mixture of the two sets of parameters. In this case, the underlying “factor” is the regime and the unconditional loss distribution may be computed as a mixture of the two conditional loss distributions (p. 457). This approach is quite useful when analyzing the credit portfolio consisting of over 100 – 120 credits that can perform differently over the specified time period. The credit portfolio can be partitioned into the clusters (or buckets) of credits (as well as possible outliers) and for each cluster low default regime and high default regime can be identified.

Glasserman and Ruiz-Mata (2006) compared the computational efficiency of ordinary Monte Carlo simulation with methods that combine simulation for the factors with the techniques used to generate conditional loss distributions. They determined that numerical transform inversion and saddle-point approximation involve some error in the calculation of conditional default probabilities. Because each replication using convolution, transform inversion, or saddle-point approximation takes longer than each replication using ordinary simulation, these methods complete fewer replications in a fixed amount of computing time. The recursive convolution method computes the full conditional loss distribution on each replication while transform inversion and saddle-point approximation must in practice be limited to a smaller number of loss thresholds. Using the saddle-point approximation requires solving for multiple saddle-point parameters on each replication. The computation time required using recursive convolution grows quickly with the number of obligors. As a consequence of these, with the total computing time held fixed, ordinary Monte Carlo often produces a smaller mean square error (Glasserman, 2008, p. 458). The number of factors plays a substantial role in choosing whether Monte Carlo or another method such as saddle-point or recursive approach should be used. When the number of factors is small, Monte Carlo simulations can be replaced by integration. For the moderate number of dimensions, a quasi-Monte Carlo sampling can be applied. On the other side, one can use approximations where a single “most important” value of the factors is used (Glasserman, 2004). Zheng (2007) suggested approximation of the conditional loss distribution using a normal distribution by matching two moments and then computing the unconditional distribution through numerical integration assuming small number of factors. Specifically Zheng (2007) suggest computing mean and variance depending on market factor Z:

$$\mu(Z) = \sum_{i=1}^{n} p_i(Z) \text{ and } \sigma^2(Z) = \sum_{i=1}^{n} p_i(Z)(1 - p_i(Z)) \tag{36}$$

so that the 0^{th} and first order approximation can be represented respectively the following way (Edgworth expansion):

$$P[L \le x] \approx E\left[\Phi\left(\frac{x-\mu(Z)}{\sigma(Z)}\right)\right] \tag{37}$$

$$P[L \le x \mid Z] \approx E\left[\Phi\left(\frac{x-\mu(z)}{\sigma(z)}\right) + H\left(\frac{x-\mu(z)}{\sigma(z)}\right)\right] \tag{38}$$

where

$$\begin{aligned} H(x) &= (1/6)\sigma(z)^{-3}\gamma(z)(1-x^2)\phi(x) \\ \gamma(z) &= \sum_{i=1}^{n} p_i(z)(1-p_i(z))(1-2p_i(z)) \end{aligned} \tag{39}$$

The higher order expansion can be expressed using Chebyshev-Hermite polynomials and moments of up to order $k + 2$. The approximation error was estimated as $o\left(n^{-k/2}\right)$. The unconditional loss distribution then can be obtained using integration.

Glasserman and Li (2005) proposed a two-step-importance sampling technique that helps compute the upper tail of the loss distribution. Such techniques include Monte-Carlo simulation methods combined with adequate variance reduction approach, and are applicable to a wide range of applications. Importance sampling is a special variance reduction technique and is used to change the distributions of the relevant risk factors in such a way that more realizations of the loss random variable are in the upper tail. Then each realization is weighted by the likelihood ratio to correct for the change in distribution.

Dembo et al. (2004) provided a large deviation approximation of the tail distribution of total financial losses on a portfolio consisting of many positions. Quantitative analysis of large losses is helpful in structuring large portfolio so as to withstand severe losses. Applications of this approach include the total default losses on a bank portfolio or the total claims against an insurer. A key assumption was that conditional on a common ‘correlating’ factor, position looses are independent. For large losses, financial distress costs are more severe if the losses occur over a relatively short period of time. Sudden losses may cause extreme cash-flow stress, and investors may require more favorable terms when offering new lines of financing over short time periods, within which they may have a limited opportunity to gather information about the credit quality and long-term prospect of a distressed financial institutions. The results provided by the authors include conditions under which a large-deviations estimate of the likelihood of a failure-threatening loss during some sub-interval of time during a given planning horizon can be calculated from the likelihood of the same size loss in a certain fixed “key time horizon”. The conditional distribution of losses on each type of position can be estimated given the large portfolio loss of concern. The authors provided some analytical guidance on the dependence of large-loss probabilities on the structure of a portfolio with a large number of positions and the ‘most likely way’ that a large loss can occur. Given a large loss, the conditional likelihood of loss on each type of position and conditional distribution of exposure in the event of loss were calculated. These conditional calculations can be interpreted in the asymptotic sense of Gibbs conditioning principle.

