Imperial College London

MASTER THESIS

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

Calibrability of First to Default correlation structure

Authors:
AMER-OUALI Nassim

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Date and signature : 8th of September, 2020

Acknowledgments

I would like to thank BNP Paribas for giving me the opportunity of doing this internship, and especially Fakher Ben Atig for his benevolence and for helping me integrating into the team in this particular period, and Sebastien Hitier and Charles Sicat for coordinating my internship.
I would also like to express my gratitude to my thesis internal supervisor Damiano Brigo for his availability, his guidance and the expertise he brought me.

Abstract

This thesis aims at studying the correlation structure of First-to-default baskets. We start by reviewing some theoretical results about copulas, which we will use in the models we will examine. After explaining general results about First-to-default baskets, we will present the Gaussian Latent Variable Model and an alternative model, the Student-*t* Copula Model. We will highlight the main differences between these two models and analyze the First-to-default baskets sensitivity to correlation under the assumptions of each of these models theoretically in a first time, and numerically in a second time.

Contents

1 Introduction **6**

2 Preliminaries **8**
2.1 Copulas . . . 8
2.1.1 Definition and first properties . . . 8
2.1.2 Fundamental copulas . . . 9
2.1.3 Implicit copulas . . . 9
2.1.4 Explicit copulas . . . 10
2.1.5 Kendall's tau . . . 11
2.1.6 Tail dependence . . . 12
2.1.7 Summary of the characteristics of the main copulas . . . 13
2.2 Linear Regression Methods . . . 13
2.2.1 Ordinary Least-squares . . . 14

3 Description of the data **17**
3.1 Default baskets pricing . . . 17
3.2 Default dependency . . . 18

4 Models **22**
4.1 The Gaussian Latent Variable Model . . . 22
4.1.1 Standard model . . . 22
4.1.2 The One-factor Copula Model . . . 22
4.1.3 Conditional Hazard Rates . . . 25
4.2 Student-t copula model . . . 27
4.3 Monte Carlo pricing of a First-to-default basket . . . 31
4.4 Section 4 Appendix . . . 32
4.4.1 Box-Muller . . . 32
4.4.2 Importance sampling . . . 34

5 Calibration results **38**
5.1 Model's results . . . 38
5.1.1 First-to-default probability . . . 39
5.1.2 FTD Spreads . . . 44
5.2 Implied correlation and historical correlation . . . 46

6 Summary and conclusions **49**

1 Introduction

Credit derivatives have started to take an important place in the derivatives market in the late 1990s. The amount of the outstanding transactions was equal to $40 billion in 1996, and by the crisis of 2007, this total notional principal was $50 trillion [2], and has declined after the crisis down to $10 trillion in 2018. Credit derivatives have been the fastest growing derivatives of the last two decades and important developments have been made in this field in terms of mathematical modeling, especially since the market has become more and more affected by new regulations and external events recently, such as the Greek debt crisis in 2015, the Great Britain's withdrawal of the European Union or even more recently, the COVID-19 pandemic. Such events have a huge impact on the industry and expose the companies to huge losses. Hence financial companies and investment banks in particular have to pay a particular attention to these situations and reinforce their risk management. A credit derivative is a contract which payoff depends on one or multiple credit events. A credit event usually refers to a default from an obligator, which means that the obligator is not able to honour the terms of a contract he is into, in particular when he cannot provide payments that he agreed to supply to a counterparty. When such an event occur, a certain amount is paid by one of the counterparties as a protection against the event. As these contracts depend on the creditworthiness of the entities involved, there exist credit ratings, which consist of a measure of the creditworthiness of a borrower. Those assessments are usually given by rating agencies, the most popular ones being Moody's, S&P, and Fitch, who assign a mark to corporate bonds. For SP and Fitch, the ratings, in the decreasing order are AAA, AA, A, BBB, BB, B, CCC, CC and C. For a more accurate measure, these ratings are usually divided into sub-ratings.

We can distinguish two categories of credit derivatives : single-name credit derivatives and multi-name credit derivatives. This first category refers to derivatives based on the credit risk of a single entity, while the second one refers to derivatives based on the credit risk of several entities at once. In this thesis, we will mainly focus on multi-name credit derivatives, and more specifically on First-to-default baskets, which consist of a contract involving two counterparties, a protection buyer and a protection seller. The buyer pays periodic fees to the seller of the protection, who, in case of a credit event, pays a certain amount to the buyer. A First-to-default basket contains multiple names (typically between 5 and 10 names of issuers) and the credit events which triggers the payment from the protection buyer is the first default from one of the issuers in the basket. We will give more details and provide a more formal definition of these type of instruments in the further sections.

The dependence structure between the default times of the names contained in a First-to-default basket has a key role and the credit derivatives market has seen the emergence of a liquid market in credit default correlation. A very popular way of modeling the default times is the Gaussian copula model, which makes the assumption that the time defaults of a basket are linked via a gaussian copula, which dramatically eases the computations but which has also been very controversial, to such a point that the Gauss Copula function, also known as Li's formula, has been qualified as "The formula that killed Wall Street" by journalist Felix Salmon in 2009.

