Credit Loss Distribution and Copula in Risk Management

Xinran Tang

MSc thesis supervised by Dr. Onno Van Gaans

Mathematical Institute
Leiden University

Defense on: 31th Aug. 2015

Abstract

In this thesis we focus on the modeling of large credit losses in corporate asset portfolio. We compare loss estimates based on the classic Vasicek's approach with the assumption of normal-distributed loss distribution, and the copula approach generating heavier-tailed loss distribution. We also provide the numeric implementations of both Vasicek's and copula modeling approaches which are widely used in bank's risk management. In addition, we demonstrate how Vasicek's approach can be adopted for estimating portfolio's concentration risk charge. The last work is my own development inspired by my internship experience at the Royal Bank of Scotland. All presented results are complemented with review of the corresponding classical works in credit risk modeling.

Acknowledgements

I would like to express my sincere gratefulness to my thesis supervisor Prof. Dr. Onno van Gaans for his guidance. Without his valuable assistance, this work would not have been completed.

I would like to express my appreciation to Mathematical Institute of Leiden University where I did my Master program of Mathematics and Science-Based Business at, and the Royal Bank of Scotland where I did my Science-Based Business internship at. Without the valuable training from them, this work would not have been completed.

At last, I would like to thank my parents and friends, and would like to express my paramount thankfulness to my former RBS colleague, Kostya Shundyak, for his endless and continuous help and guidance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Financial mathematics is one of the oldest branches of mathematical science. Modern historians believe that banking itself appeared in Babylonia around 3000 BC. It was driven out of temples and palaces storing deposits of grain, cattle, and precious metals [2]. This activity had brought into life a concept of risks. For example, a clay tablet called the Code of Hammurabi (CA.1700 BCE) prescribes that a carrier of caravan should give a receipt for the consignment, take all responsibility, and exact a receipt upon delivery, or pay fivefold in case of default [2].

Despite such a long history, management of financial risks remains an unsolved problem. Recently, Laurence H Meyer [10] conducted analysis of the Asian financial crisis of 1997. He demonstrated how the risk management at many Asian financial institutions was weakened by a decade of rapid economic growth and prosperity. Many banks were extending loans without assessment of risks or even simple cash flow analysis. Rather, lending was driven by formal availability of the collateral regardless its remoteness on default, and on the basis of relationship with the borrower. As a result, volume of loans was growing faster than borrowers' ability to pay back. Also, limits on concentrations in lending to businesses haven't been respected and loans were often large relative in comparison with bank's capital. As a result, these banks felt sharply already a very beginning of economic downturn, and created a chain effect in the banking system. This example illustrates how ignoring basic risk management can contribute to economy-wide difficulties.

Thinking about such examples from a mathematical point of view, the problems seem to appear when our *apriori* expectations of future losses (or risks, in financial terminology) turn out to be significantly lower the corresponding *apos-*

teriori estimates. Development of systematic scientific methods of correcting our optimistic expectations seems to be the only way to address this problem.

This thesis is concerned with potential credit losses (i.e. credit risk) in large homogeneous asset portfolios. Credit risk refers to the risk that a borrower will default on a debt by failing to make required payments [8]. Credit risk may impact all credit-sensitive transactions, including loans, securities and derivatives. Recent historical events, like the default of the large American investment bank Lehman Brothers and the Greek sovereign crisis, popularised two particular types of credit risk:

- Concentration risk. As defined by regulators [15], this risk is associated with any single exposure or group of exposures with the potential to produce large enough losses to threaten a bank's core operations. It may arise in the form of single name concentration (e.g. loans only to Greek banks), industry concentration (e.g. loans only to oil industry) or product concentration (e.g. only mortgages).

- Country transfer risk. The risk of loss arising from a sovereign state freezing foreign currency payments (e.g. Venezuela)

- Sovereign risk. The risk of loss when a country defaults on its obligations (e.g. Argentina).

Significant resources and sophisticated programs are used by financial institutions to analyse and manage credit risk [15]. Many of them run large credit risk departments for assessing the financial conditions of their customers, and adjust their credit practices accordingly. They either use in house capabilities to advise on reducing and transferring risk, or the credit intelligence solutions provided by the big international companies like Standard & Poor's, Moody's, Fitch Ratings or their smaller local counterparts like DBRS, Dun and Bradstreet and Bureau van Dijk.

Although the variety of financial instruments is large, the core idea of effective risk management is simple: the clients should be accepted taking into account their ability to pay, and subsequently the riskier clients need to pay more to compensate for higher potential loss for the bank. There are many practical implementations of this idea:

- As stated in [8], most lenders employ their credit scoring models to rank potential and existing customers according to their risk, and then apply appropriate strategies.

- With products such as unsecured personal loans or mortgages, lenders charge a higher price for higher risk customers (see [4]).

- It is remarked in [8] that with revolving products such as credit cards and overdrafts, risk is controlled through setting of the credit limits.
- It is also mentioned in [8] that some products also require collateral, usually an asset that is pledged to secure the repayment of the loan.

Importantly, the accurate assessment is ultimately important for large risks when many customers default on their obligations at the same time. This point brings us to formulation of the main idea of this thesis.

1.2 Thesis Idea

In this thesis we focus on the modelling of large credit losses in corporate asset portfolios. We compare loss estimates based on the classic Vasicek's approach [22] with the assumption of normal-distributed loss distribution, and the copula approach [13] generating heavier-tailed loss distributions. We also provide the numeric implementations of both Vasicek's and copula modeling approaches which are widely used in bank's risk management. In addition, we will demonstrate how Vasicek's approach can be adopted for estimating portfolio's concentration risk charge. The last work is my own development inspired by my internship experience at the Royal Bank of Scotland.

1.3 Thesis Structure

This thesis consists of the following four parts:

Firstly, we focus on derivation of a price evolution of a European call or put option under some idealized assumptions using language of stochastic differential equations. This result - called the Black-Scholes equation - has a fundamental role in credit risk modelling, since it provides a structural framework of thinking about company's default as settlement of 'in-the-money' call option on the company's residual asset value.

This model was first published by Fischer Black and Myron Scholes in their 1973 paper [18]. As pointed out in [6], the key idea behind the model is to hedge the option by buying and selling the underlying asset in just the right way to eliminate risk. This type of hedging is called delta hedging and is the basis of more complicated hedging strategies such as those engaged in by investment banks and hedge funds. From the model, one can deduce the BlackScholes formula, which gives a theoretical estimate of the price of European-style options. The formula led to a boom in options trading of the Chicago Board Options Exchange and other options markets around the world. It is widely used, although often with adjustments and corrections, by options market participants. It is also observed in [7] that many empirical tests have shown that the BlackScholes price is fairly close to the observed prices, although there are well-known discrepancies such as the option smile.

Secondly, in Chapter 4 we extend our analysis and look not at a single asset, but at the large homogeneous portfolio of assets, and derive the probability distribution of portfolio's loss. Properties of this loss distribution were first described in Oldrich Vasicek 1991 paper [22], which has fundamental importance for the credit risk management industry.
Its key observation is that in a large portfolio of loans (or large portfolio of European call options) in Black-Scholes-Merton's world with correlations governed by a single economic factor, the distribution of the portfolio's loss has a closed analytic form. This results has been extensively used in deriving approximations for more complex portfolios and instruments like derivatives such as collateralized debt obligations (CDO), as well as in regulatory capital estimates and portfolio risk management.

Importantly, Vasicek's result is based on the assumption of normality of obligor's asset value. The events of the 2008 financial crisis with many obligors defaulting within a very short time interval were extremely unlikely according to the Vasicek's model, thus corresponding portfolio losses were not covered by the capital buffers. These events draw a lot of attention to the alternative models which lead to much heavier-tailed loss distributions. A useful example of alternative models is the family of multivariate normal mixture distributions, which include Student's tdistribution and the hyperbolic distribution. Rudiger Frey, and Alexander J. McNeil and Mark A. Nyfeler in 2001 paper [13] showed that the aggregate portfolio loss distribution is often very sensitive to the exact nature of the multivariate distribution of the asset values.

We are looking at these results here using a copula approach described in Chapter 6 and Chapter 5. *Copula* is a useful tool for analysis of heavy-tail distributions by allowing the modelling of the marginals and dependence structure of a multivariate probability model separately (see [12]). For example, in our case, the individual obligor's asset can be characterised by the choice or modelling of the corresponding loss marginal distribution. As all obligors are in the same market and interact with each other, this interaction can be captured via modelling the dependency structure. Paper [13] showed that it is the copula (or dependence structure) of the obligor's asset value variables that determines the higher order joint default probabilities, and thus determines the extreme risk that there are many defaults in the portfolio.

By choosing an asset value distribution from a normal mixture family, we implicitly work with alternative copulas which often differ markedly from the normal copula. Embrechts, McNeil, and Straumann [21] in 1999 showed that some of these copulas, such as the t copula, possess tail dependence and in contrast to the normal copula, they have a much greater tendency to generate simultaneous extreme values. As discussed earlier, this effect is highly important since simultaneous low asset values will lead to many joint defaults.

In 1999, David X. Li [5] made an important contribution to the field of credit risk modelling. Instead of choosing the obligor's asset value as modeling variable, he used a random variable called *time-until-default* to denote the survival time of each defaultable entity or financial instrument. He modeled the default correlation of two entities as the correlation between their survival times using standard normal dependency but each entity's survival time was characterized by an exponential marginal distribution with a special parameter called the hazard rate. In Chapter 7, we look at three methods of estimating hazard rate:

- Obtaining historical default information from rating agencies like Moody;
- Taking the implied approach using market prices of defaultable bonds or asset swap spreads;
- Taking Merton's option theoretical approach.

Finally, we provide independent numeric R implementations for the ideas described in previous Chapters. We choose three applications which are common in bank risk management. My internship at the Economic Capital Modeling team of the Royal Bank of Scotland helped me to absorb the relevant ideas and inspired me with these implementations. Applications described in Chapter 4 are relevant for estimates of portfolio's default correlations and concentration risk charges. In Chapter 5, we implement Li's 'time-until-default' copula modeling idea [5]. The implementation method is inspired by Jun Yan [24].

Chapter 2

Relevant SDE Concepts and Methods

Formalism of stochastic differential equations in financial applications often simplifies formulation of the corresponding mathematical model. Here we provide a brief overview of SDE concepts and methods relevant for our further analyses.

2.1 Wiener Processes

A *Markov process* is a type of stochastic process for which only the present value of a variable is relevant for predicting the future. The past history of the variable and the way the present has emerged from the past is irrelevant. Wiener processes, which are particularly relevant in financial mathematics, are a special case of Markov processes.

Definition 2.1.1. A stochastic process $(Z(t))_{t \geq 0}$ is a Wiener process if it has the following properties:

- The change $Z(t + \Delta t) - Z(t)$ during a small period of time $\Delta t \geq 0$ has a normal distribution with mean 0 and variance Δt:

  $$\Delta Z = Z(t + \Delta t) - Z(t) = \epsilon\sqrt{\Delta t}, \tag{2.1}$$

  where ϵ has a standardized normal distribution $N(0, 1)$.

- The random variable $Z(t+\Delta t)-Z(t)$ and $Z(s+\Delta s)-Z(s)$ are independent, provided $0 \leq t \leq t + \Delta t \leq s \leq s + \Delta s$.

- $Z(0) = 0$.

- $t \to Z(t)$ is continuous almost surely.

A *Wiener process* is a type of Markov stochastic process with a mean change of zero and a variance rate of 1 per unit of time. In physics, it is often referred as *Brownian motion* as it has been used to describe the motion of a particle subject to a large number of small molecular shocks.

2.2 Ito Process

An Ito process is composed of a drift part and an Ito integral. Definition of an Ito process requires defining *filtration*, *adapted process* and a class of function which can be the integrand of Ito integral. see [14] for more information.

