Credit risk models

Jean-Paul Laurent
laurent.jeanpaul@free.fr
ISFA Actuarial School, Université Lyon I
50, Avenue Tony Garnier, 69007 LYON, FRANCE

August 2007
Revised: March 2008

Keywords: credit risk, loss distributions, individual model, Gaussian copula, factor models

Introduction

The computation of the distribution of aggregate losses in credit portfolios has become especially important for risk management and securitization purposes. Banks and financial institutions need to assess the risks within their credit portfolios both for regulatory requirements and for internal risk management: for instance, they may compute risk measures such as the Value at Risk or the Expected Shortfall associated with their credit risk exposure. Credit risk has been transferred from banks to other investors such as insurance companies or hedge funds which act with respect to commercial banks as reinsurance companies with respect to insurance firms. This securitization process involves some "tranching" of credit risk similar to stop-loss reinsurance. Specific products such as credit default swaps and single tranche Collateralized Debt Obligations (CDOs thereafter, see Tavakoli [1] for a description of the CDO market) have been designed for such risk transfer. The computation of the CDO tranche premiums also involves the aggregate credit loss distributions over different time horizons. The purpose of this article is to describe the most commonly used approach to compute such loss distributions. Before this, we briefly review the risks involved in a credit portfolio

I) An overview of risks involved

I.1) Default risk

Default risk is related to the inability of a borrower to reimburse a loan or a bond. More precisely, one can think of different credit events such as:

- the failure to meet payment obligations when due,
- bankruptcy (for non-sovereign entities) or Moratorium (for sovereign entities only),
- repudiation,
- material adverse restructuring of debt,
- obligation acceleration or obligation default.

The definition of a credit event hinges on the relevant bankruptcy rules that depend themselves upon geographical regions, quality of the borrower and seniority of the loan. For example, personal bankruptcy could be accessed more easily in the US than in the UK. The same comparison also holds for corporates. This will result in differences in frequency and claim severity. Let us additionally remark that as far as retail credit are concerned, the banks need to decide at which stage a loan is actually in default. For instance, a single outstanding payment on a mortgage is not usually considered as a default event. We refer to de Malherbe [2] for a detailed quantitative assessment of the credit default swaps covenants.

I.2) Change in credit quality: credit migrations, change in credit spreads

In addition to default risk, credit migrations are associated with changes in credit quality. For example, in financial markets, even if default-free interest rates remain constant, defaultable bond prices change prior to default. This is also known as credit spread risk. For example, an increase in default probability following a credit rating downgrading, adverse earnings announcements, should lead to a drop in the company's bond prices. This is not specific to corporates; default probabilities for individuals usually increase when the borrower becomes unemployed.

Let us emphasize that credit contracts may have much longer maturities than insurance contracts. For example, in a typical 30 Y mortgage, the annuities remain fixed even if the credit quality of the borrower deteriorates: the lender cannot raise payments yearly as in the insurance markets.

I.3) Recovery risk

In case of default, the lender only gets a fraction of the promised payments. In the most severe cases, no further cash-flows are being paid by the borrower. In case of bankruptcy, the assets of the defaulted firm are being sold to investors and the proceeds are used to reimburse the lenders. After liquidation, the proceeds may excess the commitments to the lenders, which means that the loss severity can be equal to zero. For simplicity, we will thereafter consider that the loss given default (LGD) is equal to the difference between the face value of the loan or the bond and the recovered value. As a consequence, losses given default are bounded.

I.4) Dependence between defaults

Since common macroeconomic factors, such as business cycles, level of unemployment, shifts in monetary policy drive both default frequency and credit loss severity, we cannot rely on the standard insurance framework and we need to cope with dependence between default events. Macroeconomic shocks can be considered as exogenous to defaults are usually lead to positive dependence. Dependence can also occur at a local level due to interactions between firms. For example, default of a bank can be seen as a bad signal with respect to the assets of its competitors in an incomplete information framework. Conversely, default of a firm could raise the market power of competitors, leading to extra-profits and an increase in credit quality. As can be seen from these examples, defaults can be informative with respect to the credit quality of survivors; this is known as contagion effects or "infectious defaults".