However, quantitative analysts have a more nuanced view on this last point and this model (amongst others) is still used today. In this paper, we will analyze this model and evaluate the impact of the dependence structure between the defaults (or default correlation, as the correlation parameter is enough to describe the whole joint distribution of the default times in the case of a Gaussian copula) over the characteristics of a First-to-default basket. We will also compare this model with the Student-t copula model, which is another popular way of modeling the default times in a baskets and try to highlight the key characteristics of the two.

2 Preliminaries

2.1 Copulas

2.1.1 Definition and first properties

Copulas are used to build a multivariate distribution by choosing separately the marginal distributions and the dependence structure. More formally, a n-dimensional copula is a joint cumulative distribution function on $[0,1]^n$ with uniform marginal distributions.

We say that a function $C : [0,1]^n \to [0,1]$ is a copula if and only if :
1. $\forall i = 1,\dots,n$, $C(u_1,\dots,u_n)$ is an increasing function of u_i
2. $\forall i = 1,\dots,n,$, $C(1,...,u_i,...,1) = u_i$ and $u_i \in [0,1]$
3. $\forall (a_1,\dots,a_d),(b_1,\dots,b_d)$ such that $\forall i = 1,\dots,d$, $a_i \leq b_i$,

$$\sum_{i_1=1}^{2} \cdots \sum_{i_{d=1}}^{2} (-1)^{\sum_{k=1}^{4} i_k} C\left(u_{1i_1},\dots,u_{di_d}\right) \geq 0 \tag{1}$$

with $u_{j1} = a_j$ and $u_{j2} = b_j$ $\forall j = 1,\dots,d$.

The following theorem, known as Sklar's theorem, is a very important one as it enables to build a multivariate distribution by selecting any marginal cumulative distribution functions and a dependence structure to link them with a copula.

Theorem 1 (Sklar) Let $X_1,\dots,X_n$ be random variables with a joint distribution F and marginals $F_1,\dots,F_n$. Then, there exist a copula $C : [0,1]^n \to [0,1]$ such that for all $(x_1,\dots,x_n) \in \overline{\mathbb{R}}^n$,

$$F(x_1,\dots,x_n) = C(F_1(x_1),\dots,F_n(x_n) \tag{2}$$

Moreover, if the marginals $F_1,\dots,F_n$ are continuous, the copula is uniquely defined. Otherwise, the copula is uniquely defined on $F_1(\overline{\mathbb{R}}) \times \dots \times F_n(\overline{\mathbb{R}})$.

Copulas have other key properties commonly used which we list below :

Proposition 1 (Invariance to strictly increasing transformations) Let $(X_1,\dots,X_n)$ be a vector of random variables with continuous marginal cdfs with copula C. Let $\{T_i\}_{i=1,\dots,n}$ be a family of strictly increasing functions, then the copula of $(T_1(X_1,\dots,T_n(X_n))$ is also C.

Proposition 2 (Fréchet-Hoeffding bounds) For any copula C, the following inequalities hold

$$\max\left\{\sum_{i=1}^{n} u_i + 1 - n, 0\right\} \leq C\left(u_1,\dots,u_n\right) \leq \min\left\{u_1,\dots,u_n\right\} \tag{3}$$

for any $(u_1,\dots,u_d) \in [0,1]^n$.

We can classify the copulas into three possible categories : fundamental copulas, implicit copulas and explicit copulas.

2.1.2 Fundamental copulas

There exist three fundamental copulas, each corresponding to a specific dependence structure. The first one is the independence copula which is the dependence structure of any mutually independent continuous random variables. It is given by

$$C(u_1,\ldots,u_n) = \prod_{i=1}^{n} u_i \tag{4}$$

The independence copula is the distribution function of mutually independent uniform random variables $(U_1,\ldots,U_n)$.
The second fundamental copula is the comonotonicity copula given by

$$C(u_1,\ldots,u_n) = min(u_1,\ldots,u_n) \tag{5}$$

This copula corresponds to a perfect positive dependence structure and it is the joint cumulative distribution function of $(U,\ldots,U)$ with $U \sim Unif(0,1)$.
Finally, the third fundamental copula only makes sense in two dimensions and is called the countermonotonicity copula given by

$$C(u_1,u_2) = max(u_1 + u_2 - 1, 0)(6)$$

This copula corresponds to a perfect negative dependence structure and it is the joint cumulative distribution function of $(U, 1-U)$ with $U \sim Unif(0,1)$.
We will show in the following sections that the fundamental copulas can be build using implicit or explicit copulas in limit cases scenarios of their parameters.

2.1.3 Implicit copulas

The implicit copulas do not have simple expressions as they are build with known multivariate distribution functions via Sklar's theorem. Hence, they have to be simulated. The two copulas we will work with through this project are the Gaussian copula and the Student-t copula, which are both implicit copulas.

The Gaussian copula is the most frequently used copula, particularly in the context of modelling default times, as we will detail further in this thesis.
Let $(X_1,\ldots,X_n)$ a vector of Gaussian random variables with means $\nu = (\nu_1,\ldots,\nu_n)$ and variances $\sigma^2 = (\sigma_1^2,\ldots,\sigma_n^2)$, with correlation matrix $\Sigma = (\rho_{i,j})_{i=1,\ldots,n;j=1,\ldots,n}$. Define

$$U_i := \Phi(\frac{X_i - \mu_i}{\sigma_i}) \text{ for } i = 1,\ldots,n \tag{7}$$

Φ being the cumulative distribution function of a univariate standard normal distribution.
Then, the distribution function $C_{\Sigma}^{GC}(u_1,\ldots,u_n)$ of the $\{U_i\}_{i=1,\ldots,n}$ is the Gaussian copula associated to the correlation matrix Σ.