Definition 2.2.1. Given a measurable space $(\Omega, \mathcal{F})$, a *filtration* is a family of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$ with $\mathcal{F}_t \subseteq \mathcal{F}$ for each t such that when $s \leq t$, we have $\mathcal{F}_s \subseteq \mathcal{F}_t$.

Definition 2.2.2. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be an increasing family of σ-algebras of subset of Ω. A process $g : [0, \infty) \times \Omega \to \mathbf{R}^n$ is called $\mathcal{F}_t$*-adapted* if for each $t \geq 0$ the function

$$\omega \to g(t, \omega)$$

is $\mathcal{F}_t$-measurable.

Using these definitions, we come to the definition of Ito process:

Definition 2.2.3. (1-dimensional Ito processes [14]) Let B_t be 1-dimensional Brownian motion on probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_{t\geq 0}$ be a filtration such that B is $\mathcal{F}$-adapted and $B(t) - B(s)$ is independent of $\mathcal{F}_s$ for all $t \geq s \geq 0$. An *Ito process* is a stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ of the form

$$X_t = X_0 + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB_s, \tag{2.2}$$

where μ and σ are $\mathcal{F}$-adapted and

$$P[\int_0^t \sigma(s)^2 ds < \infty, \forall t \geq 0] = 1$$

and

$$P[\int_0^t |\mu(s)|ds < \infty, \forall t \geq 0] = 1.$$

If X_t is an Ito process of the form (2.2), it is sometimes written in the shorter differential form

$$dX_t = \mu(t)dt + \sigma(t)dB_t \tag{2.3}$$

and $(t, w) \to \mu(t, \omega)$ is then called the expected drift rate, and $(t, w) \to \sigma(t, \omega)$ is called the volatility.

2.3 Asset Price Process

A standard example of an Ito process in financial mathematics is a model of the price of a non-dividend-paying stock. Here we demonstrate the process of mapping properties of financial product into the parameters of Ito formula:

- Stochastic variable X should be identified not with the stock price A, but with its instantaneous return dA/A. The first model would imply that the expected stock return does not depend on the stock's price – investors would require a 12% per annum expected return when the stock price is \$10, as well as when it is \$30, which is not realistic.

- Assuming zero uncertainty and a constant expected drift rate μ and integrating between time 0 and time T, we get

$$X_T = X_0 e^{\mu T}, \tag{2.4}$$

  where X_0 and X_T are the stock price at time 0 and T, respectively. Thus, μ can be identified with a continuously compounded growth rate of the stock price per unit of time

- In [3] it is remarked that it is reasonable to assume that the variability of the instantaneous return in a short period of time does not depend on the stock price. This suggests that the standard deviation of the change in a short period of time should be proportional to the stock price.

Thus, we arrive at the most widely used model of stock price behaviour:

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \tag{2.5}$$

with μ being the stock's expected rate of return, and σ being volatility of the stock price.

2.4 Ito Formula

Ito's formula is widely employed in mathematical finance, and its best known application is in the derivation of the Black-Scholes equation for option values.

2.4.1 The 1-dimensional Ito formula

Theorem 2.4.1. ***(1-dimensional Ito formula [14])*** *Assume X_t is a Ito drift-diffusion process that satisfies the stochastic differential equation*

$$dX_t = \mu dt + \sigma dB_t,$$

where B_t is a Brownian motion. Let $f \in C^2([0,\infty)\times\mathbf{R})$ be a twice-differentiable scalar function, then we have

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right)dt + \sigma\frac{\partial f}{\partial x}dB_t, \tag{2.6}$$

where the derivatives of f are evaluated at (t, X_t).

It is obvious that $f(t, X_t)$ is an Ito process with a drift rate of

$$\left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right)(t, X_t),$$

and a volatility of $\sigma\frac{\partial f}{\partial x}(t, X_t)$.

Example 2.4.2. (Source: B. Øksendal [14]) What is the value of

$$\int_0^t sdB_s?$$

From classical calculus it seems that the term tB_t should appear, so we let $f(t,x) = tx$, then $f(t, B_t) = tB_t$. Then by Ito formula,

$$df(t, B_t) = d(tB_t) = B_t dt + tdB_t,$$

so

$$tB_t = \int_0^t B_s ds + \int_0^t sdB_s,$$

so

$$\int_0^t sdB_s = tB_t - \int_0^t B_s ds.$$

Example 2.4.3. Now consider the Stochastic Differential Equation $dX_t = \mu X_t dt + \sigma X_t B_t$ with $X_0 = x_0 \in \mathbf{R}$. We let $f(t,x) = x_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma x}$, therefore

$$X_t = f(t, B_t) = x_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B_t},$$

then by Ito formula,

$$\begin{aligned} dX_t = df(t, B_t) &= \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial x}\, dB_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\,(dB_t)^2 \\ &= (\mu - \frac{1}{2}\sigma^2)X_t dt + \sigma X_t dB_t + \frac{1}{2}\sigma^2 X_t dt \\ &= \mu X_t dt + \sigma X_t dB_t. \end{aligned}$$

So all conditions in the definition of solution (using Ito formula) are satisfied. Hence the complete solution is

$$X_t = x_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t}.$$

2.4.2 The multi-dimensional Ito formula

Translating the previous arguments into the higher dimensions, we arrive at the following

Definition 2.4.4. [14] Let $B(t) = (B_1(t), \ldots, B_m(t))$ denote m-dimensional Brownian motion. If each of the processes $\mu_i(t)$ and $\sigma_{ij}(t)$ satisfies the condition given in Definition 2.2.3 ($1 \le i \le n$,$1 \le j \le m$) then we can form the following n Ito processes

$$\begin{cases} dX_1(t) = \mu_1 dt + \sigma_{11} dB_1(t) + \cdots + \sigma_{1m} dB_m(t) \\ dX_2(t) = \mu_2 dt + \sigma_{21} dB_1(t) + \cdots + \sigma_{2m} dB_m(t) \\ \vdots \\ dX_n(t) = \mu_n dt + \sigma_{n1} dB_1(t) + \cdots + \sigma_{nm} dB_m(t) \end{cases},$$

Or, in matrix notation

$$dX(t) = \mu dt + \sigma dB(t),$$

where

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nm} \end{pmatrix}, dB(t) = \begin{pmatrix} dB_1(t) \\ \vdots \\ dB_m(t) \end{pmatrix}. \tag{2.7}$$

Such a process $X(t)$ is called an n-dimensional Ito process, we use notation $X(t)$ instead of X_t to indicate the difference between n-dimensional Ito process and 1-dimensional Ito process.

It is natural to ask what is the Ito formula for n-dimensional Ito process. An answer is provided by the following:

Theorem 2.4.5. ***(General Ito formula [14])*** *Let*

$$dX(t) = \mu dt + \sigma dB(t)$$

be an n-dimensional Ito process as above. Let $(t, x) \to f(t, x) = (f_1(t, x), \ldots, f_p(t, x))$ be a C^2 map from $[0, \infty) \times \mathbf{R}^n$ into $\mathbf{R}^p$. Then the process

$$Y(t, \omega) = f(t, X(t))$$

is again an Ito process, whose n_{th} component, Y_k, is given by

$$dY_k(t) = \frac{\partial f_k}{\partial t}(t, X_t)dt + \sum_i \frac{\partial f_k}{\partial x_i}(t, X)dX_i(t) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f_k}{\partial x_i \partial x_j}(t, X_t)dX_i(t)dX_j(t)$$

where $dX_i(t)$ is expanded as in Definition 2.4.4 with the convention that $dB_i dB_j = \delta_{ij} dt$, $dB_i dt = dt dB_i = 0$.

Chapter 3

The Black-Scholes Model

3.1 The Black-Scholes world

Ideas of Ito calculus were formulated in the 1950s, however their application in finance required a set of additional modelling which were made approximately 30 years later in works of the economists R. Merton, F. Black and M. Scholes [18].

F. Black and M. Scholes tried to formulate a minimal description of the 'fair' world, and asked a question whether in this world the stochasticity (or risk) of the value of a financial instrument can be completely eliminated by holding in a portfolio a small number of other financial instruments. The answer to this question turned out to be positive, and this led to a significant advance in financial mathematics.

Black and Scholes set the following assumptions for their idealised world:

- a riskless profit cannot be made on the market, i.e. two different types of financial instruments which provide the same payoff for the investors should have the same price (in financial jargon, there is no arbitrage opportunity)
- Any amount, even fractional, of cash can be borrowed and lent at the market at the riskless rate r
- As well as any amount, even fractional, of any financial instrument can be bought and sold on the market. Note that this includes so called 'short selling': you can sell a financial instrument which is not in your portfolio at this moment, and purchase it only at the time of its delivery
- The financial transactions on the market are costless (in financial jargon, the market is frictionless).

To study the properties of this 'fair' world, Black and Scholes assumed that a company's stock has the following properties:

- The instantaneous log returns of the stock price S is a geometric Brownian motion with constant drift rate μ and volatility σ, i.e.

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t, \tag{3.1}$$

  where W_t is a Brownian motion, and t is time.

- The stock does not pay dividends. Even though this sounds not realistic, there are companies in the real world like Google which never paid dividends so far.

The riskless rate of borrowing money form the market is realized in Black-Scholes 'fair' world by a bond for any required lifetime T of the transaction which pays a constant rate r per unit of time.

Finally, a second stochastic financial instrument in Black-Scholes's world is called a call option on a stock. This is a financial contract which gives a buyer a right but not an obligation to buy a stock at fixed price K at expiry time $t = T$ regardless of the stock price $S(T)$ at that time point. We view the price of the option as a function of the stock price and of time, denoting the price of this call option as $C(S,t)$ and its payoff $V(S,T)$ at expiry. Note that since the price of the stock is stochastic, the price of the call option on the stock is stochastic as well. However, Black and Scholes observed that the nature of this stochasticity is identical for both of them and the right combination of those two instruments should eliminate this stochasticity completely.

3.2 The Black-Scholes equation

John C. Hull in [3] gave a derivation of The Black-Scholes equation. Consider a portfolio consisting of one call option on a stock and a certain fraction α of the stock itself. In order to find the fraction α of the stock which eliminates the stochasticity of the option's price, we need to know how the payoff price of the option at expiry V changes as a function of S and t over an infinitesimaly small time increment dt. By Ito's lemma and (3.1), we have:

$$dV_t = \left(\mu S_t \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S_t \frac{\partial V}{\partial S}\, dW_t. \tag{3.2}$$

Thus, if we allow α to vary in time, we can choose $\alpha = \frac{\partial V}{\partial S}$, which makes the price Π of the portfolio equal to

$$\Pi = -V + \frac{\partial V}{\partial S} S. \tag{3.3}$$

Over the small time interval $[t, t+\Delta t]$, the change of the value of the portfolio is approximately:

$$\Delta \Pi = -\Delta V + \frac{\partial V}{\partial S} \Delta S. \tag{3.4}$$

We will assume that this is a self-financing portfolio, i.e. the infinitesimal change in its value is only due to the infinitesimal changes in the values of its assets, and not due to changes in the positions in the assets.

Discretizing the equations for dS/S and dV,

$$\Delta S = \mu S \, \Delta t + \sigma S \, \Delta W, \tag{3.5}$$

$$\Delta V = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \Delta t + \sigma S \frac{\partial V}{\partial S} \Delta W, \tag{3.6}$$

we get the following expression for $\Delta \Pi$:

$$\Delta \Pi = \left(-\frac{\partial V}{\partial t} - \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \Delta t. \tag{3.7}$$

Since the volatility part has been completely eliminated now, the rate of return for this portfolio must be equal to the rate of return on a bond, i.e.

$$r \Pi \, \Delta t = \Delta \Pi. \tag{3.8}$$

Otherwise, the assumption of no arbitrage would be violated.