We will further detail the most commonly used approaches to the dependence between defaults as far as the computation of loss distributions is involved. Dependence usually results in much fatter tails and an increase in senior tranches CDO premiums, which correspond to stop-loss reinsurance premiums in insurance markets.

II) Computation of loss distributions

II.1) Marginal default probabilities

The individual default probabilities of the obligors within a credit portfolio are a key input in the computation of aggregate loss distributions.

As far as retail portfolios are involved, banks could either use internal data or if available public information about defaults and obligors individual characteristics. Usually, obligors are grouped in homogeneous clusters and thus marginal default probabilities depend upon static observables, such as age, marital status and so on. Credit scoring techniques based upon logistic regression or classification techniques are routinely used. For instance in the US, FICO scores have almost become a market standard.

When it comes to corporate obligors, the number of obligors is usually much smaller, and especially for investment grade bonds. Thus, the number of observed defaults is much smaller which leads to some statistical difficulties. To assess the magnitude of default probabilities, creditors need to rely on the analysis of financial statements. This can be done at the lender's level, usually within credit department of lending banks or thanks to the rating agencies. While the financial analysis is forward looking and rates obligors along a discrete grid, further data analysis is required to translate credit ratings into default probability as illustrated by Moody's historical default rates of bond issuers.

Another approach that can be thought of, especially for publicly traded companies, is the use of the Stock market. In the structural model of Merton, default occurs if the market value of the company's assets falls below a prespecified threshold. The lower the value of the stock, the smaller the distance to the default barrier and the more likely default will occur. This is the route followed by Kealhofer [3].

Historical and market implied default probabilities

In frictionless and arbitrage-free financial markets, prices of financial securities are computed as the expectation of the discounted risky cash-flows. Let us consider a simplified defaultable bond with a single promised cash-flow of 1 \$ to be paid at time *t*. In case of default, we assume that no payment will be made (recovery rate is equal to zero). Let us denote by $\bar{B}$ the market price for such a bond and by *r*, the (continuously compounded) discount rate for maturity *t*. $\bar{B} = e^{-rt} Q(\tau > t)$, where τ denotes the default date and $Q(\tau > t)$ is the survival probability. If the previous defaultable discount bond is actually traded, the survival probability can be readily computed as $Q(\tau > t) = \bar{B} \times e^{rt}$. It appears that the default probabilities extracted from bond prices are higher that default probabilities computed from historical data on defaults. Investors require a risk premium to hold defaultable bonds, reflecting imperfect diversification of default risk: if default events of obligors were independent and the number of obligors very large, default risk could be perfectly diversified thanks to the law of large numbers. This is the classical framework in insurance theory and in such a competitive market, the computation of bond prices should follow the "pure premium actuarial rule", meaning that default probabilities implied from financial markets should equal the "historical default probabilities". For high quality names, say AAA bonds, default are rare events, the number of traded bonds small and the dependence between default events is larger than for high yield names since extreme macroeconomic factors drive defaults of such firms.

Not surprisingly, the ratio of market implied and historical default probabilities is larger for these high quality names compared with speculative bonds.

Credit risk measurement usually involves historical default probabilities, while securitization and risk transfer (typically the pricing of CDO tranches) requires the use of market implied probabilities. We will thereafter denote by $F_1, \ldots, F_n$ the marginal distribution functions associated with the names in the credit portfolio, that can be either historical or market implied depending on the context.

II.1) Individual and collective models

Let us denote by $\tau_1, \ldots, \tau_n$ the default dates of *n* obligors, by $N_1(t) = 1_{\{\tau_1 \le t\}}, \ldots, N_n(t) = 1_{\{\tau_n \le t\}}$ the corresponding default indicators for some given time horizon *t* and by LGD_i the loss given default on name *i*. We assume that the maximum loss is normalized to unity, the aggregate loss on the credit portfolio for time horizon *t* is then given by:

$$L(t) = \frac{1}{n} \sum_{i=1}^{n} LGD_i \times N_i(t)$$

Let us remark that we only cope with defaults (or "realized losses") and not losses due to changes in credit quality of non defaulted bonds. As can be seen from the previous equation, the standard credit risk model is an individual model.