Sidenius et al. (2008) presented the SPA framework in which models are specified by a two-layer process. The first layer models the dynamics of portfolio loss distributions in the absence of information about default times. This background process can be explicitly calibrated to the full grid of marginal loss distributions as implied by initial CDO tranche values indexed on maturity, as well as to the prices of suitable options. The authors gave sufficient conditions for consistent dynamics. The second layer models the loss process itself as a Markov process conditioned on the path taken by the background process. The choice of loss process is non-unique. Sidenius et al. (2008) presented a number of choices, and discussed their advantages and disadvantages. Several concrete model examples were given, and valuation in the new framework was described in detail. Among the specific securities for which algorithms are presented were CDO tranche options and leveraged super-senior tranches.

In practice, one can estimate the implied loss distribution from observed market data. For example, Krekel and Partenheimer (2006) described how to determine the implied loss surface of a credit portfolio from CDO tranche quotes. The approach can be applied for pricing of CDO tranches and N^{th}-to-default swaps and for risk management of CDO tranches. It also can serve as an initial distribution for dynamic loss models. The calibration can be performed numerically by solving a nonlinear optimization problem using sequential quadratic programming method.

Analysis, conclusion and recommendations

The most important issue in generating loss distributions for credit portfolios is to find a fast and accurate algorithm which can be easily implemented. There is always trade-off between these characteristics of the described algorithms. Some of the analyzed algorithms can be implemented easily for solving practical problems; on the other side, the disadvantage of implementing them is in a number of assumptions that can make the model unrealistic incurring inaccurate decision making. For a given credit portfolio, it is advantageous considering various loss distributions models such as bucketing approach by Hull and White (2004), recursive method by Andersen et al. (2003), Fourier analytical approach, and saddle-point approximation and use the ones that provide better solution to the specific problem and the given credit portfolio.

Among the algorithms described above, the fastest algorithm is LHP, but the accuracy of this algorithm depends on the number of obligors in the portfolio and the credit spread distribution. The more concentrated the credit spreads are and the more obligors the credit portfolio has, the more applicable LHP is. If we have a smaller number of obligors (say, less than 100 obligors) we can use finite homogeneous approach. Finite homogeneous approach is slower than LHP since we need to integrate conditional loss distribution over the market factor to obtain unconditional loss distribution. Unfortunately, in practice, homogeneity assumption and assumption that the portfolio is large are not the case and LHP can be rejected by practitioners seeking more accurate loss distribution generation. LHP as well as finite homogeneous portfolio methods can't be used as an approximation of heterogeneous portfolios – a loss distribution generated using a recursive method for a heterogeneous portfolio can be quite different from the loss distribution generated using LHP when the input parameter for the LHP is the weighted average spread of the heterogeneous portfolio.

On the other extreme, when more accurate generation of loss distribution is required, the best method is Monte Carlo simulation which is, at the same time, the slowest. The Monte Carlo simulation is especially valuable in cases when the level of heterogeneity is very high, in other words when the credit spread distribution is very diverse.

Monte Carlo simulations (as well as other loss generation methods) can be made faster and will require less memory if we cluster credit default spreads into homogeneous partitions of credit spreads. Given a heterogeneous credit portfolio, the task is to distribute into a group of clusters in such a way that the objects within each cluster are homogeneous and highly correlated while the clusters themselves are different. In fact, this task has two difficulties – firstly, we need to identify an optimal partition of the credit spreads into different clusters; secondly the spread clusters should be explicitly determined.