Finally, equating the two formulas for $\Delta \Pi$ we obtain:

$$\left(-\frac{\partial V}{\partial t} - \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \Delta t = r \left(-V + S \frac{\partial V}{\partial S} \right) \Delta t, \tag{3.9}$$

and simplifying, we arrive at the celebrated Black - Scholes partial differential equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + r S \frac{\partial V}{\partial S} - r V = 0. \tag{3.10}$$

In [3] it is remarked that under the assumptions of the Black - Scholes world, this equation holds for any type of option if its price $V(S, t)$ is twice differentiable with respect to S and once with respect to t.

3.3 The Black-Scholes formula

John C. Hull in [3] provided a derivation of the The Black-Scholes formula. For the call option on a stock, the Black-Scholes PDE has the following boundary conditions:

$$C(0,t) = 0 \text{ for all } t \tag{3.11}$$
$$\lim_{S\to\infty} \frac{C(S,t)}{S} = 1 \text{ for all } t \tag{3.12}$$
$$C(S,T) = \max\{S-K,0\} \tag{3.13}$$

The Black Scholes PDE can be transformed into a standard diffusion equation

$$\frac{\partial u}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} \tag{3.14}$$

by using the following transformations:

$$\tau = T - t, \tag{3.15}$$
$$u = C(S,t)e^{r\tau}, \tag{3.16}$$
$$x = \ln\left(\tfrac{S}{K}\right) + \left(r - \tfrac{1}{.}2\sigma^2\right)\tau \tag{3.17}$$

The terminal condition $C(S,T) = \max\{S-K,0\}$ transforms into an initial condition,

$$u(x,0) = u_0(x) = K(e^{\max\{x,0\}} - 1), \quad x \in \mathbf{R}. \tag{3.18}$$

Using the standard textbook method for solving a diffusion equation, we have

$$u(x,\tau) = \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} u_0(y) \exp\left[-\frac{(x-y)^2}{2\sigma^2\tau}\right] dy, \tag{3.19}$$

which, after some manipulations, yields

$$u(x,\tau) = Ke^{x+\frac{1}{2}\sigma^2\tau}N(d_1) - KN(d_2), \tag{3.20}$$

where

$$d_1 = \tfrac{1}{\sigma\sqrt{\tau}}\left[\left(x + \tfrac{1}{2}\sigma^2\tau\right) + \tfrac{1}{2}\sigma^2\tau\right], \tag{3.21}$$
$$d_2 = \tfrac{1}{\sigma\sqrt{\tau}}\left[\left(x + \tfrac{1}{2}\sigma^2\tau\right) - \tfrac{1}{2}\sigma^2\tau\right], \tag{3.22}$$

and $N(x)$ is the standard normal cumulative distribution function.

Reverting u, x, τ to the original set of variables, yields the final expression for the solution of the Black-Scholes equation,

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)}, \tag{3.23}$$

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right], \tag{3.24}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}. \tag{3.25}$$

Chapter 4

Portfolio Loss Distribution

A main idea behind the Black-Scholes equation was a perfect hedging of risk for an option by a fraction of the underlying stock in a portfolio. However, it was very well known that in large portfolios of even identical financial instruments, there are risks, and those risks tend to materialise at different time points (see [23]). This observation inspired another fundamental results which gives the likelihood of losses in a large portfolio of options.

Note that the binary nature of an option at expiry can be used in a context of corporate finance to model an indicator of default. If the stock price is above zero, the company is performing, whereas if it becomes equal to zero, the investors choose to liquidate the company and declare its default. This analogy was made rigorous in work of R. Merton. However, we will not elaborate on his results here, and simply assume that an obligor i defaults if the value of its assets A_i at time $t = T$ falls below the contractual value B of its obligations. As in Merton's model (2.5), the value of its assets $A_i(t)$ will be described by the process

$$dA_i(t) = \mu A_i(t)dt + \sigma A_i(t)dW_t,$$

where W_t is a Brownian motion and μ and σ are positive constants. As $A_i(0) > 0$, $A_i(T)$ can be formally solved as (see example 2.4.3 in Chapter 2)

$$\log A_i(T) = \log A_i(0) + \mu T - \frac{1}{2}\sigma^2 T + \sigma\sqrt{T}W, \tag{4.1}$$

with X_i now being a standard normal variable. The probability of default of obligor i is equal to

$$p_i = \mathbb{P}[A_i(T) < B_i] = \mathbb{P}[X_i < c_i] = N(c_i)$$

where

$$c_i = \frac{\log B_i - \log A_i - \mu T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$

and $N(x)$ is the cumulative normal distribution function.

Now consider a portfolio consisting of n identical loans with the same term T and identical obligors. Let the probability of default on each loan be $p_i = p$, and assume the same correlation ρ between asset values of any two obligors. For the sake of simplicity, let's assume that the gross loss L_i on the i-th loan is $L_i = 1$ if the i-th borrower defaults and $L_i = 0$ otherwise. It is convenient to denote by L the portfolio fraction gross loss,

$$L = \frac{1}{n}\sum_{i=1}^{n} L_i$$

For independent defaults and in a limit of $n \to \infty$, the distribution of L would converge, by the central limit theorem, to a normal distribution. In a more realistic model, the defaults are not independent. The conditions of the central limit theorem are not satisfied and L is not asymptotically normal. It turns out, however, that the closed form analytical expression for the portfolio probability loss distribution $P(L)$ can be derived.

4.1 The Limiting Distribution of Portfolio Losses

Vasicek in [22] assumes that the pairwise correlations ρ between obligors log returns in one time unit can be represented by splitting the variables X_i in Equation (4.1) as

$$X_i = Y\sqrt{\rho} + Z_i\sqrt{1-\rho} \tag{4.2}$$

where Y and Z_i are mutually independent standard normal variables, and $\rho \in [0, 1]$ is a constant, $i = 1, \cdots, n$. The variable Y can be interpreted as a portfolio common economic factor over the interval $(0, T)$, whereas $Z_i\sqrt{(1-\rho)}$ characterises the company's specific risk. We also call such a model a 'factor model' (see [22]).

Since $p = \mathbb{P}[X_i < c_i] = N(c_i)$, we have $c_i = N^{-1}(p)$. For the fixed common factor Y, the conditional probability of loss on any one loan is

$$p(Y) = \mathbb{P}[L_i = 1|Y] = N\left(\frac{N^{-1}(p) - Y\sqrt{\rho}}{\sqrt{1-\rho}}\right). \tag{4.3}$$

Conditional on the value of Y, the variables L_i are independent equally distributed Bernoulli variables, hence with a finite variance. The portfolio loss conditional on Y converges, by the law of large numbers, to its expectation

$p(Y)$ as $n \to \infty$. Then in a limit as $n \to \infty$ and hence approximately for large portfolios,

$$\begin{aligned}
\mathbb{P}[L \le x] &= \mathbb{P}[p(Y) \le x], \\
&= \mathbb{P}\left[N\left(\frac{N^{-1}(p) - Y\sqrt{\rho}}{\sqrt{1-\rho}}\right) \le x\right], \\
&= \mathbb{P}\left[-Y \le \frac{\sqrt{1-\rho}N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right], \\
&= N\left(\frac{\sqrt{1-\rho}N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right).
\end{aligned}$$

Thus, the cumulative distribution function of loan losses in a large portfolio's limit is

$$\mathbb{P}[L \le x] = N\left(\frac{\sqrt{1-\rho}N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right). \tag{4.4}$$

Note that the assumption of equal obligor's weights in the portfolio is not critical. Let the portfolio weights be $w_1, w_2, \cdots, w_n$ with $\sum_{i=1}^n w_i = 1$. It has been shown by Vasicek in [22] that the portfolio loss $L = \sum_{i=1}^n w_i L_i$ conditional on Y converges to its expectation $p(Y)$ whenever (and this is a necessary and sufficient condition) $\sum_{i=1}^n w_i^2 \to 0$ and the portfolio loss distribution converges to the form 4.4. In other words, if the portfolio contains a sufficiently large number of loans without it being dominated by a few loans much larger than the rest, the limiting distribution provides a good approximation for the portfolio loss.

4.2 Properties of the Loss Distribution

As stated in (4.4), the portfolio loss distribution is given by the cumulative distribution function

$$F(x; p, \rho) = N\left(\frac{\sqrt{1-\rho}N^{-1}(x) - N^{-1}(p)}{\sqrt{\rho}}\right). \tag{4.5}$$

In this section we discuss some more properties of this distribution following [22], we find that

- When $\rho \to 1$:
  We have $\mathbb{P}(L \le x) = 1 - p = \mathbb{P}(L = 0)$ for all $x \in (0, 1)$, and
  $\mathbb{P}(L = 1) = p$.
  Thus all loans default with probability p.

- When $\rho \to 0$:
  We have $\mathbb{P}(L \le x) \to 0$ for $x < p$, and

$\mathbb{P}(L \leq x) = 1$ for $x \geq p$.
This yields $\mathbb{P}(L = p) = 1$.

The corresponding probability density can be derived by calculating the derivative of $F(x; p, \rho)$ with respect to x, which is

$$\begin{aligned} f(x; p, \rho) &= \frac{\partial F(x; p, \rho)}{\partial x}, \\ &= \sqrt{\frac{1-\rho}{\rho}} \times \exp \\ &\quad \left(-\frac{(1-2\rho)(N^{-1}(x))^2 - 2\sqrt{1-\rho}N^{-1}(x)N^{-1}(p) + (N^{-1}(p))^2}{2\rho} \right), \\ &= \sqrt{\frac{1-\rho}{\rho}} \exp \left(\frac{1}{2}(N^{-1}(x))^2 - \frac{1}{2\rho}(N^{-1}(p) - \sqrt{1-\rho}N^{-1}(x))^2 \right). \end{aligned}$$

Proposition 4.2.1. *[22] For any given level of confidence α, the α-quantile $q_\alpha(L)$ of a random variable L with distirbution function $F(x; p, \rho)$ is given by*

$$q_\alpha(L) = p(-q_\alpha(Y)) = N(\frac{N^{-1}(p) + \sqrt{\rho}q_\alpha(Y)}{\sqrt{1-\rho}})$$

where Y has distribution $N(0, 1)$ and $q_\alpha(Y)$ denotes the α-quantile of the standard normal distribution.

Proposition 4.2.2. *[22] The expectation and the variance of a random variable L with distribution function $F(x; p, \rho)$ are given by*

$$\mathbb{E}(L) = p$$

and

$$\mathbb{V}(L) = N_2(N^{-1}(p), N^{-1}(p); \rho) - p^2,$$

where $N_2(\cdot, \cdot; \rho)$ denotes the cumulative bivariate normal distribution function with correlation ρ.

4.3 Use of Vasicek's distribution

In this section, we give two examples and the corresponding R programming implementations of the practical use of Vasicek's portfolio loss distribution function, i.e. Equation (4.4). Ideas of these examples were inspired by my internship at the Economic Capital Modeling team of the Royal Bank of Scotland.