Structural models and Gaussian copulas

Inspired by the structural approach of Merton, defaults occur whenever assets fall below a prespecified threshold (see Gupton et al [4], Finger [5], Kealhofer [3]). In the multivariate case, dependence between default dates ensues dependence between asset price processes. On the other hand, the most commonly used approach states that default dates are associated with a Gaussian copula (Li [6]). Thus, default indicators follow a multivariate Probit model. For simplicity, the Gaussian copula model can be viewed as a one period structural model. Consequently, Hull, Pedrescu & White [7] show that from a practical point of view, the copula and structural approaches lead to similar loss distributions. Burtschell *et al.* [8] point out that in many cases, the computation of the loss distribution is rather robust with respect to the choice of copula.

For simplicity, we will thereafter assume that LGD_i is non stochastic. We will thus concentrate upon the modeling of dependence between default dates rather than upon the recovery rates. For large credit portfolios, taking into account stochastic losses given defaults has only a small impact on the loss distribution provided that losses given defaults are independent of default indicators. Altman *et al.* [9] analyze the association between default and recovery rates on corporate bond over the period 1982-2002; they show negative dependence, i.e. defaults are more severe and frequent during recession periods. An analysis of the changes in the loss distributions due to recovery rates possibly correlated with default dates is provided in Frye [9], [10] or in Chabaane *et al.* [11]; this usually results in fatter tails of the loss distributions.

The assumption that default indicators are independent given a low dimensional factor is another key ingredient in credit risk models (see Wilson [12], [13], Gordy [14], Crouhy *et al.* [15], Pykhtin & Dev [16], Frey & McNeil [17]). This dramatically reduces the numerical complexities when computing loss distributions. While commercial packages usually involve several factors, we will further restrict to the case of a single factor which eases the exposition. This is also the idea behind the regulatory Basel II framework (see Gordy [18]). Thus, in the Gaussian copula (or multivariate Probit) approach, the latent variables associated with default indicators $N_i(t) = 1_{\{\tau_i \le t\}}, i = 1,\dots,n$ can be written as: $V_i = \rho_i V + \sqrt{1-\rho_i^2}\bar{V}_i, i = 1,\dots,n$ where $V, \bar{V}_1,\dots,\bar{V}_n$ are independent standard Gaussian variables. The default times are then expressed as: $\tau_i = F_i^{-1}\left(\Phi(V_i)\right)$ where Φ denotes the Gaussian cdf. In other words default of name *i* occurs before *t* if and only if $V_i \le \Phi^{-1}\left(F_i(t)\right)$. It can be easily checked that default dates are independent given the factor *V* and that the conditional default probabilities can be written as $P\left(\tau_i \le t \middle| V\right) = \Phi\left(\frac{\Phi^{-1}\left(F_i(t) - \rho_i V\right)}{\sqrt{1-\rho_i^2}}\right)$, $i = 1,\dots,n$. Let us remark that thanks to the theory of stochastic orders, increasing any of the correlation parameters $\rho_1,\dots,\rho_n$ leads to an increase in the dependence of the default times $\tau_1,\dots,\tau_n$ with respect to the supermodular order. The corresponding copula of default times is known as the one factor Gaussian copula.

Clearly, determining the correlation parameters is not an easy task especially when the number of names involved in the credit portfolio is large. Extra simplicity consists in grouping names in homogeneous portfolios with respect to sector or geographical region. We refer to Gregory & Laurent [19] for a discussion of this approach. The easier to handle approach consists in assuming some kind of homogeneity at the portfolio level. For instance, we can assume that the correlation parameter is name independent. This is known as the “flat correlation” approach. This both underlies the computations of risk measures in the Basel II agreement framework and of CDO tranches premiums.

