

# A discrete model for correlated default times and its application to CDO

*Ivan Mitov*

Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria

**Keywords:** Default times. Default correlation. Dependence. Portfolio Loss Distribution. CDO.

**AMS:** 62P05

## Abstract

A simple discrete time model for the dependence between the default times of assets (firms) in a given portfolio is discussed. The model is based on multivariate geometric distributions. A possible way for calibration of the model is discussed too. In the framework of this model portfolio loss distribution is calculated and a Monte Carlo based estimation of the expected losses in a given CDO is done.

## 1. Introduction

The problem for modelling default time distribution and the correlation between the times to default of assets (firms) in a given portfolio is of great interest. The investor is exposed to the risk of default of one or more firms. It is clear that the default of some of the securities depend on different factors. These factors can be separated into three main types. Factors, specific to each firm. Common factors like currency rates, interest rates, macroeconomic factors etc. The third type of factors are those specific for a set of firms, for example firms in a given industry sector. Different models for default correlation are discussed in the literature (see e.g. [3], [4]). The mathematical basis of CDO is also discussed widely (see e.g. [1]). In the talk we represent a simple discrete time model for default times of the assets, based on the multivariate geometric distribution. The model is related to that considered in [2] in continuous time. A possible application of the model to the CDO losses is considered. In the last section we represent a Monte-Carlo simulation for the aggregate losses in a given CDO.

## 2. Model Description

Let us have a portfolio consisting of  $N \geq 2$  assets (firms). Assume that:

1. The time is discrete  $t = 0, 1, 2, \dots$ . The calendar year consists of  $T > 2$  periods (units of time);

2. The default event of an asset (firm) can be caused by different types of shocks. In general, we assume that  $M = \sum_{l=1}^N \binom{N}{l} = 2^N - 1$  types of shocks

can occur:  $\binom{N}{1} = N$  shocks specific for each firm separately;  $\binom{N}{2}$  shocks specific for each pair of firms ...  $\binom{N}{N} = 1$  shock which causes a default event to all the firms in the portfolio.

3. The instances of shocks  $\xi_1, \xi_2, \dots, \xi_M$  are independent random variables with distributions  $\Pr(\xi_j = k) = q_j^{k-1} p_j$ , where  $0 < p_j < 1$ ,  $q_j = 1 - p_j$ ,  $k = 0, 1, 2, \dots$ ,  $j = 1, 2, \dots, M$ .

Define the matrix  $A = ||a_{ij}||_{N \times M}$  such that each row corresponds to a firm, and each column corresponds to a shock. If the  $j$ -th shock causes the default event to  $i$ -th firm then  $a_{ij} = 1$ , and  $a_{ij} = 0$  otherwise. The time to default  $\tau_i$  of the  $i$ -th firm is the time when the first possible shock occurs, i.e.

$$\tau_i = \min_{a_{ij}=1, j=1, 2, \dots, M} \{\xi_j\}, \quad i = 1, \dots, N.$$

The times to default are not independent in general, but they also have geometric distributions.

**Proposition 1** *The marginal distribution of  $\tau_i$  is*

$$\Pr(\tau_i = t) = Q_i^{t-1} P_i, \quad t = 1, 2, \dots, \quad Q_i = \prod_{j=1}^M q_j^{a_{ij}}, \quad P_i = 1 - Q_i, \quad i = 1, 2, \dots, N.$$

*The joint survival function of  $(\tau_1, \tau_2, \dots, \tau_N)$  is*

$$s(t_1, \dots, t_N) = \Pr(\tau_1 > t_1, \dots, \tau_N > t_N) = \prod_{j=1}^M q_j^{\max(a_{1j}t_1, \dots, a_{Nj}t_N)}.$$

The distribution of the vector  $(\tau_1, \tau_2, \dots, \tau_N)$  is known as  $N$ -variate geometric distribution. In general the model contains  $M = 2^N - 1$  parameters which have to be estimated. On the other hand the information available from the rating agencies (e.g. Standard & Poors, Moody's KVM) is restricted to the default probability for particular period  $T$ , i.e.  $\Pr\{\tau_i \leq T\}$  for each firm, and the so called Moody's KVM default correlation, which is in fact the correlation coefficient between the indicators  $I_{\{\tau_i \leq T\}}$  and  $I_{\{\tau_j \leq T\}}$ . Due to this reason, the parameters can be estimated only in some particular cases of the general model.