	V1	V2	V3	V4	V5	V6
1	-0.46016284	-2.388099357	-1.672394161	-1.568652210	-0.7198926288	-1.34006611
2	-1.17559669	-0.458309968	-0.539554131	-0.292844823	-1.4224019654	-0.52797879
3	1.07050446	-0.402853357	0.348621166	0.672726412	0.0029852489	1.05555638
4	0.83023690	0.874382137	0.633159582	0.457235426	0.7073049354	0.54558034
5	1.89580825	0.903555845	2.049568040	1.904875655	1.6201421133	1.57604889
6	-0.94184744	-0.465969830	-1.038212702	-0.132383782	-2.2616306867	-1.54461518
7	-0.10829635	-0.437857259	-0.956303341	-0.554764242	-0.4084855297	-0.77370502
8	-1.54211042	-0.419208176	-1.511868723	-1.209224883	-1.7658007901	-0.70611556
9	-1.26333648	-0.450296731	-0.167246674	-1.638648376	-0.3944178520	-1.06318266
10	-0.82537053	-0.905390843	-1.890200166	0.446197446	-0.1277498390	-0.97355569
11	0.23649862	0.363431008	0.838822644	-0.643690742	-0.3210261417	-0.22971728
12	0.55000756	0.509799255	-0.061107957	0.644344540	1.1289425378	0.41786844
13	-1.68249568	-0.299426133	-1.173373567	-1.515719794	-1.3917803872	-1.17980513
14	-1.91001377	-1.851116599	-1.809398685	-1.573691875	-1.7416143921	-1.59054932
15	0.60633221	1.284962830	0.102024757	0.368731026	-0.7502315884	0.72906181

Figure 4.1: Obligors' standardized asset in 15 time points

4.3.1 Correlation estimation

The value of the pairwise correlations ρ are not known in practical situations, and it needs to be determined from empirical time series of portfolios default rates. These correlations ρ are usually also used in the required level of the bank's economic capital to ensure its solvency with specified confidence at a future time point. Since we have no real historic data, we will first generate a data set.

The algorithm is as follows:

1. Randomly generate 2000 obligors with assets satisfying normal distribution with mean 0 and correlation 0.7, in 100 time points. Figure 4.1 shows part of the obligors' asset in 15 time points.

2. Set corresponding 2000*100 standard liability threshold matrix, with threshold $N^{-1}(p)$, where p is the default probability. In our case, $p = 0.1$, therefore the threshold equals to -1.28.

3. Compare asset and liability threshold matrix, finding 'default or not' matrix. (see Figure 4.2)

4. For each time point (i.e. each realization of Y), calculate the default rate, which equals to the fraction (the number of default)/(total number of obligors), i.e., $p(Y)$ in Equation (4.3)

5. After getting $p(Y)$, we can estimate the correlation ρ based on a small derivation:

	V1	V2	V3	V4	V5	V6
1	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE
2	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
3	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
4	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
5	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
6	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
7	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
8	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
9	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE
10	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
11	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
12	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
13	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE
14	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
15	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

Figure 4.2: Obligor's default-or-not matrix in 15 time points

(a) We know from Equation (4.3) that:

$$p(Y) = \mathbb{P}[L_i = 1|Y] = N\left(\frac{N^{-1}(p) - Y\sqrt{\rho}}{\sqrt{1-\rho}}\right).$$

(b) Next, we introduce a variable: Distance-to-Default(DD).

$$DD = N^{-1}(p(Y)) = \frac{N^{-1}(p) - Y\sqrt{\rho}}{\sqrt{1-\rho}}.$$

(c) Then

$$\mathrm{var}(DD) = \frac{\rho}{1-\rho}\mathrm{var}(Y) = \frac{\rho}{1-\rho}$$

due to the assumption of $\mathrm{var}(Y) = 1$.

(d) Thus, the estimated value of ρ is

$$\rho = \frac{\mathrm{var}(DD)}{1+\mathrm{var}(DD)}.$$

The whole R code is as follows:

```
library(MASS)

#number of time points
T <- 100

#number of obligors
Nobl <- 2000

#average value of assets is 0
mu <- rep(0,Nobl)

#correlation rho = 0.7
rho = 0.7

#correlation matrix
Sigma <- matrix(rho, nrow=Nobl, ncol=Nobl) + diag(Nobl)*(1-rho)

#values of assets
assets <- mvrnorm(n=T, mu=mu, Sigma=Sigma)

#portfolio default rate
p <- 0.1

#thus the corresponding standard liability threshold
Lthr <- qnorm(p)

#and all liab tresholds in matrix form
liab <- matrix(rep(Lthr, T*Nobl), nrow=T,ncol=Nobl)

#observed defaults
defaults <- assets < liab

#portfolio default rate at each time point, i.e. left side of Eq
    .(4.3)
portDefaultRate <- rowSums(defaults)/Nobl

#time average portfolio default rate, should be close to p
avDefaultRate <- mean(portDefaultRate)

#so called distance to default series (excluding points of no
    default), i.e. it is N^{-1}(p(Y)), (p(Y)) is in Eq.(4.3)
distanceToDefault <- qnorm(portDefaultRate[portDefaultRate > 0])

#variance of distance to default
varAvDefaultRate <- var(distanceToDefault)

#Vasicek's estimate of rho, should be close to rho
rhoVasicek <- varAvDefaultRate / (1 + varAvDefaultRate)
```

corrVasicek.R

4.3.2 Concentration risk charge

The implementation of the example of a concentration risk charge discussed below is my original work, inspired by discussions with my former RBS colleagues.

The *Concentration Risk Charge* covers the risk of losses which a regulated institution acquires via excessive exposures to a particular asset, counterparty or group of related counterparties. For example, people were very optimistic about the economy prior to the 2008 financial crisis, and were simply borrowing tomorrow's money on today. Many banks acquired large exposures on mortgages, focusing on profit and ignoring increasing credit risk. Therefore, during the crisis, when defaults on mortgages were high, these banks easily got bankrupt due to the large exposure on mortgage business.

Regulators recognize this problem, however they assume that it cannot be solved within a standardized regulatory approach. The regulatory capital requirements are calculated using a Vasicek-like formula with predefined parameters which (mostly) don't take into account differences between different sectors of the economy. Thus, a portfolio with all exposures in a single sector attracts the same amount of regulatory capital as a portfolio with exposures spread over multiple sectors. Therefore, regulators require from banks to develop their internal methodology of calculating their concentration risk charge and report the number to the central bank. Finding the right exposure of each business sector in a bank's portfolio, is the key to calculate right concentration risk charge.

The algorithm is as follows:

1. Estimate the value of the portfolio's quantile risk measure for the current portfolio composition.

2. Choose a risk diversification measure. A common choice is a variance of portfolio's loss since it is typically used for portfolio risk management.

3. Estimate the value of the chosen risk diversification measure for the marginal (sector) loss distribution for the current portfolio composition. (see Figure 4.3)

4. Choose risk diversification strategy. For example, a diversification strategy might consist of shifting a fraction of the portfolio proportional to sector's loss variance from a sector with high variance to a sector with low variance.

5. Repeat execution of risk diversification strategy until a stationary state or a predefined business constraint is reached.

6. Estimate the value of the portfolio's quantile risk measure for the diversified portfolio. (see Figure 4.4)

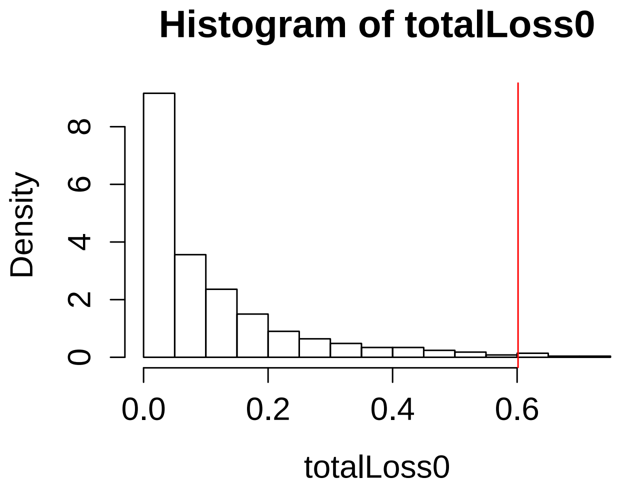

Figure 4.3: The initial portfolio losses with 99% quantile line

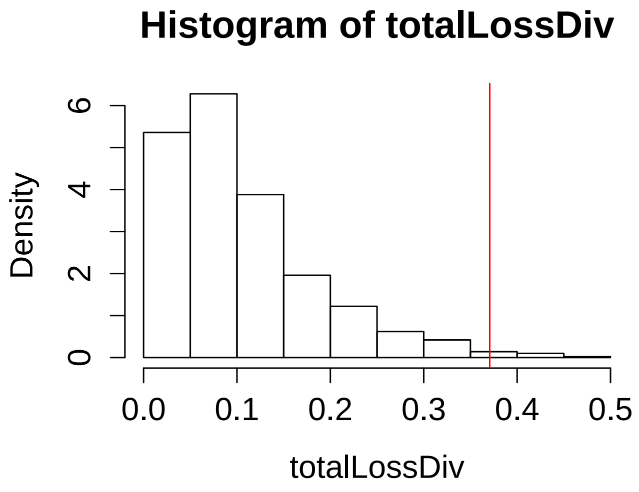

Figure 4.4: The diversified portfolio losses with 99% quantile line

7. The difference between the quantile's risk measure for the current and diversified portfolios constitutes a concentration risk charge. We can see from Figure 4.3 and Figure 4.4 that under 99% quantile's measure, the concentration risk charge is 0.6-0.37=0.23.

The following code provides a method to calculate concentration risk charge:

We Assume the portfolio consists of two different sectors: mortgage and credit-card.

```
require(lattice)
library(mvtnorm)

#probability of default of obligor
PD <- 0.1

#bank's respective exposure of mortgage portfolio and creditcard
    portfolio
exposure <- c(0.9,0.1)

#corresponding obligor correlation within mortgage portfolio and
    creditcard portfolio
rho1 <- c(0.4,0.1)

#difine a function caluculating default probability, i.e. Equation
    (4.3)
vasCondPD <- function(PD,rho,Y){
  return ( pnorm( (qnorm(PD) - sqrt(rho) * Y) / sqrt(1 - rho) ))
}

#number of scenarios
nScenarios <- 1000

#correlation between mortgage portfolio and creditcard porfolio
rho <- 0.6

#generating 1000 scenarios standardized normal-distributed mortgage
     portfolio and creditcard porfolio with correlation
mat <- rmvnorm(nScenarios, mean=c(0,0), sigma=(1-rho)*diag(2)+rho)
mat <-as.matrix(mat)


#calculating initial losses for 2 portfolios respectively, loss=
    exposure*default probability
lossMort0 <- exposure[1] * vasCondPD(PD,rho1[1],mat[,1])
lossCards0 <- exposure[2] * vasCondPD(PD,rho1[2],mat[,2])

#total initial losses
```

```
totalLoss0 <- lossMort0 + lossCards0

#calculating 99% quantile of the initial total losses
q99 <- quantile(totalLoss0,probs=0.99)

#histogram of the initial total losses with 99% quantile line
hist(totalLoss0, freq=FALSE)
abline(v=q99)

#concentration adjustment processes:
#1. set the initial portfolio losses for mortgage and creditcard
    respectively
lossMort <-lossMort0
lossCards <- lossCards0

#2. everytime adjust the exposure with the change proportional on
    the variance of the losses from previous round until it reaches
     the balance
for(t in 1:1000){
  lossMort <- exposure[1] * vasCondPD(PD,rho1[1],mat[,1])
  lossCards <- exposure[2] * vasCondPD(PD,rho1[2],mat[,2])

  dExpMort <- var(lossMort)*exposure[1]
  dExpCards <- var(lossCards)*exposure[2]

  exposure[1] <- exposure[1] - dExpMort + dExpCards
  exposure[2] <- exposure[2] + dExpMort - dExpCards
}

#total losses after diversification
totalLossDiv <- lossMort + lossCards

#calculating 99% quantile of the diversified total losses
q99Div <- quantile(totalLossDiv,probs=0.99)

#histogram of the diversified total losses with 99% quantile line
hist(totalLossDiv, freq=FALSE)
abline(v=q99Div)

#the difference of the quantile between the initial portfolio and
    diversified portfolio, i.e. the so-called concentration charge
concCharge = q99 - q99Div
```

vasicek.R

Chapter 5

Copula

5.1 Mathematical Definition of Copula

As stated in [13], a copula function is simply a multivariate joint distribution function of random vectors with standard uniform marginal distributions. As mentioned in Chapter 6, a copula gives a way of putting marginal distributions of several individual obligor's asset returns, or survival time in Li's case, together to form a joint distribution of groups of risks. We can simply say that a copula is a tool which enables us to model the marginal distribution, as well as the dependency structure of a vector of latent variables separately.