If $\Phi_{n,\Sigma}$ denotes the n-dimensional normal distribution, we have

$$C_{\Sigma}^{Ga}(u_1,\ldots,u_n) = \Phi_{n,\Sigma}(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_n)) \tag{8}$$

There is no closed form for this copula, but we can write it as

$$C_{\Sigma}^{Ga}(u_1,\ldots,u_n) = \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_n)} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}} \sum^{-1} (x-\mu)\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{9}$$

We can note that the Gaussian copula does not capture tail dependence, and that for any permutation σ,

$$C_{\Sigma}^{Ga}(u_{\sigma(1)},\ldots,u_{\sigma(n)}) = C(u_1,\ldots,u_n) \tag{10}$$

Another important and frequently used implicit copula is the Student-t copula, which has the advantage of taking into account tail dependence in the dependence structure.
Let $(X_1,\ldots,X_n)$ a vector of standard normal random variables with correlation matrix Σ, $\mu = (\mu_1,\ldots,\mu_n)$ and $\xi_\nu \sim \chi^2_\nu$ independent from the $\{X_i\}_{i=1,\ldots,n}$. Define

$$U_i := t_\nu\left(\mu_i + \frac{\sqrt{\nu}}{\sqrt{\xi_\nu}} X_i\right) \tag{11}$$

with t_ν denoting the cumulative distribution function of the univariate Student-t distribution with ν degrees of freedom.
Then, the distribution function $C_{\nu,\Sigma}(u_1,\ldots,u_n)$ of the $\{U_i\}_{i=1,\ldots,n}$ is the Student-t copula with ν degrees of freedom associated to the correlation matrix Σ.
If $t^n_{\nu,\Sigma}$ denotes the n-dimensional Student-t distribution, we have

$$C_{\nu,\Sigma}^{t}(u_1,\ldots,u_n) = t^n_{\nu,\Sigma}(t_\nu^{-1}(u_1),\ldots,t_\nu^{-1}(u_n)) \tag{12}$$

We can write the Student-t copula as

$$C_{\nu,\Sigma}(u_1,\ldots,u_n) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_n)} \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\pi\nu)^n \det(\Sigma)}} \left(1 + \frac{1}{\nu}(x-\mu)^{\mathrm{T}} \sum^{-1} (x-\mu)\right)^{-\frac{\nu+n}{2}} \mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{13}$$

We will discuss how to simulate these copulas in the context of basket time-to-default modelling in further sections.

2.1.4 Explicit copulas

The explicit copulas have a closed-form expressions for the joint distribution function which satisfy the initial criteria mentioned in the beginning of our section. An important class of explicit copulas is the Archimedean copulas and we will briefly describe three copulas derived from this class.

The Archimedean copulas can be constructed the following way

$$C(u_1, u_2, \ldots, u_n) = \psi^{-1}\left(\psi(u_1) + \psi(u_2) + \ldots + \psi(u_n)\right) \tag{14}$$

The function $\psi : [0,1] \to [0,\infty)$ is called the generator of the copula and the following properties hold :
1. ψ is convex
2. ψ is decreasing
3. $\psi(1) = 0$. If, moreover, $\psi(0) = \infty$, then ψ is called a strict generator and the copula is called a strict copula.

The Clayton copulas are built by choosing $\psi(t) = \frac{t^{-\theta}-1}{\theta}$ with $\theta \in [-1,\infty)\backslash\{0\}$ Then, the 2-dimensional Clayton copula reads

$$C_\theta^{Cl}(u_1, u_2) = \max\left(u_1^{-\theta} + u_2^{-\theta} - 1\ , 0\right)^{-1/\theta} \tag{15}$$

In particular, if $\theta > 0$, ψ is a strict generator and we have

$$C_\theta^{Cl}(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta} \tag{16}$$

The Gumbel copula use the function $\psi = (-ln(t))^\theta$ as its generator. The two-dimensional Gumbel copula reads

$$C_\theta^{\mathrm{Gu}}(u_1, u_2) = \exp\left(-\left((-\log u_1)^\theta + (-\log u_2)^\theta\right)^{1/\theta}\right) \tag{17}$$

with $\theta \in [1,\infty)$.

The last explicit copula we present in this section is the Frank copula, which strict generator is $\psi(t) = -\frac{e^{-\theta t}-1}{e^{-\theta}-1}$. The two-dimensional Frank copula reads

$$C_\theta(u_1, u_2) = -\frac{1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right) \tag{18}$$

with $\theta \in \mathbb{R}\backslash\{0\}$.