There are a number of algorithms that allows clustering of credit spreads – K-means, Diana, Clara, fuzzy analysis, self-organizing maps, model-based clustering. There are also internal and stability validation based on internal and stability measures. Taking the credit spreads data and clustering partition as an input, internal measures are used to assess the quality of clustering based on intrinsic information. Stability measures assess clustering consistency by comparing it with the clusters obtained after each column is removed. However, clustering methods should be aligned with the goals pursued by researcher-practitioner since clustering methods alone may not give appropriate number of clusters for accurate pricing of credit derivatives. For example, if an estimation of sum of squares within groups using k-means cluster analysis and validation measures suggest 4 clusters of credit spreads, this number of clusters may not enough to estimate CDO fair spreads for each tranche given the particular level of tolerance. This is because we still may have high values for the within-group sum of squares, low level of homogeneity in each cluster or, equivalently, high level of heterogeneity within the clusters. High level of heterogeneity requires further increase in number of clusters to achieve convergence of estimated CDO fair spreads to the real one. The optimal number of clusters depends of the tranche being priced, credit spread distribution of the credit portfolios, and the recovery rates of the obligors. In practice, the number of clusters may be greater than the one suggested by, for example, the k-means cluster analysis; however this number of clusters still will be several times less than the number of obligors in the credit portfolio thereby increasing the speed of loss distribution generation. The disadvantage of using cluster analysis is that additional time is required to perform such analysis. The movements of credit spreads should be monitored, and the cluster analysis might be required to be performed after each switch regime change.

In risk management of tranches, loss distributions for the credit portfolios need to be extensively recalculated for estimation of credit risk measures such as, for example, idiosyncratic delta of a credit or a group of credits. One of the methods to speed up the building of loss distribution is perturbation method suggested by O'Kane (2008, pp. 362-364). The perturbation method coupled with partitioning of credit spreads can help further speed up the loss distribution building and allows more efficient estimation of sensitivity of CDO tranche prices to the instantaneous movement in credit spreads of the cluster. This is because such approach is more consistent with what we usually observe in the market when changes in spreads occur in a group of highly correlated spreads rather than in one particular credit.

Not all methods may be easily implemented in situations when the portfolio is heterogeneous. For example, in the well known algorithm proposed by Andersen et al. (2003) one has to estimate the value of common divisor of all potential losses. The algorithm depends on the statistical characteristics of the credit spreads (which can be diverse with high level of heterogeneity) and in some cases the value of the common divisor can be very small substantially affecting the speed of the algorithm and requiring more memory for estimation of the loss distribution. A tolerance level of the common divisor of all potential losses can be chosen to speed up the process, but in this case the resulting loss distribution depends on the chosen tolerance level, which may incur discreteness of the loss distribution shape and even may not be accurate affecting credit risk management and credit derivatives pricing.

Easiness of implementation of specific algorithms as well as accuracy depends on the access of specialized mathematical libraries. For example, to use a complex FFT algorithm for loss distribution generation, an analyst would need an access to an effective FFT algorithm. Moreover, this algorithm requires more time for generation of loss distribution than the recursive algorithms described above and this one of the main reasons why FFT algorithms are not widely used in practice. When estimating a loss distribution using finite homogeneous portfolio, an analyst needs to choose an effective integration method as well as the integration ranges to reach the required tolerance.

Generation of loss distribution should be aligned with other quantitative and qualitative analysis of the obligors and informed decision should be made in order to properly estimate the input parameters. For example, credit quality of each credit can change over the specified period of time, and, therefore, it can affect the loss distribution that will change during this period of time due to changes in credit quality of the obligors.

The analyzed methods assumed constant recovery rates. The value for recovery rates is often assumed to be equal to 40%, during financial crisis even less – 10 to 20%. Moreover, these values are often assumed to be the same for all the obligors with different ratings and credit spread term structure. Estimation of recovery rates for each obligor based on financial statement analysis, for example, estimation of recovery rates for a portfolio consisting of 140 obligors is time consuming; on the other side, the generated loss distribution based on the carefully estimated recovery rates would be more realistic.

Since recovery rates are changing over the particular time period, the loss distribution generation methods can be improved by assuming that recovery rates for each obligor (and, therefore, loss given default) are random variables. The probabilities of default and recovery rates are negatively correlated and they can be exogenously modeled as stochastic processes which then can be incorporated into existing loss distribution generation models.

The described methods don’t consider credit spread distribution of the credit portfolio. Loss distribution of the portfolio or CDO tranche depend on the credit spread distributions observed in the market. Analysis of credit spread distributions can help identify which credit spreads are most likely affect specific CDO tranche, and which credit spreads wouldn’t. For example, for a given tenor, say 3 years, a credit portfolio may have credits with low credit spreads concentrating around the value giving time to default which are more than 10 years. Such credit spreads don't affect loss distribution of the considered CDO tranche, and, therefore can be excluded from consideration. If a number of such credit spreads is very high, then by excluding these credits from consideration we can substantially decrease computational time of the loss distribution generation. In other words, loss distribution generation can be faster and substantially improved, if we perform preliminary statistical analysis of credit spread distribution, recovery rates and default probabilities of the obligors.