### 3. Common shock model

Consider the following case: Each firm is exposed to two types of shocks. The first shock is firm specific and it does not depend on the condition of the other firms. The second one is economy-wide shock event. Its occurrence leads to default of all firms. In this case  $M = N + 1$  and the matrix  $A = ||a_{ij}||_{N \times M}$  reduces to  $a_{ii} = 1$ ,  $a_{iM} = 1$ ,  $i = 1, 2, \dots, N$ , and  $a_{ij} = 0$  otherwise. Denoting

by  $\xi_j, j = 1, 2, \dots, N$  the instants when the specific shocks occur and  $\xi$  - the instant when the economy-wide shock occurs we have

$$\tau_i = \min(\xi_i, \xi), \quad i = 1, \dots, N.$$

Suppose that  $\xi_j \sim Ge(p_j), 0 < p_j < 1, q_j = 1 - p_j, j = 1, \dots, N, \xi \sim Ge(p), 0 < p < 1, q = 1 - p$ . Now, Proposition 1 (with the corresponding changes in notations) gives the marginal survival function

$$s(t_i) = \Pr(\tau_i > t_i) = (q_i q)^{t_i}, t_i = 1, 2, \dots; i = 1, 2, \dots, N. \quad (1)$$

Under the conditions in this case the correlation coefficient between the indicators  $I_{\{\tau_i \leq T\}}$  and  $I_{\{\tau_j \leq T\}}$  has the following form

$$\rho(I_{\{\tau_i \leq T\}}, I_{\{\tau_j \leq T\}}) = \frac{(q_i q_j q)^T - (q_i q)^T (q_j q)^T}{\sqrt{(q_i q)^T (q_j q)^T (1 - (q_i q)^T) (1 - (q_j q)^T)}}. \quad (2)$$

It was mentioned above that the calibration of the model is a difficult problem since the lack of enough available information. We calibrate the model using the Moody's KMV probabilities  $\hat{p}_i, i = 1, \dots, N$  for one year default and the so called Moody's KMV default correlation  $\hat{\rho}_{i,j}, i = 1, \dots, N - 1, j = i + 1, \dots, N$ , which is in fact the correlation coefficient between the indicators. It was assumed earlier that the year is divided into  $T$  periods. Using the available data and equations (1) and (2) the following system is obtained

$$\hat{p}_i = P(\tau_i \leq T) = 1 - (q q_i)^T, i = 1, 2, \dots, N,$$

$$\hat{\rho}_{i,j} = \rho(I_{\{\tau_i \leq T\}}, I_{\{\tau_j \leq T\}}), i = 1, 2, \dots, N - 1, j = i + 1, \dots, N.$$

The system contains  $N(N + 1)/2$  equations for  $N + 1$  unknown parameters. The least square method gives the following approximate solution

$$q = \frac{2}{N(N - 1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left( \frac{(1 - \hat{p}_i)(1 - \hat{p}_j)}{\hat{\rho}_{i,j} \sqrt{\hat{p}_i(1 - \hat{p}_i)\hat{p}_j(1 - \hat{p}_j)} + (1 - \hat{p}_i)(1 - \hat{p}_j)} \right)^{\frac{1}{T}}$$

$$q_i = (1 - \hat{p}_i)^{\frac{1}{T}} q^{-1}, \quad i = 1, 2, \dots, N. \quad (3)$$

#### 4. What is the CDO

A Collateralized Debt Obligation (CDO) is an asset backed security whose collateral consists mainly of a portfolio of defaultable instruments like loans, junk bonds, mortgages, etc. The main economic reason for the existence of CDOs is to address market imperfections. The bank issues notes with different subordination, so called tranches, that are sold to investors. The principal payment and interest income (LIBOR + spread) are allocated to the notes

according to the following rule: *senior* notes are paid before *mezzanine* and lower rated notes. Any residual cash is paid to the *equity* notes. The equity note is also called the first-loss position because it is the first to be affected by a default in the portfolio. Therefore it offers a larger spread than the more senior notes. In practice we can say that the CDO investors are selling protection, choosing a tranche according to their risk - return preference. A consequence of this payment subordination is that a senior note is less risky than the collateral. As we move down the CDO's notes, the risk level increases. Each tranche, except for the equity, is rated by credit rating agencies. The senior notes usually have Aaa/AAA rating. The originating bank has an incentive to keep a significant share of the equity tranche to demonstrate its commitment to the CDO performance.