For example, the obligor's asset return can be described by modeling the marginals. Also, as all obligors are in the same market, each obligor's action and financial status has an interaction effect with other obligors. This interaction effect can be described by modeling the dependency structure.

Definition 5.1.1. [12] Given a random vector $(X_1, X_2, \ldots, X_d)$ with continuous marginal distributions $F_i(x) = \mathbb{P}(X_i \leq x), x \in \mathbf{R}, i = 1, \ldots, d$, a copula function is a multivariate distribution function such that its marginal distributions are standard uniform. A common notation for a copula is:

$$C(u_1, u_2, \ldots, u_d) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_d \leq u_d), \tag{5.1}$$

where $(U_1, U_2, \ldots, U_d) = (F_1(X_1), F_2(X_2), \ldots, F_d(X_d))$.

The marginal diatribution F_i contains all the information about each variable X_i, whereas the copula C contains all the information about the dependency structure.

If for each i the inverse of F_i exists, equation (5.1) can also be written as follows:

$$C(u_1, u_2, \ldots, u_d) = \mathbb{P}(X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2), \ldots, X_d \leq F_d^{-1}(u_d)), \tag{5.2}$$

where F_i^{-1} is the inverse of F_i.

Example 5.1.2. [13] If the vector of latent variables $\mathbf{X}$ has a multivariate Gaussian distribution with correlation matrix R, then the copula of $\mathbf{X}$ may be represented by

$$C_R^{Ga}(u_1, \ldots, u_m) = \mathbf{N}_R(N^{-1}(u_1), \ldots, N^{-1}(u_m)), \tag{5.3}$$

where $\mathbf{N}_R$ denotes the joint distribution function of a centered m-dimensional normal random vector with correlation matrix R, and N is the distribution function of univariate standard normal. C_R^{Ga} is known as the *Gaussian copula* with the correlation matrix R.

Example 5.1.3. [23] The *independence copula* is defined by

$$C(u_1, \ldots, u_d) = u_1 \times \cdots \times u_d.$$

5.1.1 Sklar's theorem

A copula is powerful because of Sklar's theorem, which enables the seperation of modeling marginal distributions and dependency structure.

Theorem 5.1.4. ***(Sklar [19], see also [13])*** *Let $\mathbf{F}$ be a multivariate d-dimensional distribution function with marginals $F_1, \ldots, F_d$. Then there exists a copula C such that*

$$\mathbf{F}(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \quad (x_1, \ldots, x_d \in \mathbf{R}).$$

Moreover, if the marginal distributions $F_1, \ldots, F_d$ are continuous, then C is unique.

The converse is also true:

Proposition 5.1.5. *[13] For any copula C and marginal distribution functions $F_1, \ldots, F_d$, the function*

$$\mathbf{F}(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \quad (x_1, \ldots, x_d \in \mathbf{R})$$

defines a multivariate distribution function with marginals $F_1, \ldots, F_d$.

Summarizing Theorem 5.1.4 and Proposition 5.1.5, one can say that every multivariate distribution with continuous marginals admits a unique copula representation. Also copulas and distribution functions are the building blocks to derive new multivariate distributions with prescribed correlation structure and marginal distributions.

5.2 Copula for Fatter-tail Loss Distribution

We are interested in the model which would generate a fatter-tailed portfolio loss distribution, for the reason, see Chapter 6. In this section, we will consider some examples given in [23] of how the copula approach can be used for constructing loss distributions with fatter tails than it would be for the normally distributed asset value log-returns introduced in Chapter 4.

We look at the vector of asset value log-return but replace the assumption of multivariate normal distribution with multivariate *t distribution* for the reasons given in Section 6.3 of Chapter 6. We first recall some basic test distributions from statistics (see [17]).

Definition 5.2.1. (The Chi-square distribution) Given an i.i.d. sample $X_1, \dots, X_n \sim N(0,1)$, $X_1^2 + \cdots + X_n^2$ is said to be χ^2*-distributed* with n degrees of freedom.

Definition 5.2.2. (The Student's t-distribution) Given a standard normal variable $Y \sim N(0,1)$ and a χ^2-distributed variable $X \sim \chi^2(n)$, such that Y and X are independent. Then the variable Z defined by $Z = Y/\sqrt{X/n}$ is said to be t-distributed with n degrees of freedom.

In general the t-distribution has more mass in the tails than a normal distribution. Due to the property of a t-distribution that as the degree of freedom parameter ν goes to infinity, it converges to the normal distribution. If we start with an approximately normal distribution, we can gradually move away from this model by choosing smaller values of ν step-by-step.

Definition 5.2.3. (The multivariate t-distribution) Given a multivariate Gaussian vector $\mathbf{Y} = (Y_1, \dots, Y_m) \sim N(0, R)$ with correlation matrix R, the scaled vector $\theta\mathbf{Y}$ is said to be *multivariate t-distributed* with n degrees of freedom, if $\theta = \sqrt{n/X}$ with $X \sim \chi^2(n)$ and X is independent of $\mathbf{Y}$. And

5.2.1 t-Copula

Definition 5.2.4. [23] Given $n \geq 3$ and F_n a t-distribution function with n degrees of freedom, given the multivariate t-distribution function with n degrees of freedom and correlation matrix R, denoted by $F_{n,R} \sim t(n, R)$, we define a *t-copula* function as follows:

$$C_{n,R}(u_1, \ldots, u_d) = F_{n,R}(F_n^{-1}(u_1), \ldots, F_n^{-1}(u_d)), \quad u_1, \ldots, u_d \in (0, 1). \tag{5.4}$$

The copula $C_{n,R}$ incorporates a multivariate t-dependency structure. Similarly with Gaussian copulas, we can combine with a t-copula with any marginal distributions we like. For example, if we want to build a factor model which would generate the fatter-tailed loss distribution, we can choose a t-copula with Gaussian marginals, or a Gaussian copula with t-marginals just like the following:

Example 5.2.5. [23] A multivariate distribution function with t-dependency and Gaussian marginals can be defined by

$$F(x_1, \ldots, x_d) = C_{n,R}(N(x_1), \ldots, N(x_d)), \quad x_1, \ldots, x_d \in \mathbf{R},$$

where $N(\cdot)$ denotes the standard normal distribution function.

Example 5.2.6. [23] A multivariate distribution function with Gaussian dependency and t-marginals can be defined by

$$F(x_1, \ldots, x_d) = C_R^{Ga}(F_n(x_1), \ldots, F_n(x_d)), \quad x_1, \ldots, x_d \in \mathbf{R},$$

where $F_n(\cdot)$ denotes the student's t-distribution function.

Replacing Gaussian dependency by t-dependency, or replace Gaussian marginal by t-marginals will both significantly shift mass into the tail of the loss distribution arising from a corresponding factor model.

Figure 5.1 contrasts the lack of tail dependence of the normal copula with the strong tail dependence of the t copula with $n = 3$ degrees of freedom. The left hand plot shows 7000 points from bivariate normal distribution, whereas the right hand plot shows 7000 points from a bivariate t distribution. The correlation in each plot is 0.7. Clearly, in the lower left and upper right quadrants, the t dependence structure produces more joint extreme values close to the diagonal.

5.3 Common Copula Families

Article [24] divide copulas into two main families, the most frequently used copula families are Elliptical copulas and Archimedean copulas.

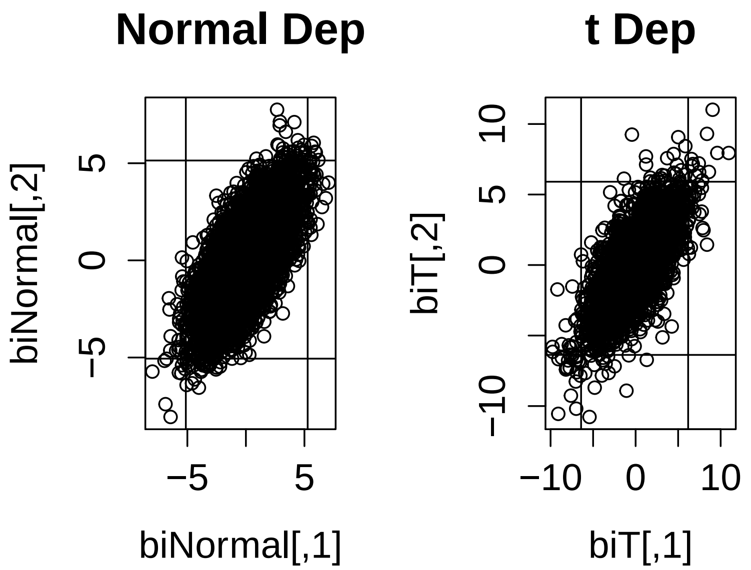

Figure 5.1: Normal dependence vs. t dependence. Vertical and horizontal lines at 99.5% and 0.5% quantiles of marginal distribution

5.3.1 Elliptical copula

An elliptical copula has a correlation matrix inherited from the corresponding elliptical distributions, which determines the dependency structure. J. Yan [24] implemented the four commonly used dependency structures in R: autoregressive of order 1, exchangeable, toeplitz, and unstructured, depending on the correlation between each variable. For example, in the case of dimension $d = 3$, the corresponding correlation matrices are as follows (see [24]):

$$\begin{pmatrix} 1 & \rho_1 & \rho_1^2 \\ \rho_1 & 1 & \rho_1 \\ \rho_1^2 & \rho_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_1 \\ \rho_1 & \rho_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}, \tag{5.5}$$

where ρ_j's are correlation parameters.

5.3.2 Archimedean copula

Another common copula family is Archimedean copula. Archimedean copulas are popular because they allow modeling dependency structure in arbitrarily high dimensions with only one parameter [12].

Family	Parameter α	Generator $\varphi(t)$	Inverse $\varphi^{-1}(s)$
Clayton	$\alpha \geq 0$	$t^{-\alpha} - 1$	$(1+s)^{-1/\alpha}$
Frank	$\alpha \geq 0$	$-\ln \frac{e^{-\alpha t}-1}{e^{-\alpha t}-1}$	$-\alpha^{-1} \ln(1 + e^{-s}(e^{-\alpha} - 1))$
Gumbel	$\alpha \geq 1$	$(-\ln t)^{\alpha}$	$\exp(-s^{1/\alpha})$

Table 5.1: Summary of three One-Parameter Archimedean copulas for $d > 2$

Definition 5.3.1. [24] A copula is called Archimedean if it has the following representation:

$$C(u_1, \ldots, u_d) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_d)), \tag{5.6}$$

where φ is a continuous, strictly decreasing and convex function, and is the so-called generator, φ^{-1} is the inverse of the generator φ. A generator uniquely determines an Archimedean copula.

Table 5.1 gives three common one-parameter multivariate Archimedean copulas ($d > 2$). It is worth to know that Archimedean copulas with dimension 3 or higher only allow positive correlation, whereas negative correlation is allowed for bivariate Archimedean copulas. Figure 5.2 compares four common bivariate copula, bivariate Gaussian (normal), Student-t, Gumbel, and Clayton copula. We can see from the figure that all of Student-t, Gumbel and Clayton copula have a fatter tail dependence than Gaussian copula.

5.4 R Implementation

In this section we will implement Li's model introduced in Chapter 7. Part of the approach was initially explored by Jun Yan in paper [24]. We do the following steps in our code:

1. We firstly plot 1000 random points generated from a trivariate normal copula and a trivarite t copula. (see Figure 5.3)
2. Then we generate a 200-sized object from a trivariate normal copula with exponential margins with respective parameter (hazard rate): 0.015, 0.02, 0.025. We set the correlation parameter as 0.5.
3. Finally, we use maximum likelihood method to fit this copula-based model. The result is shown in Figure 5.4.