2.1.5 Kendall's tau

The Kendall's tau is a measure of dependence between two random variables X and Y. Denote (X', Y') an independent copies of the vector (X, Y). The pair of vectors (X, Y) and (X', Y') are said to be concordant if

$$(X - X')(Y - Y') > 0 \tag{19}$$

and discordant if

$$(X - X')(Y - Y') < 0 \tag{20}$$

In order to measure the dependency between X and Y, Kendall's tau evaluates the difference between their probabilities of concordance and discordance.
Let X and Y two random variables and (X', Y') an independent copy of (X, Y). Their Kendall's tau is given by

$$\tau(X,Y) = \mathbb{P}\{(X-X')(Y-Y') > 0\} - \mathbb{P}\{(X-X')(Y-Y') < 0\} \tag{21}$$

If (X, Y) is a vector of continuous random variables with copula C, their Kendall's tau can also be written

$$\tau(X,Y) = 4\iint_{[0,1]^2} C(u,v)dC(u,v) - 1 \tag{22}$$

Kendall's tau also has interesting properties.
The first one is its invariance to strictly increasing transformations of X and Y.
Moreover, there exists a link between its value and the dependence parameters of the corresponding copula. For instance, in the case of a Student-t or a Gaussian copula with correlation matrix $\Sigma = (\rho_{ij})_{i=1,\dots,n;j=1,\dots,n}$, we have the relation

$$\tau(X_i, X_j) = \frac{2}{\pi} arcsin(\rho_{ij}) \tag{23}$$

In particular, this property allows us to estimate the correlation coefficients from an estimator of the Kendall's tau.

2.1.6 Tail dependence

Consider again a vector of two random variables (X, Y). The tail dependence between X and Y is a measure of the likelihood of extreme movements in both X and Y. Large positive joint movements is qualified as an upper tail dependence and large negative joint movements is qualified as a lower tail dependence.
In order to measure their upper tail dependence between X and Y, we evaluate the probability that Y is in the upper tail of its cumulative distribution function F_Y knowing that X is in the upper tail of its cumulative distribution function F_X. Analogously, we measure their lower tail dependence by evaluating the probability that Y is in the lower tail of F_Y knowing that X is in the lower tail of F_X. More formally, the coefficient of upper tail dependence reads

$$\lambda_{\mathrm{U}}(X,Y) = \lim_{q \to 1-} \mathbb{P}\left[Y > F_Y^{-1}(q) \mid X > F_X^{-1}(q)\right] \tag{24}$$

and the coefficient of lower tail dependence reads

$$\lambda_L(X_1, X_2) = \lim_{q \to 0+} \mathbb{P}\left[Y \leq F_Y^{-1}(q) \mid X \leq F_X^{-1}(q)\right] \tag{25}$$

X and Y are said to be upper tail dependent with parameter λ_U if $\lambda_U > 0$ and asymptotically independent if $\lambda_U = 0$, and lower tail dependent with parameter λ_L if $\lambda_L > 0$.

λ_U and λ_L can also be expressed using the copula C of (X, Y). We have

$$\lambda_U = \lim_{q \to 1-} \frac{1 - 2q + C(q,q)}{1 - q} \tag{26}$$

and

$$\lambda_L = \lim_{q \to 0+} \frac{C(q,q)}{q} \tag{27}$$

2.1.7 Summary of the characteristics of the main copulas

In this section, we will list the main characteristics of the copulas we have seen in the previous sections by highlighting particularly their tail dependence and how to vary their parameters so that we obtain a fundamental copula.

Copula	Independence	Comonotonicity	Countermonotonicity
Gaussian	$\rho = 0$	$\rho \to 1$	$\rho \to -1$
Student-t	$\nu \to \infty$ and $\rho \to 0$	$\nu \to \infty$ and $\rho \to 1$	$\nu \to \infty$ and $\rho \to -1$
Clayton	$\theta \to 0$	$\theta \to \infty$	$\theta = -1$
Gumbel	$\theta = 1$	$\theta \to \infty$	No negative dependence
Frank	$\theta \to 0$	$\theta \to \infty$	$\theta \to -\infty$

Copula	Upper tail dependence	Lower tail dependence
Gaussian	No	No
Student-t	Yes and $\lambda_U = t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right) > 0$	Yes and $\lambda_L = \lambda_U$
Clayton	No	Yes if $\theta > 0$ and $\lambda_U = 2^{-1/\theta} > 0$
Gumbel	Yes and $\lambda_U = 2 - 2^{1/\theta} > 0$	No
Frank	No	No

2.2 Linear Regression Methods

Linear regression analysis is a supervised learning method which objective is to establish a relationship between a variable $y \in \mathbb{R}$ which we want to predict from an input $x = \left(x_1, \ldots, x_p\right)^T \in \mathbb{R}^p$. Linear regressions assume a linear relationship between y and x, e.g. y is modeled as $\theta_0 + \theta_1 x_1 + ... + \theta_p x_p$, with $\left(\theta_1, \ldots, \theta_p\right)$ the parameters of the prediction function.
Let us consider a general linear regression problem in which we observe $n \geqslant 1$ pairs (x_i, y_i), with $y_i \in \mathbb{R}$ and $x_i = \left(x_{i,0}, x_{i,1}, \ldots, x_{i,p}\right) \in \mathbb{R}^p$. These pairs (x_i, y_i) form our "training set".