References

Andersen, L. J., Sidenius, J. & Basu, S. (2003, November). All your hedges in one basket. *Risk*, 67-72.

Arvanitis, A. & Gregory, J. (1999). A credit risk toolbox. In D. Shimko (ed.), *Credit risk. Models and management.* pp. 283 - 289. London: Risk books.

Arvanitis, A. & Gregory, J. (2001). *Credit: The complete guide to pricing, hedging and risk management.* London: Risk Books.

Dembo, A., Deuschel, J., & Duffie, D. (2004). Large portfolio losses. *Finance & Stochastics*, *8*(1), 3-16.

Glasserman, P. & Li, J. (2003). Importance sampling for a mixed Poisson model of portfolio credit risk. In: S. Chick, P.J. Sanchez, D. Ferrin, and D.J. Morrise (Eds.), *Proceedings of the 2003 Winter Simulation Conference*, pp. 267 – 275.

Glasserman, P. & Ruiz-Mata, J. (2006). A comparison of approximation techniques for portfolio credit risk. *Journal of Credit Risk 2*, 33-66.

Glasserman, P. (2004). Tail approximations for portfolio credit risk. *Journal of Derivatives 12*, 24-42.

Gordy, M. (2003, July). A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, *12*(3), 199 – 232.

Gordy, M. (2004). Granularity adjustment in portfolio risk measurement. In G. Szegö (ed.), *Risk Measures for the 21st Century*. pp. 109-121. John Wiley & Sons.

Gordy,M.B. (2002). Saddlepoint approximation of CreditRisk+. *Journal of Banking and Finance 26*(7), 1337-1355.

Gregory, J., & Laurent, J. (2003, June). I will survive. *Risk*, *16*(6), 103-107.

Gregory, J., & Laurent, J. (2004, October). In the core of correlation. *Risk 17*(10), 87-91

Grundke, P.(2007). Computational aspects of integrated market and credit portfolio models. *OR Spectrum*, 29(2), 259 – 294.

Hull, J. & White, A. (2004). Valuation of a CDO and an nth to default CDS without Monte Carlo simulations. *Journal of Derivatives, 12*(2), 8-23

Kalemanova, A., Schmid, B. & Werner, R. (2007, September). The normal inverse Gaussian distribution for synthetic CDO Pricing. *Journal of Derivatives*, *14*(3), pp. 80-93.

Krekel, M. & Partenheimer, J. (2006). The implied loss surface of CDOs. Retrieved January 16, 2009, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=950376

Laurent, J., & Gregory, J. (2005, Summer). Basket default swaps, CDOs and factor copulas. *Journal of Risk*, *7*(4), 1-20.

Laurent, J.-P. (2004). *Fast analytic techniques for pricing synthetic CDOs*. Presentation at Credit Risk Summit Europe. October 13, 2004.

Martin, R. (2006, December). The saddlepoint approximation and portfolio optionalities. *Risk*, 93-95.

Merino, S. & Nyfeler, M. (2002, August). Calculating portfolio loss. *Risk 15*(8), 82-86

O'Kane, D. (2008). *Modeling single-name and multi-name credit derivatives*. New York, NY: John Wiley & Sons

Pykhtin, M. (2004, March). Multi-factor adjustment. *Risk 17*(3). 85 – 90.

Reiβ, O. (2003). *Mathematical methods for the efficient assessment of market and credit risk.* PhD Thesis. Department of Mathematics, University of Kaiserslautern.

Schloegl, L., & O'Kane, D. (2005, October). A note on the large homogeneous portfolio approximation with the Student-t copula. *Finance & Stochastics*, *9*(4), 577-584

Schönbucher, P.J. (2002). *Taken to the limit: Simple and not so simple loan loss distributions.* Working Paper, Department of Statistics, Bonn University.

Sidenius, J., Piterbarg, V., & Andersen, L. (2008, March). A new framework for dynamic credit portfolio loss modeling. *International Journal of Theoretical & Applied Finance*, *11*(2), 163-197.

Vasicek, O. (1987). Probability of loss on loan portfolio. Working paper, KMV Corporation.

Vasicek, O. (2002, December). Probability of loss on a loan portfolio. *Risk*, 160 – 162.

Zheng, H. (2007). Portfolio distribution modeling and computation. *Workshop on Fast Financial Algorithms.* Tanaka Business School & Imperial College. Retrieved January 20, 2009, from http://www3.imperial.ac.uk/pls/portallive/docs/1/18529698.PDF