The R code is as follows:

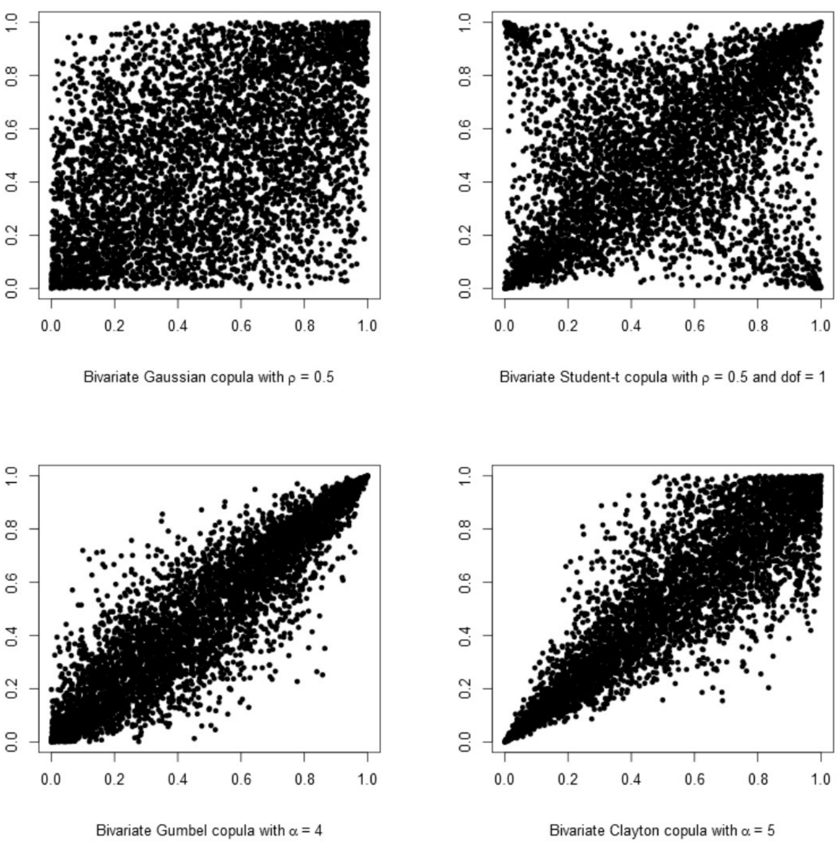

Figure 5.2: An example of the bivariate Gaussian (normal), Student-t, Gumbel, and Clayton copula. Source: [12]

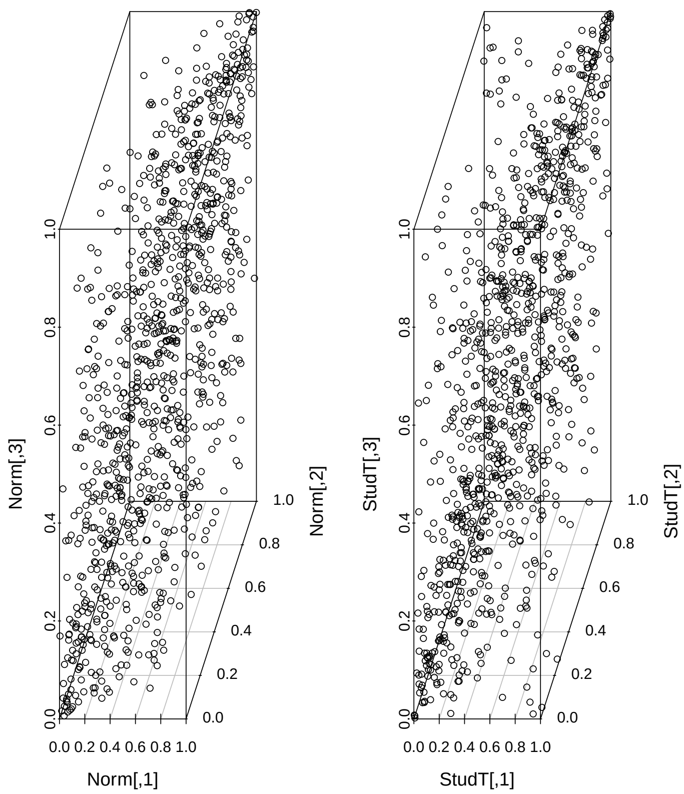

Figure 5.3: 3d scatter plots of random numbers from a normal copula and a t-copula

```
> loglike.Fit <- fitMvdc(dat, cop.normExp, c(0.015,0.020,0.025,0.5))
> loglike.Fit
The Maximum Likelihood estimation is based on 200 observations.
Margin 1 :
        Estimate Std. Error
m1.rate  0.01615      0.001
Margin 2 :
        Estimate Std. Error
m2.rate   0.0203      0.001
Margin 3 :
        Estimate Std. Error
m3.rate  0.02778      0.002
Copula:
      Estimate Std. Error
rho.1   0.4866      0.034
The maximized loglikelihood is -2845.281
Optimization converged
Number of loglikelihood evaluations:
function gradient
      99        8
```

Figure 5.4: The estimation of copula correlation and marginal parameters by maximum likelihood method

```
library(copula)
require(scatterplot3d)

###1. plot 1000 random points from a trivariate normal copula and a trivarite t copula
#create a 3-dim normal copula without specifying magins with correlation 0.5
myCop.norm <- ellipCopula(family=normal, dim=3, dispstr=ex, param=0.5)

#create a 3-dim t copula without specifying margins with correlation 0.5 and 3 degree of freedom
myCop.t <- ellipCopula(family=t, dim=3, dispstr=ex, param = 0.5, df=3)

#random generate 1000 points from a trivariate normal copula and a trivarite t copula
n <- 1000
Norm <- rcopula(myCop.norm, n)
StudT <- rcopula(myCop.t, n)

#plot 1000 random points
par(mfrow=c(1,2))
scatterplot3d(Norm)
scatterplot3d(StudT)

###2. generate a 200-sized sample from a trivariate normal copula with exponential margins with different parameter
```

```
cop.normExp <- mvdc(copula=myCop.norm, margins=c(exp,exp,exp),
    paramMargins = list(0.015,0.02,0.025))
dat <- rmvdc(cop.normExp, 200)

###3. using maximum likelihood method to fit the copula-based model
     generated from process 2
#the loglikelihood at the true parameter value:
loglike.True <- loglikMvdc(c(0.015,0.020,0.025,0.5), dat, cop.
    normExp)

#estimate the copula correlation and margins' parameter by moments
    estimate.
loglike.Fit <- fitMvdc(dat, cop.normExp, c(0.015,0.020,0.025,0.5))
```

copula.R

Chapter 6

Latent Variable Models

6.1 Latent Variable Models

The formalism of a latent variable underlies essentially all credit risk models derived from Merton's firm value model, as discussed in Chapter 4. Each latent variable model of credit risk contains three components: a latent variable, a threshold, and a corresponding binary default indicator with value of 0(non-default) or 1(default). A latent variable is typically chosen as a value of obligor's assets, and a threshold is associated with values of the long-term liabilities. If the value of the latent variable falls below the threshold, we know that the default happened, and as a result the value of the indicator will be 1.

Consider a portfolio of m obligors. At time $t = 0$ all obligors are assumed to be in a non-default state. Following Frey, McNeil and Nyfeler [13], we give the following formal definition of a latent variable model:

Definition 6.1.1. [13] Let $X = (X_1, \dots, X_m)$ be an m-dimensional random vector with continuous marginal distributions representing the latent variables at time T, and let $(D_1, ..., D_m)$ be a vector of deterministic cut-off levels. We call $(X_i, D_i)_{1\leq i\leq m}$ a *latent variable model* for the binary random vector $\mathbf{Y} = (Y_1, \dots, Y_m)$ if the following relationship holds:

$$Y_i = 1 \Leftrightarrow X_i \leq D_i.$$

In the factor model introduced in Chapter 4, the latent variables X_i are assumed to be Gaussian random variables and are interpreted as the relative changes in asset's log-returns.

For modelling of a portfolio's credit risk, Definition 6.1.1 should be complemented with a specification of the dependency structure between the latent variables. This dependency plays a crucial role in determining large losses in the portfolio. Again, in the factor model introduced in Chapter 4, the dependency between latent variables X_i is a multivariate Gaussian with a uniform correlations ρ (we have looked at a way of estimating ρ in Chapter 4).

It is natural to specify this dependency using a *copula* approach. Copula gives a way of putting marginal distribution of an individual obligor's asset return, or survival time in Li's case [5], together to form a joint distribution of groups of risks. A good introduction to copulas is provided in [12].

Copula approach allows to provide an alternative definition of the equivalence of latent variable models as was also pointed out by Frey, McNeil and Nyfeler [13].

6.2 Equivalence for Latent Variable Models

Consider the following definition of the structurally equivalent latent variable models.

Definition 6.2.1. [13]**(Equivalence for latent variable models)**
Let $(X_i, D_i)_{1\leq i\leq m}$ and $(X_i', D_i')_{1\leq i\leq m}$ be two latent variable models generating default indicator vectors $\mathbf{Y}$ and $\mathbf{Y}'$. The models are called equivalent if $\mathbf{Y} \stackrel{\mathrm{d}}{=} \mathbf{Y}'$, i.e. $\mathbf{Y}$ and $\mathbf{Y}'$ has the same distribution.

In other words, the equivalence in distribution for the corresponding default indicators defines the equivalence of the latent variable models.

As stated in [13], a sufficient condition for two latent variable models to be equivalent is that individual default probabilities are the same in both models and the copulas of the latent variables are the same. The following propositions was proved in Frey, McNeil and Nyfeler [13].

Proposition 6.2.2. *[13] Consider two latent variable models $(X_i, D_i)_{1\leq i\leq m}$ and $(X_i', D_i')_{1\leq i\leq m}$ with default indicator vectors $\mathbf{Y}$ and $\mathbf{Y}'$. These two models are equivalent if:*

1. $P(X_i \leq D_i) = P(X_i' \leq D_i'), i \in \{1, \ldots, m\}$, *and*

2. $\mathbf{X}$ *and* $\mathbf{X}'$ *have the same copula.*

Thus, as stated in [13], even if the terms defining the model $(X_i, D_i)_{1\leq i\leq m}$ are interpreted and calibrated in different ways, the models still can be structurally equivalent.

Gaussian copula, whose definition we have already given in Chapter 5, is the latent variable dependence structure which implicitly underlies all standard industry models.

6.3 Change of Dependence Structure

Many factor models in credit risk assume the multivariate normality of the corresponding latent variables. However, this choice is not supported by solid empirical evidence - unfortunately, the world is not Gaussian. Also, it has been shown in [13] that the aggregate portfolio loss distribution can be very sensitive to the exact nature of the multivariate distribution of the latent variables.

The models which lead to heavy-tailed loss distributions, can be developed even keeping the individual default probabilities of obligors and the matrix of latent variable correlations fixed.

It is elegant to use a *copula*, which we will introduce in Chapter 5, as a bridge to connect a multivariate latent variable distribution with the portfolio loss distribution which banks are mostly interested in. If we want a model which can generate a heavier-tailed loss distribution, which represents the higher simultaneous joint default probability, we can simply choose a copula which has the property of heavier tail dependence.

To illustrate this point, Frey, McNeil and Nyfeler [13] used the t-distribution for the following reasons:

- As the degree of freedom parameter ν goes to infinity, the t-distribution converges to the normal distribution. Therefore if we start with an approximately normal distribution, we can gradually move away from this model by choosing smaller values of ν.

- The t copula is very different to the Gaussian copula. It has the property of tail dependence, so that it tends to generate simultaneous extreme events, such as bigger losses (see Figure 6.1 [11]), with higher probabilities than the Gaussian copula [21]. This is exactly what we want, since we want a realistic credit model which is able to give sufficient weight to scenarios where large joint defaults occur.