We can also formulate the problem using matrix notations. Denote

$$X = \begin{pmatrix} x_1^T \\ \vdots \\ x_n^T \end{pmatrix} = \begin{pmatrix} x_{1,0} & \dots & x_{1,p} \\ \vdots & & \vdots \\ x_{n,0} & \dots & x_{n,p} \end{pmatrix} \in \mathbb{R}^{n\times(p+1)}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n \tag{28}$$

We can rewrite this relationship as :

$$\mathbf{Y} = X\boldsymbol{\theta} + \boldsymbol{\varepsilon}$$

- Y represents a vector of size n of observed values often called labels.
- X is a matrix of dimension n by p with n equal to the number of observations and p equal to the number of features or explanatory variables.
- ϵ is the error term of the regression. It captures all the other features which could explain Y and that are not modeled in X.

2.2.1 Ordinary Least-squares

First, we can think about the ordinary least square method which is a method used to find the unknown parameters in a linear regression model and based on minimizing the sum of the squares of the differences between the label and the prediction made by the linear model.

$$L(\theta) = \frac{1}{2}||Y - X\theta||_2^2$$

$$\hat{\theta} = \text{argmin}_{\theta\in\mathbb{R}^p} L(\theta)$$

Assume the following hypothesis :

- No colinearity of the explanatory variables X which means that that there is not an explanatory variable that can be expressed as a linear combination of the other explanatory variables
- The errors ϵ are independent.
- The explanatory variables are exogenous which means that they are not correlated with the error term ϵ
- The error term ϵ is homoscedastic, in other words the variance of the error term is constant.

Under these assumptions, the Gauss Markov Theorem states that the Ordinary Least Squares estimator is optimal in the class of linear unbiased estimators.
Hence this estimator has the lowest sampling variance within the class of linear unbiased estimators. We can find an explicit expression for $\hat{\theta}$.

$$\min_{\theta} L(\theta) = \min_{\theta} \frac{1}{2}\|Y - X\theta\|_2^2$$

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{1}{2}\frac{\partial}{\partial \theta}(Y - X\theta)^T(Y - X\theta) = 0$$

$$X^\top Y = \left(X^\top X\right)\theta$$

$$\hat{\theta} = \left(X^\top X\right)^{-1} X^\top Y$$

Note that the matrix formulation of the problem gives us an easy geometric interpretation of the OLS estimator : $\hat{Y} = X\hat{\theta}$ is the closest point to Y in the linear subspace $span(\mathrm{X}) \subset \mathbb{R}^n$.
We can also find the solution of the OLS problem geometrically. $\mathbb{R}^n$ being a Hilbert space and $\mathrm{span}(X)$ a linear subspace of $\mathbb{R}^n$ (thus closed and convex), we can apply the Hilbert projection theorem, which implies that $\hat{Y}$ is unique and $Y - \hat{Y}$ is orthogonal to $\mathrm{span}(X)$. In particular, the second point implies that $X^T(Y - \hat{Y}) = 0$ and thus $X^T X\hat{\theta} = X^T Y$. When $X^T X$ is invertible, we obtain the closed form formula $\hat{\theta} = (X^T X)^{-1} X^T Y$.
If we are in the case where we have only one explanatory variable, the problem simplifies to:

$$\mathrm{Argmin}_{\theta_0,\theta_1} \sum_{i=1}^{n} (y_i - \theta_1 x_i - \theta_0)^2$$

And then by setting the gradient equal to 0 we find that :

$$\hat{\theta}_1 = \frac{\sum x_i \sum y_i - n\sum x_i y_i}{(\sum x_i)^2 - n\sum x_i^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\hat{\theta}_0 = \frac{\sum y_i - \hat{\theta}_1 \sum x_i}{n} = \bar{y} - \hat{\theta}_1 \bar{x}$$

Then we have to assess the quality of the prediction. We will use the determination coefficient R^2 defined as :

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

The closer the determination coefficient is closed to 1, the better the prediction is.

The OLS has the advantage of having easily interpretable parameters and a low computational time.

However, it also has some drawbacks such as the fact that a large number of explanatory variables could lead to over-fitting.
This issue could be tackle with shrinkage method such as LASSO regression or RIDGE regression where we add a regularization term to the loss function, which we decide not to discuss here as these methods haven't been used during the project.

3 Description of the data

This brief section aims at describing the instruments we will work with along this thesis. We have already introduced the main principles of credit derivatives in the Introduction and made the distinction between single-name credit derivatives and multi-name credit derivatives, and we will describe our quantities of interest more formally.
The most liquid credit derivatives are the credit default swaps (CDS), which is a basic protection contract. The CDS contracts ensure protection against a default from an entity. Assume that two companies A (which will be the protection buyer) and B (which will be the protection seller) enter into a credit default swap with maturity T agree that if a reference entity C defaults (i.e. there is a credit event). The premium leg consists of a stream of periodic payments from the protection buyer A to the protection seller B, which lasts until the maturity time T of the contract or until the reference entity C defaults. We illustrate these cash flows in the figure below :