Let us also remark that when the marginal default probabilities are equal, i.e. $F_1(t) = \dots = F_n(t)$, then the default indicators $N_1(t),\dots,N_n(t)$ are exchangeable. Conversely, when the default indicators are exchangeable, one can think of using de Finetti’s theorem which states the existence of univariate factor such that the default indicators are conditionally independent given that factor. In other words, for “homogeneous portfolios”, the assumption of a one dimensional factor is not restrictive. The only assumption to be made is upon the distribution of conditional default probabilities. We refer to Burtschell *et al.* [8], [20] for some discussion of different mixing distributions.

Computation of loss distributions

Let us now discuss the computation of aggregate loss distribution in the previous framework. The simplest case corresponds to the previous homogeneous case. We then denote by $\tilde{p}_t = P\left(\tau_i \le t \middle| V\right)$ the unique conditional default probability for time horizon *t*. Thanks to the homogeneity assumption, the probability of *k* defaults within the portfolio ($k = 0,1,\dots,n$) or equivalently the probability that the aggregate loss $L(t)$ equals $\frac{k}{n}$ can be written as $\binom{n}{k}\int \tilde{p}^k \left(1-\tilde{p}\right)^{n-k} \nu_t\left(d\tilde{p}\right)$ where ν_t is the distribution of $\tilde{p}_t$. In other words, the loss distribution is a binomial mixture. Let us denote by φ the density function of a standard Gaussian variable. We can equivalently write $P\left(L(t)=\frac{k}{n}\right)=\int_{\mathbb{R}} P\left(\tau_i \le t \middle| v\right)^k \left(1-P\left(\tau_i \le t \middle| v\right)\right)^{n-k} \varphi(v)dv$ which can be computed numerically thanks to a Gaussian quadrature.

An interesting feature of the above approach is the simplicity of distributions for large homogeneous portfolios. Thanks to de Finetti's theorem, the aggregate loss $L(t)$ converges almost surely and in mean to $\tilde{p}_t = \Phi\left(\frac{\Phi^{-1}\left(F(t)-\rho V\right)}{\sqrt{1-\rho^2}}\right)$ as the number of names tends to infinity. In the credit risk context this idea was firstly put in practice by Vasicek [21]. It is known as the Large Portfolio Approximation. Further asymptotic developments, such as the saddlepoint expansion techniques have been used, starting from Martin *et al.* [22].

Let us now consider the computation of the aggregate loss distribution for a given time horizon within the Gaussian copula framework without any homogeneity or asymptotic approximation. This is based on the computation of the characteristic function of the aggregate loss. We further denote by $\varphi_{L(t)}(u)=E\left[\exp\left(iuL(t)\right)\right]$. Thanks to the conditional independence upon the factor V, we can write $\varphi_{L(t)}(u)=\int_{\mathbb{R}}\left(\prod_{j=1}^{n}\left(1+\Phi\left(\frac{\Phi^{-1}\left(F_j(t)-\rho_j v\right)}{\sqrt{1-\rho_j^2}}\right)\left(e^{iu\frac{LGD_j}{n}}-1\right)\right)\right)\varphi(v)dv$. The previous integral can be easily computed by using a Gaussian quadrature. Let us also remark that the computation of the characteristic function of the loss can be adapted without extra complication when the losses given default LGD_i are stochastic but (jointly) independent together with the latent variables $V, V_1, \ldots, V_n$. The computation of the loss distribution can then be accomplished thanks to the inversion formula and some Fast Fourier Transform algorithm (see Laurent & Gregory [23]). A slightly different approach, based on recursions is discussed in Andersen, Sidenius & Basu [24]. The previous approach is routinely used for portfolios of approximately one hundred names.

Let us remark that since the standard assumption states that the copula of default times is Gaussian, we are able to derive aggregate loss for different time horizons. This is of great practical importance for the computation of CDO tranche premiums, which actually involves loss distributions over different time horizons.

Conclusion

The computation of the loss distributions is of great importance for credit risk assessment and the pricing of credit risk insurance.

Standard risk measures involved in the credit field, such as the Value at Risk and the Expected Shortfall, can be easily derived from the loss distribution. At this stage, let us remark that the risk aggregation of different portfolios is not straightforward when these do no share the same factor. The Basel II framework makes the strong assumption that the same factor applies to all credit portfolios. Moreover thanks to some large portfolio approximation, the aggregate losses are comonotonic: they only depend upon the unique factor. In that simplified framework, Value at Risk and Expected Shortfall of the wholly aggregate portfolio are obtained by summation thanks to the comonotonic additive property of the previous risk measures. This departs from the Solvency II framework in insurance.