6.4 Alternative Latent Variable Model Proposed by Li (1999)

We leave the detailed mathematical definition of copula to Chapter 5. In this section we only explain a brief idea of Li's model in 1999 for the preparation of the next chapter.

In 1999, David X. Li [5] took another route of relaxing constraints of Guassian models, which we will discuss in Chapter 7.

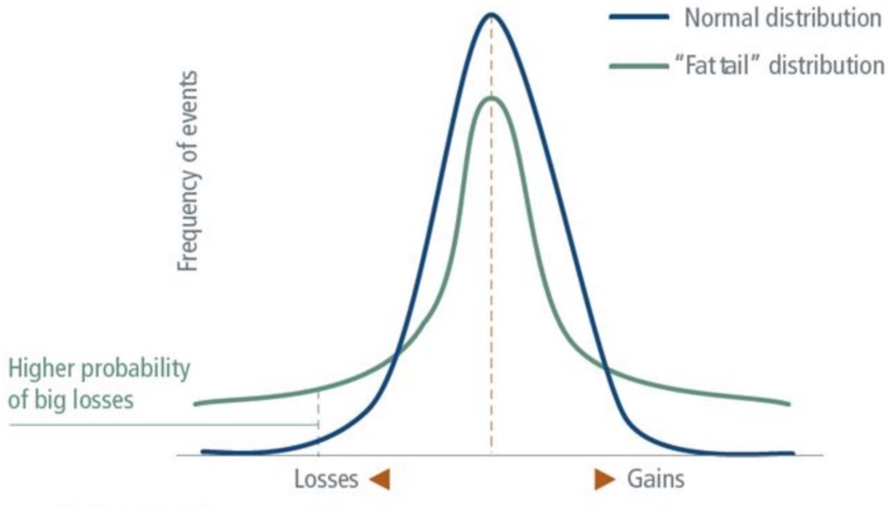

Figure 6.1: Normal vs. fat-tailed loss distribution (Source: [11])

Using the converse of Sklar's Theorem (see Proposition 5.1.5), assuming there are m entities, a copula of the form:

$$C_R^{Ga}(1 - exp(\lambda_1 x_1), \ldots, 1 - exp(\lambda_m x_m)),$$

where $\lambda_1, \ldots, \lambda_m$ are the parameters of each marginal exponential distribution, can be used to model the whole dependency structure.

We will elaborate on Li's model and the meaning of the parameter λ_i, which is actually the hazard rate in survival analysis, in Chapter 7.

Chapter 7

Li's Model and Hazard Rate

7.1 Main Idea of Li

In 1999, David X. Li [5] built a special latent variable model, where he modeled the survival time of each defaultable entity i as a latent variable X_i. He assumed that each X_i satisfies an exponential distribution marginally, but globally the vector of these latent variables $\mathbf{X} = (X_1, \ldots, X_m)$ has a Gaussian copula, with the correlation as the same as the correlation between the variables X_i.

So basically Li applied the essence of survival analysis widely used in biological statistical modeling, to the credit risk modeling. He first made a very interesting choice of the latent variable, which is the survival time rather than the asset return of an entity, and built a model still using normal dependence, but with exponential marginal distribution with a special parameter, which is so-called hazard rate.

7.1.1 Survival function

Inspired by [5], let us first consider an individual obligor. This obligor's time-until-default, T, is a continuous random variable, which measures the length of time from today to the time when default occurs.

First we give the definition of the *survival function* given in [5]:

$$S(t) = \mathbb{P}(T > t), \quad t \geq 0. \tag{7.1}$$

It gives the probability that an obligor will survive until time t.

Let $F(t)$ denote the distribution function of T,

$$F(t) = \mathbb{P}(T \leq t) = 1 - S(t), \quad t \geq 0. \tag{7.2}$$

If the obligor has already survived x years, the future life time for this obligor has the conditional distribution of $T - x$ given $T > x$. Li [5] also introduced two more notations,

$$q(t, x) = \mathbb{P}[T - x \leq t | T > x], \quad t \geq 0, x \geq 0. \tag{7.3}$$

$$p(t, x) = 1 - tq_x = \mathbb{P}[T - x > t | T > x], \quad t \geq 0, x \geq 0. \tag{7.4}$$

$q(t, x)$ can be interpreted as the conditional probability that the obligor will default within the next t years conditional on its survival for x years, and $p(t, x)$ can be interpreted as the conditional probabilty that the obligor will still survive t years more on its survival for x years. In the special case of $x = 0$, we have $p(t, 0) = S(t), \quad t \geq 0$.

If $t = 1$, we have following definition:

Definition 7.1.1. [5]

$$q_x = \mathbb{P}[T - x \leq 1 | T > x]$$

is called the *marginal default probability*, which gives the probability of default in the next year conditional on the survival until the beginning of this year.

7.1.2 Hazard rate function

The hazard rate is defined as the instantaneous default probability rate for an obligor that has survived until age x. It can be seen as the conditional default rate during the next instant of time.

Definition 7.1.2. The function

$$h(x) = \frac{f(x)}{1 - F(x)}$$

is called *hazard rate function*, where $f(x) = F'(x)$ is the probability density function of T.

The relationship of the hazard rate function with the survival function is as follows:

$$h(x) = -\frac{S'(x)}{S(x)}, \quad x \geq 0.$$

So, the survival function can be expressed in terms of the hazard rate function:

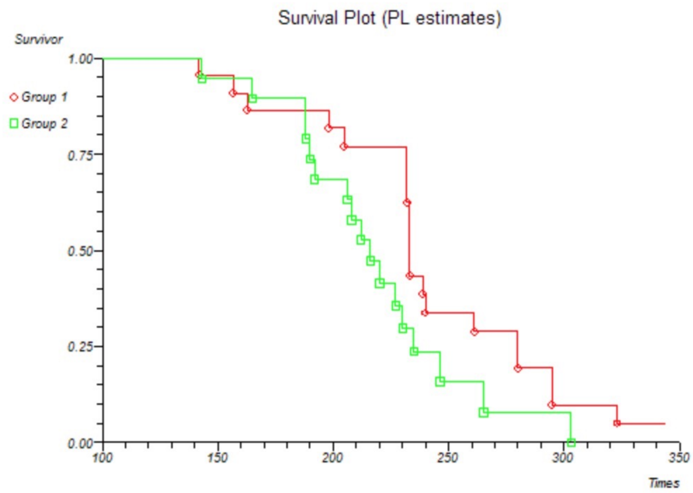

Figure 7.1: two obligor's survival plot (Source: [1])

$$S(t) = e^{-\int_0^t h(s)ds}, \quad t \geq 0. \tag{7.5}$$

In addition the distribution function is

$$F(t) = 1 - S(t) = 1 - e^{-\int_0^t h(s)ds}, \quad t \geq 0. \tag{7.6}$$

The hazard rate has many similarities with the short-term interest rate [5]. Therefore many modeling techniques for the short-term interest rate can be applied to model the hazard rate. In reality, people usually assume that the hazard rate is a constant, h. This is the key assumption for our following section, where we provide three ways to calculate this constant. In this case, the density function is

$$f(t) = F'(t) = he^{-ht},$$

which shows that the survival time follows an exponential distribution with parameter h.

Generally, survival analysis involves the modeling of time to 'event', for example, death or failure, therefore it is widely used in biological organizations and medical institutions. To visualize the survival time and its relationship with hazard rate, we found two plots from [1] (see Figure 7.1, Figure 7.2), it is obvious that the survival declines with time.

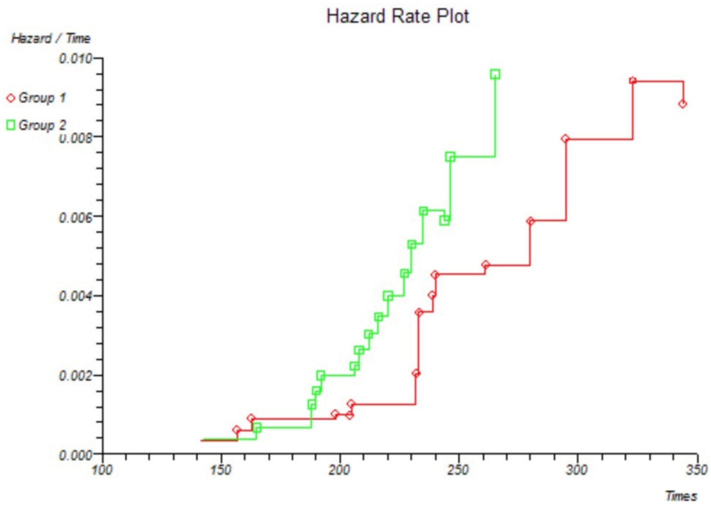

Figure 7.2: Two obligor's hazard rate plot (Source: [1])

7.1.3 Joint survival function

Finally, we give the definition of the joint survival function for two obligors introduced in [20]. Given obligors A and B, based on their survival times T_A and T_B, we have the joint survival function

$$S_{T_A T_B}(s,t) = \mathbb{P}[T_A > s, T_B > t],$$

and the joint distribution function is

$$F(s,t) = \mathbb{P}[T_A \leq s, T_B \leq t] = 1 - S_{T_A}(s) - S_{T_B}(t) + S_{T_A T_B}(s,t).$$

7.2 Estimation of Constant Hazard Rate

The term structure of default rates can be obtained in three significantly different ways:

- From time series of historical default rates provided by rating agencies like Moody's and Fitch.
- From market prices of defaultable bonds or asset swap spreads.
- From a framework of the Merton model.

We will next discuss each of these approaches.

Years	1	2	3	4	5	7	10	15	20
Aaa	0.000	0.012	0.012	0.037	0.105	0.245	0.497	0.927	1.102
Aa	0.022	0.059	0.091	0.159	0.234	0.384	0.542	1.150	2.465
A	0.051	0.165	0.341	0.520	0.717	1.179	2.046	3.572	5.934
Baa	0.176	0.494	0.912	1.404	1.926	2.996	4.851	8.751	12.327
Ba	1.166	3.186	5.583	8.123	10.397	14.318	19.964	29.703	37.173
B	4.546	10.426	16.188	21.256	25.895	34.473	44.377	56.098	62.478
Caa	17.723	29.384	38.682	46.094	52.286	59.771	71.376	77.545	80.211

Table 7.1: Average cumulative default rates (%), 1970 - 2009. Source: Moody's.

7.2.1 Method 1: Historical default probability

Rating agencies, such as Moody's, S&P, and Fitch, systematically rate creditworthiness of corporate bonds. For example, Moody's long-term rating scale consists of 9 categories, Aaa, Aa, A, Baa, Ba, B, Caa, Ca, and C, in order from the best to the worst rating category. Assignment of a rating category is not purely model based, and takes into account a whole range of qualitative and quantitative factors.

Table 7.1 gives a typical example of data provided by the rating agencies. It shows the cumulative default rates for the corporate bonds with a particular rating between 1 and 20 years within a 39-years observation window. As illustrated in [3], for example, a bond with an A credit rating has a 0.051% chance of defaulting during the first year, a 0.165% chance of defaulting by the end of the second year, and so on. The probability of a bond defaulting during a particular year can be calculated from the table. For example, the probability that an initially rated A bond will default during the second year is $0.165\% - 0.051\% = 0.114\%$.