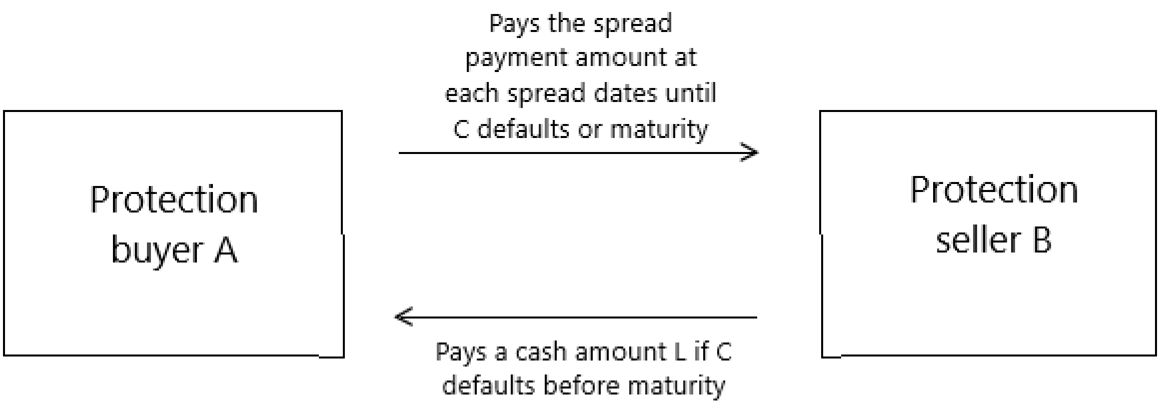

Figure 1: Credit Default Swap between a protection seller A and a protection buyer B on the default of a reference entity C

The credit events which count as a default from C and trigger the payment from the protection seller, as stated by the International Swaps and Derivatives Association (ISDA), are the following :
. Bankruptcy of C
. Payment default from C
. Debt restructuring of C, i.e. a reorganization of its debts to reinforce them
. Obligation acceleration, i.e. when C has to pay a debt earlier as a result of a default or another similar event
. Repudiation, i.e. if C rejects the validity of its contract and no longer respects it

3.1 Default baskets pricing

A default basket is very similar to a credit default swap. It is also a contract involving a protection leg and a premium leg, but the difference is that the credit event which triggers the payment from the protection leg is the n-th default in a basket of N issuers. Let's index the issuers in the basket with $i = 1,\dots,N$ and note $i(n)$ the index of the n-th issuer to default. Then, assuming a contract with unit nominal \$1, the amount L paid by the protection seller to the protection buyer after the n-th default is equal to $(1 - R_{i(n)})$.

Let's denote τ_n^* the time of the n-th default, T the maturity time of the contract, D the stochastic discount factor such that $D(t_1,t_2)$ represents the discount factor between dates t_1 and t_2 (depending on the interest rates) and assume a contract with unit nominal. Then at time τ_n^* the present value of the default leg is given by :

$$\mathrm{PV}_{\mathrm{Def}}(t) = \begin{cases} D\left(t,\tau_n^*\right)\left(1 - R_{i(n)}\right) & \text{if } \tau_n^* \leq T \\ 0 & \text{if } \tau_n^* > T \end{cases} \tag{29}$$

If we denote t_0 as the beginning date of the contract, $(t_1,\dots,t_p)$ the payment dates for the premium leg, $\alpha_i = t_i - t_{i-1}$ and $(S_1,\dots,S_p)$ the corresponding spread payments values, the present value of the premium leg at time t is given by :

$$\mathrm{PV}_{\mathrm{Prem}}(t) = \begin{cases} \sum_{i=1}^{m} S_i\alpha_i D\left(t,t_i\right) + \mathrm{PV}_{\mathrm{Accr}}\left(t\right) & \text{if } t_m \leq \tau_n^* < t_{m+1} \\ \sum_{i=1}^{p} S_i\alpha_i D\left(t,t_i\right) & \text{if } \tau_n^* \geq t_p \end{cases} \tag{30}$$

with $\mathrm{PV}_{\mathrm{Accr}}\left(t\right)$ corresponding to a compensation to the protection seller for the protection given from the last payment date t_m before the default occur to τ_n^*. We have

$$\mathrm{PV}_{\mathrm{Accr}}(t) = \begin{cases} S_m D(t,\tau_n^*)(\tau_n^* - t_m) & \text{if } t_m \leq \tau_n^* < t_{m+1} \\ 0 & \text{if } \tau_n^* \geq t_p \end{cases} \tag{31}$$

In order to price a n-th to default basket at a time t, we compute the expectation of the difference between the present values (at time t) of the premium leg and the default leg. From the protection seller point of view, this leads to

$$V_{NTD}(t) = E[\mathrm{PV}_{\mathrm{Prem}}(t) - \mathrm{PV}_{\mathrm{Def}}(t)]$$

We will see in the next section how this quantity can be computed numerically in the case of a First-to-default basket.