As far as the pricing of credit insurance and more precisely of CDO tranches is involved, one needs to compute stop-loss premiums $E\left[\left(L(t)-k\right)^{+}\right]$ for different time horizons t and "attachment points" k. Thanks to the semi-analytical techniques detailed above, these computations can be carried out very quickly and have now become the standard framework used by market participants.

References

[1] Tavakoli, J. (2003). Collateralized debt obligations & structured finance: new developments in cash and synthetic securitization, John Wiley & Sons.

[2] De Malherbe, E. (2006). A simple probabilistic approach to the pricing of default swap covenants, J. Risk, **8**(3), 1-29.

[3] Kealhofer, S. (2003). Quantifying credit risk I: default prediction, Financial Analysts Journal, 59(1), 30-44

[4] Gupton, G., Finger, C. & Bahia, M. (1997). CreditMetrics, technical document, J.P. Morgan

[5] Finger, C. C. (2001). The one-factor CreditMetrics model in the new Basel capital accord, RiskMetrics Journal, 2(1), 9-18

[6] Li, D. (2000). On default correlation: a copula function approach, J. of Fixed Income, **9**(4), 43-54

[7] Hull, J., Pedrescu, M. & White, A. (2006). The valuation of correlation-dependent credit derivatives using a structural model, Working Paper, University of Toronto

[8] Burtschell, X., Gregory, J. & Laurent, J-P. (2005). A comparative analysis of CDO pricing models, Working Paper, ISFA Actuarial School, Université Lyon 1 & BNP Paribas

[9] Altman, E. I., Brooks, B., Resti, A. & Sironi, A. (2004). The link between default and recovery rates: theory, empirical evidence, and implications, J. Business, **78**, 2203-2228

[10] Frye, J. (2000). Collateral damage, RISK, April, 91-94

[11] Frye, J. (2003). A false sense of security, RISK, August, 63-67

[12] Chabaane, A., Laurent, J-P. & Salomon, J. (2005). Credit risk assessment and stochastic LGDs, in "*Recovery Risk: The Next Challenge in Credit Risk Management*", E. Altman, A. Resti and A. Sironi (eds.), *Risk Publications*

[13] Wilson, T. (1997). Portfolio credit risk I, RISK, September, 111-117

[14] Wilson, T. (1997). Portfolio credit risk II, RISK, October, 56-61

[15] Gordy, M. (2000). A comparative anatomy of credit risk models, J. Banking and Finance, **24**, 119-149

[16] Crouhy, M., Galai, D. & Mark, R. (2000). A comparative analysis of current credit risk models, J. Banking and Finance, 24, 59-117

[17] Pykhtin, M. & Dev, A. (2002). Analytical approach to credit risk modelling, RISK, March, S26 - S32

[18] Frey, R. & McNeil, A. (2003). Dependent defaults in models of portfolio credit risk, J. of Risk, **6**(1), 59-92

[19] Gordy, M. (2003). A risk-factor model foundation for ratings-based bank capital rules, J. of Financial Intermediation, **12**(3), 199–232
[20] Gregory, J. & Laurent, J-P. (2004). In the core of correlation, RISK, October, 87-91
[21] Burtschell, X., Gregory, J. & Laurent, J-P. (2007). Beyond the Gaussian copula: stochastic and local correlation, Journal of Credit Risk, **3**(1), 31-62.
[22] Vasicek, O., (2002). Loan Portfolio Value, RISK, December, 160-162
[23] Martin, R., Thompson, K. & Browne, C. (2001). Taking to the Saddle, RISK, June, 91-94
[24] Laurent, J-P. & Gregory, J. (2005). Basket default swaps, CDOs and factor copulas, J. of Risk, 7(4), 103-122
[25] Andersen, L., Sidenius, J. & Basu, S. (2003). All your hedges in one basket, RISK, November, 67-72