Note that the probability of default within a year can be both increasing and decreasing function of time. Typically, it is an increasing function for investment-grade bonds (e.g., the probabilities of an A-rated bond defaulting during years 0-5, 5-10, 10-15, and 15-20 are 0.717%, 1.329%, 1.526%, and 2.362%, respectively): the bond issuer is initially considered to be creditworthy, and the factors affecting its financial health arrive rather randomly over a long period of time. For bonds with a poor credit rating, the probability of default can be a decreasing function of time (e.g., the probabilities that a B-rated bond will default during years 0-5, 5-10, 10-15, and 15-20 are 25.895%, 18.482%, 11.721%, and 6.380%, respectively). For these bond issuers the factors which might be leading to default are already identifiable, and the next year or two may be critical for all of them. However, the longer such an issuer survives, the greater the chance of improvement for its financial health.

Hazard rate estimation from historical default probability

The hazard rates can be easily estimated from Table 7.1. For example, the *unconditional default probability* for a Caa-rated bond during the third year as seen at time 0 is 38.682 - 29.384=9.298%. Its survival probability in the first two years is 100 - 29.384 = 70.616%. The default probability during the third year conditional on no earlier default is therefore 0.09298/0.70616, or 13.17%. These conditional default probabilities are the *hazard rates* entering equation (7.5).

Obviously, we can transform the default probability function (7.6)

$$F(t) = 1 - e^{-\int_0^t h(s)ds}$$

into

$$F(t) = 1 - e^{\overline{h}(t)t}, \tag{7.7}$$

where $\overline{h}(t)$ is the average hazard rate(or default intensity) between time 0 and time t.

Example 7.2.1. For an A-rated company, if we want to calculate the default intensity using historical data and based on equation (7.7), when $t = 7$, we have

$$\overline{h}(7) = -\frac{1}{7}\ln[1 - F(7)].$$

The value of $F(7)$ is taken directly from table 7.1, which is 0.01179. The average 7-year hazard rate is therefore

$$\overline{h}(7) = -\frac{1}{7}\ln[0.98821] = 0.0017.$$

7.2.2 Method 2: Estimate hazard rate from bond price

This methods looks at estimating the hazard rate without using historical default rates but based on the other market information.

Recovery rate

The recovery rate for a bond is typically defined as the bond's market value (as a fraction of its face value) right after a default. Table 7.2 provides historical average recovery rates for different categories of bank loans and bonds in the United States. It shows that bank loans with a first lien on assets had the best average recovery rate, 65.6%. For bonds, the average recovery rate ranges from 49.8% for those that are both senior to other lenders and secured to 24.7% for

Class	Average recovery rate (%)
First lien bank loan	65.6
Second lien bank loan	32.8
Senior unsecured bank loan	48.7
Senior secured bond	49.8
Senior unsecured bond	36.6
Senior subordinated bond	30.7
Subordinated bond	31.3
Junior subordinated bond	24.7

Table 7.2: Recovery rates on corporate bonds as a percentage of face value, 1982-2009. Source: Moody's.

those that rank after other lenders with a security interest that is subordinate to other lenders.

Interestingly, recovery rates are significantly negatively correlated with default rates. This means that a bad year for the default rate is usually doubly bad because it is accompanied by a low recovery rate. Moody's looked at average recovery rates and the average default rates each year between 1982 and 2009, and found [16] that the following relationship provides a good fit to the data:

$$\text{Average recovery rate} = 0.503 - 6.3 \times \text{Average default rate}. \tag{7.8}$$

Hazard rate estimation from bond price

In the simplest approach, the only reason for the price of a corporate bond being lower than a similar risk-free bond (this difference in financial jargon is called 'credit spread') is the possibility of default. Consider first the following approximate calculation, John C. Hull in [3] supposes that a bond yields 2% more than a similar risk-free bond and that the expected recovery rate in the event of a default is 40%, from the expectation to lose 2% per year from defaults and the recovery rate of 40%, an estimate of the probability of a default per year conditional on no earlier default is 0.02/(1-0.4), or 3.33%. In a more formal form,

$$h = \frac{s}{1-R}, \tag{7.9}$$

where h is the hazard rate per year, s is the annualised spread of the corporate bond yield over the risk-free rate, and R is the expected recovery rate.

To calculate average hazard rates from bond prices, we use equation (7.9) and bond yields published by Merrill Lynch (see [3]). The recovery rate is assumed to be 40%. To calculate the bond yield spread, we assume that the risk-free interest rate is the 7-ear swap rate minus 10 basis points (see [3]).

Example 7.2.2. [3] For an A-rated bond, the average Merrill Lynch yield was

5.995%. The average 7-year swap rate was 5.408%, so that the average risk-free rate was 5.308%. This gives the average 7-year hazard rate as

$$\frac{0.05995 - 0.05308}{1 - 0.4} = 0.0115,$$

or 1.15%.

7.2.3 Comparison between method 1 and method 2

The described approaches differ significantly and might lead to quite different results. People intend to use market information rather than historical information for the following reasons:

- A bank market unit is required to base its calculation of profit and loss on current market information. This information reflects the market expectations about the future which will determine the actual profit and loss. This forward-looking view is not present in historical default data.
- Market's response in anticipation of future credit quality modifications is much faster than the rating agencies. A typical example is the rating agencies reaction to the Asian financial crisis in 1997.
- Factors influencing deterioration of corporate's credit quality change over long periods of time which make the longer term estimates of default probabilities unstable.

7.2.4 Method 3: Using Black-Scholes formula to estimate hazard rate

K. Merton [9] pioneered an approach of modelling a company's equity as an option on the assets of the company. Suppose that a firm has one zero-coupon bond outstanding with maturity time T. Let $A(t)$ be the value of the company's assets at time t, $E(t)$ the value of its equity, and $D(T)$ be the debt repayment due at time T, and let σ_A and σ_E be the volatility of the assets (assumed constant) and instantaneous volatility of the equity, respectively.

If $A(T) = D$, the value of its equity is zero, therefore it is (at least in theory) rational for the company to default on the debt at time T. If $A(T) > D$, the company should make the debt repayment at time T, therefore, the value of the firm's equity at time T in Merton's model is

$$E_T = \max(A(T) - D, 0).$$

John C. Hull in [3] gives a way to estimate hazard rate using Black-Scholes formula. He shows that the equity is a call option on the value of the assets with a strike price equal to the repayment required on the debt. The Black-Scholes-Merton formula gives the value of the equity today as

$$E(0) = A(0)N(d_1) - De^{-rT}N(d_2), \tag{7.10}$$

where

$$d_1 = \frac{\ln(A(0)/D) + (r + \sigma_A^2/2)T}{\sigma_A\sqrt{T}},$$

and

$$d_2 = d_1 - \sigma_A\sqrt{T},$$

and N is the univariate standard normal distribution function. The value of the debt today is $A(0) - E(0)$.

The risk-neutral probability that the company will default on the debt is $N(-d_2)$. To calculate this, we require $A(0)$ and σ_A. Neither of these are directly observable. However, if the company is publicly traded, we can observe $E(0)$ and σ_E. Also, from Ito's formula and explanation in [3],

$$\sigma_E E(0) = \frac{\partial E}{\partial A}\sigma_A A(0). \tag{7.11}$$

This provides another equation that must be satisfied by $A(0)$ and σ_A, so combining equation (7.10) and (7.11), we can get $A(0)$ and σ_A.

Example 7.2.3. (Source:[3]) The value of a company's equity is \$3 million and the volatility of the equity is 80%. The debt that will have to be paid in 1 year is \$10 million. The risk-free rate is 5% per annum. In this case $E(0) = 3$, $\sigma_E = 0.80$, $r = 0.05$, $T = 1$, and $D = 10$. Solving equations (7.10) and (7.11) yields $A(0) = 12.40$ and $\sigma_V = 0.2123$. The parameter d_2 is 1.1408, so that the probability of default is $N(-d_2) = 0.127$, or 12.7%. The market value of the debt is $A(0) - E(0)$, or 9.40. The present value of the promised payment on the debt is $10e^{-0.05\times 1} = 9.51$. The expected loss on the debt is therefore $(9.51 - 9.40)/9.51$, or about 1.2% of its no-default value. The expected loss (EL) equals the probability of default (PD) times one minus the recovery rate. It follows that the recovery rate equals one minus EL/PD. In this case, the recovery rate is $(12.7 - 1.2)/12.7$, or about 91%, of the debt's no-default value.

Chapter 8

Further work

8.1 Further work

We found there are still some interesting problems which can be researched on in the future, for example,

- For the asset price model described in (2.5) of Chapter 2, we have assumed that the stochastic part is a Brownian Motion. Based on this simple diffusion model, Vasicek derived the portfolio credit loss distribution function. However, we are interested in a kind of asset model, whose stochastic part is not standard Brownian Motion, but with jumps. Also in the future we hope to derive a loss distribution based on jump-diffusion models.

- In this thesis and in practice, hazard rate is usually being assumed to be a constant. This definitely simplified the risk modeling. However, we are interested in a stochastic hazard rate, and we want to find a way to model hazard rate in the future research.

- In Chapter 5, Figure 5.2 gives an example of the bivariate Gaussian, Student-t, Gumbel, and Clayton copula. It is obvious that all of Student-t, Gumbel, and Clayton copula have a fatter tail than the Gaussian copula. However, we also see that the tails of the Gumbel and Clayton copula are fatter in a different way. In the future we want to go deeper to the comparison between Archimedean copulas, in order to use them better to model dependency structure.

Bibliography

[1] David Collett. *Modelling Survival Data in Medical Research.* Chapman and Hall, 1994.

[2] Glyn Davies. *A History of Money: from Ancient Times to the Present Day.* University of Wales Press, 1996.

[3] John C. Hull. *Options, Futures and Other Derivatives (7th ed.).* Prentice Hall, 2008.

[4] Investopedia. Risk-based mortgage pricing. *Investopedia*, 2014.

[5] David X. Li. On default correlation: A copula function approach. *The RiskMetrics Group Working Paper Number 99-07*, 2000.

[6] Donald Mackenzie. *An Engine, Not a Camera: How Financial Models Shape Markets.* Cambridge, MA: MIT Press, 2006.

[7] Bodie Zvi; Alex Kane; Alan J. Marcus. *Investments (7th ed.).* New York: McGraw-Hill/Irwin, 2008.

[8] Damiano Brigo; Massimo Masetti. *Risk Neutral Pricing of Counterparty Risk.* Risk Books, 2006.

[9] Robert Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29 (1974): 449–70.*

[10] Laurence H. Meyer. Why risk management is important for global financial institutions. 2000.

[11] Amit Minhas. Good ceo, bad ceo and fat tails. *LinkedIn*, 2015.

[12] Roger B. Nelsen. *An Introduction to Copulas.* Springer, 1999.

[13] Rudiger Frey; Alexander J. McNeil; Mark A. Nyfeler. Copulas and credit models. 2001.

[14] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2010.

[15] Basel Committee on Banking Supervision. Sound credit risk assessment and valuation for loans. June 2006.

[16] D.T. Hamilton; P. Varma; S. Ou and R. Cantor. Default and recovery rate of corporate bond issuers. *Moody's Investor's Service*, 2004.

[17] Wiebe R. Pestman. *Mathematical Statistics.* Walter de Gruyter (2 edition), 1998.

[18] Fischer Black; Myron Scholes. The pricing of options and corporate liabilities. 1973.

[19] A. Sklar. Fonction de repartition an dimension et leur marges. *Publications de l'Insitute Statistique de l'Universit e de Paris*, pages 8:229–231, 1959.

[20] Peter J. Smith. *Analysis of Failure and Survival Data.* CRC Press, 2002.

[21] Paul Embrechts; Alexander Mcnell; Daniel Straumann. Correlation: pitfalls and alternatives. *Risk, 12(5): 93-113*, 1999.

[22] Oldrich Vasicek. The distribution of loan portfolio value. *Risk Vol.15, No.12*, 2002.

[23] Christian Bluhm; Ludger Overbeck; Christoph Wagner. *An Introduction to Credit Risk Modeling.* CRC Press Company, 2003.

[24] Jun Yan. Enjoy the joy of copula: With a package copula. *Journal of Statistical Software*, 2007.