3.2 Default dependency

The data we are going to use consist of basket containing several issuers and we have to take into account their default dependency, that is the likelihood that two or more issuers default approximately at the same time, which may occur for many reasons, among which the geographical sector in which the issuers are located which may cause them to be impacted the same way by external events or their sector of activity, such that a default from one company can lead another one to default as well. The aim of this section is to illustrate this relationship between the default dependency of the issuers in the basket and the probability that a default occurs.
For clarity purposes, we will use the same approach as in [1], chapter 12, by taking the example of a simple basket containing only two names A and B. Let's denote

their corresponding default times respectively τ_A and τ_B. The probability of default by time T of A and B are defined as

$$P_A(T) = \mathbb{E}[\mathbb{1}_{\tau_A \leq T}] \tag{32}$$

and

$$P_B(T) = \mathbb{E}[\mathbb{1}_{\tau_B \leq T}] \tag{33}$$

Both these individual default probabilities can be found in practice by extracting them from market. Then, we can define their joint default probability, i.e. the probability that both A and B default before time T, as

$$P_{AB}(T) = \mathbb{E}[\mathbb{1}_{\tau_A \leq T}\mathbb{1}_{\tau_B \leq T}] \tag{34}$$

The First-to-default default probability at time T of the basket depends not only on the probability of default of A and B but also on their joint default probability. Indeed, in the case of a 2 names basket, the credit event triggering the payment from the protection leg is the default of A, B or A and B, i.e.

$$\begin{aligned} P_{FTD}(T) &= 1 - \mathbb{E}[\mathbb{1}_{\tau_A \geq T}\mathbb{1}_{\tau_B \geq T}] \\ &= 1 - \mathbb{E}[(1 - \mathbb{1}_{\tau_A \leq T})(1 - \mathbb{1}_{\tau_B \leq T})] \\ &= P_A(T) + P_B(T) - P_{AB}(T) \end{aligned} \tag{35}$$

And the probability of a second-to-default event is simply the probability that both A and B default by time T, i.e.

$$P_{STD}(T) = P_{AB}(T) \tag{36}$$

We can think of three particular scenario which show how the first-to-default and second-to-default probability can be affected by the joint default probability of A and B, which we illustrate through the following diagrams :

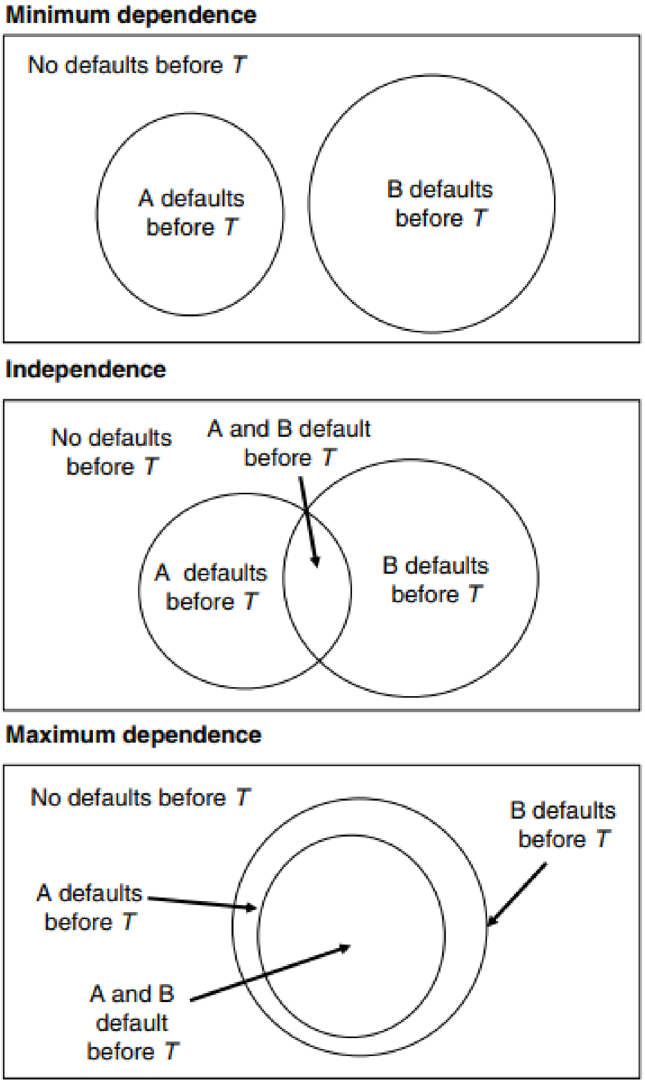

Figure 2: Minimum dependence, independence and maximum dependence scenarios for the joint distribution of A and B. Figure taken from [1]

1. Minimum dependence scenario : $P_{AB}(T) = max(P_A(T) + P_B(T) - 1, 0)$, which leads to two cases : either $P_A(T) + P_B(T) \geq 1$ in which case

$$\begin{aligned} P_{FTD}(T) &= 1 \\ P_{STD}(T) &= P_A(T) + P_B(T) - 1 \end{aligned} \tag{37}$$

or $P_A(T) + P_B(T) < 1$, and

$$\begin{aligned} P_{FTD}(T) &= P_A(T) + P_B(T) \\ P_{STD}(T) &= 0 \end{aligned} \tag{38}$$

2. Independence : $P_{AB}(T) = P_A(T)P_B(T)$. Then,

$$\begin{aligned} P_{FTD}(T) &= P_A(T) + P_B(T) - P_A(T)P_B(T) \\ P_{STD}(T) &= P_A(T)P_B(T) \end{aligned}$$