## 5. Portfolio Loss Distribution

We will thereafter consider a CDO with some given maturity  $T_m$ . It is based upon  $N$  reference credits,  $j = 1, \dots, N$ , with nominals  $\nu_j$ ,  $i = 1, \dots, N$  and maturity also equal to  $T_m$ . Let  $\delta_j$  be the recovery rate for credit  $j$  and  $(1 - \delta_j)\nu_j$  is the corresponding Loss Given Default. As above, we denote by  $\tau_j$  the default time of asset  $j$  and by  $I_j(t) = I_{\{\tau_j \leq t\}}$  the corresponding counting process. Let  $L(t)$  be the cumulative loss on the credit portfolio at time  $t$ , i.e.

$$L(t) = \sum_{j=1}^N \nu_j (1 - \delta_j) I_j(t) \quad (4)$$

which is also a pure jump process. Assuming  $\nu_j$  and  $\delta_j$ ,  $j = 1, 2, \dots, N$  to be constants, we can see that the possible values of  $L(t)$  are the numbers of the following type  $\sum_{j=1}^k \nu_{i_j} (1 - \delta_{i_j})$  for all subsets  $\{i_1, i_2, \dots, i_k\}$  of the set  $\{1, 2, \dots, N\}$ . It is the case when no defaults occur in the interval  $[0, t]$ ,  $k = 0$  and the sum is equal to zero. Suppose  $0 \leq k \leq N$  and  $\{i_1, i_2, \dots, i_k\}$  is a subset of  $\{1, 2, \dots, N\}$  and  $\{j_1, j_2, \dots, j_{N-k}\}$  is its complement subset. Then the portfolio loss distribution is given by

$$\Pr \left( L(t) = \sum_{j=1}^k \nu_{i_j} (1 - \delta_{i_j}) \right) = \left( q \prod_{l=1}^{N-k} q_{j_l} \right)^t \prod_{s=1}^k (1 - q_{i_s}^t)$$

Under the following additional assumptions  $\lambda_1 = \lambda_2 = \dots = \lambda_N = \mu$ ,  $\nu_1 = \nu_2 = \dots = \nu_N = \nu$ , and  $\delta_1 = \delta_2 = \dots = \delta_n = \delta$ , the above formula takes a rather more simple form

$$\Pr(L(t) = k\nu(1 - \delta)) = \binom{N}{k} q^t q_1^{(N-k)t} (1 - q_1^t)^k.$$

## 6. Cash flow structure

Assume further that the CDO has three tranches, denoted by equity, mezzanine and senior. Usually any initial portfolio losses affect the holders of the equity tranche. It means that the investors get their regular payments until a default occurs. After a default occurrence the equity tranche payments are reduced according to the size of the default loss. If the losses reach a particular predefined threshold  $A$ , the equity tranche holders don't receive payments anymore. Any further losses are covered by the mezzanine tranche investors. If the cumulative losses exceed other predefined threshold  $B > A$ , then the payments to the mezzanine are stopped too. All losses over  $B$  are covered by the senior tranche. For the thresholds  $0 \leq A < B \leq \sum_{j=1}^N \nu_j$ , and usually  $3\% \leq \frac{A}{\sum_{j=1}^N \nu_j} \leq 6\%$  and  $13\% \leq \frac{B}{\sum_{j=1}^N \nu_j} \leq 16\%$ . Denote by  $E(t), M(t), S(t)$  the aggregate losses for the three tranches respectively. These aggregate losses can be expressed in terms of portfolio loss as follows

$$E(t) = \begin{cases} L(t), & \text{if } L(t) \leq A, \\ A, & \text{if } L(t) > A, \end{cases} \quad M(t) = \begin{cases} 0, & \text{if } L(t) \leq A, \\ L(t) - A, & \text{if } A < L(t) \leq B, \\ B - A, & \text{if } L(t) > B, \end{cases}$$

and

$$S(t) = \begin{cases} 0, & \text{if } L(t) \leq B, \\ L(t) - B, & \text{if } L(t) > B. \end{cases}$$

## 7. Monte Carlo simulation

In this section a simulation of the aggregate losses in the different CDO tranches is done. We are interested in the dependence of the losses from the correlation between default times of the assets. Assume the CDO is based on a portfolio with  $N = 100$  assets each with \$100 millions notional. Let the recovery rate for each firm be  $\delta = 4\%$  of the notional. The thresholds are chosen  $A = 5\%$  and  $B = 15\%$ . Suppose the individual shocks have the same probability  $p_1 = p_2 = \dots = p_{100}$  and the economy-wide shock occurs with probability  $p$ . The estimates of probabilities  $p_1, q_1, p$ , and  $q$  are given by the equations (3), which are reduced to  $q = 1 - p = ((1 - \hat{p}) / (1 - \hat{p}(1 - \hat{\rho})))^{1/T}$  and  $q_1 = 1 - p_1 = (1 - \hat{p}(1 - \hat{\rho}))^{1/T}$ . Given the estimates  $q$  and  $q_1$  proceed to the following algorithm for mean aggregate loss estimation in the different CDO tranches:

- 1) Generate  $N$  independent, geometric r.v.'s  $\xi_1, \xi_2, \dots, \xi_N$  with parameters  $(p_1, q_1)$  and a geometric r.v.  $\xi$  with parameter  $(p, q)$ ;
- 2) Find  $\tau_i = \min(\xi_i, \xi), i = 1, \dots, N$ ;
- 3) Find the number of the times  $\tau_i$ , which are less than  $T$ ;

- 4) Calculate  $L(T)$  and respectively  $E(T)$ ,  $M(T)$  and  $S(T)$ ;  
 5) Repeat steps 1) - 5)  $m$  times and finally calculate

$$\hat{E}(T) = \frac{1}{m} \sum_{k=1}^m E^{(k)}(T), \hat{M}(T) = \frac{1}{m} \sum_{k=1}^m M^{(k)}(T), \hat{S}(T) = \frac{1}{m} \sum_{k=1}^m S^{(k)}(T)$$

where  $E^{(k)}(T)$ ,  $M^{(k)}(T)$  and  $S^{(k)}(T)$  for  $k = 1, \dots, m$  are the aggregate losses for the three tranches in each of the  $m$  realizations. The results are given in the table below for  $\hat{\rho} = 0.03, 0.1, 0.3, 0.5, 0.7, 0.9$  and  $m = 50000$ .

	$\hat{\rho} = 0.1$	$\hat{\rho} = 0.3$	$\hat{\rho} = 0.5$	$\hat{\rho} = 0.7$	$\hat{\rho} = 0.9$
Equity	99.81%	97.66%	85.34%	58.52%	28.71%
Mezzanine	29.20%	16.57%	9.58%	10.35%	12.98%
Senior	0.84%	2.39%	3.95%	5.47%	6.87%

The numbers are displayed as a percentage of the maximum possible loss in each tranche. The results in the table show that the losses in the different tranches strongly depend on the correlation between default times.

**Acknowledgements:** This work is partially supported by NFSI grant No.VU-MI-105/2005.

## 8. Bibliography

- [1] D. Duffie and N. Garleanu, (2001), *Risk and Valuation of Collateralized Debt Obligations.*, Financial Analyst's Journal, 57, 41-59.  
 [2] K. Giesecke, (2002), *An exponential model for dependent defaults*, (working paper), Humboldt - Universitat zu Berlin.  
 [3] D. Li, (2000), *On default correlation: A copula approach*, Journal of Fixed Income 9(4), 43-54.  
 [4] Fan Yu, (2003), *Dependent default in intensity-based models*, (working paper), University of California, Irvine.