3. Maximum dependence : $P_{AB}(T) = \min[P_A(T), P_B(T)]$. And we have

$$P_{FTD}(T) = \max(P_A(T), P_B(T))$$
$$P_{STD}(T) = \min(P_A(T), P_B(T))$$

We note here that the First-to-default default probability decreases from $P_A(T)+P_B(T)$ to $max(P_A(T), P_B(T))$ and that the Second-to-default probability increases from 0 to $min(P_A(T), P_B(T))$ as the dependency between A and B increases.
We define the FTD spread as the fixed spread which has to be paid by the protection buyer to the protection seller in order to equalize the expectation of the present value of the premium leg and the default leg. Its value depends on the number of issuers, their individual spreads, the maturity of the First-to-default basket and the default correlation between the issuers. We will see in section 5.1.2 that the First-to-default spread's behaviour is similar to the First-to-default default probabilities in terms of boundaries it lies between the minimum of the individual issuers spread and the sum of all the individual spreads, and also in terms of sensitivity to default correlation.

4 Models

4.1 The Gaussian Latent Variable Model

4.1.1 Standard model

The Gaussian Latent Variable Model is the most famous and the most frequently used model when dealing with correlation products.

Let's introduce the random variable A, which we call a "latent variable", associated to a credit issuer. Here, this latent variable is associated to a time-to-default τ in so far as we say that a default occurs by time T if the value A is less than a threshold $C(T)$.
In particular, assuming a standard normal distribution for A we have

$$Pr(\tau < T) = Pr(A < C(T)) = \Phi^{-1}(C(T)) \tag{39}$$

Φ being the cumulative distribution function of the standard normal distribution.

Hence, the values of A entirely determine the values of τ, and we have $A = C(\tau)$. Denote $Q(T)$ the issuer's probability of survival past time T. The relation above implies that

$$Q(T) = 1 - \Phi(C(T)) \tag{40}$$

and in particular,

$$Q(\tau) = 1 - \Phi(A) \tag{41}$$

Using this relation and the property that if $X \sim N(0,1)$, then $\Phi(X) \sim U(0,1)$ (and therefore $1 - \Phi(X) \sim U(0,1)$), which we will use as it is more computationally efficient to draw from a uniform distribution than a gaussian distribution, we can simulate the default time τ using a Monte Carlo approach the following way :

Algorithm 4.1

- For each simulation $k = 1, 2, ..., M$, draw $u_k \sim U(0,1)$ independently
- For each $k = 1, \ldots, M$, solve $\tau^k = Q^{-1}(u_k)$
- Average over all the simulations to return $\tau = \frac{1}{M}\sum_{k=1}^{M} \tau^k$

4.1.2 The One-factor Copula Model

The previous model is convenient in the case of modelling a single-name credit derivative, but in the context of our project, we will be working with baskets of issuers, which requires a model that takes into account the correlation factor between the issuers.

Let's assume we are working with a basket of N issuers, and note A_i the latent variable associated to the i-th issuer. In this model, we make the assumption of a Gaussian copula between the default times $\{\tau_i\}_{i=1,...,N}$. In order to construct a Gaussian copula between the $\{\tau_i\}_{i=1,...,N}$ we can first notice that if $p_i(T)$ is the default probability of the issuer i by time T, then $U_i := p_i(\tau_i)$ is a uniform random variable. Thus, if $C(u_1,...,u_n)$ is a multivariate uniform distribution and $U_1,...,U_N$ are multivariate uniform with distribution C the multivariate distribution for the default time can be :

$$\tau_1 := p_1^{-1}(U_1), \ldots, \tau_M := p_N^{-1}(U_N)$$

In order to have a Gaussian dependence structure between the default times $\{\tau_i\}_{i=1,...,N}$, we assume that

$$[U_1, \ldots, U_N] = [\Phi(A_1), \cdots, \Phi(A_N)]$$

Hence we construct the $\{\tau_i\}_{i=1,...,N}$ from the $\{A_i\}_{i=1,...,N}$ via

$$\tau_i = p_i^{-1}(\Phi(A_i))$$

which implies in particular that

$$\{\tau_i < T\} = \{A_i < \Phi^{-1}(p_i(T))\}$$

Finally, we have

$$\begin{aligned} \forall i,j, \Pr\left(\tau_i < t_i, \tau_j < t_j\right) &= \Pr\left(A_i < \Phi^{-1}(p_i(t_i)), A_j < \Phi^{-1}(p_i(t_j)); \rho_{ij}\right) \\ &= \Phi_{2,\rho_{ij}}\left(\Phi^{-1}(p_i(t_i)), \Phi^{-1}\left(p_i(t_j)\right)\right) \\ &= C\left(p_i(t_i), p_j\left(t_j\right)\right) \end{aligned} \tag{42}$$

with $\rho_{ij} = \mathrm{Corr}(A_i, A_j)$.
So the default copula is the Gaussian bi-variate cumulative distribution function

$$C^{Def}_{\rho_{i,j}}\left(u_i, u_j\right) = \Phi_2\left(\Phi^{-1}(u_i), \Phi^{-1}\left(u_j\right), \rho_{ij}\right) \tag{43}$$

Moreover, the survival copula reads

$$\begin{aligned} C^{Surv}_{\rho_{i,j}}\left(u_i, u_j\right) &= \Phi_2\left(\Phi^{-1}(1-u_i), \Phi^{-1}\left(1-u_j\right), \rho_{ij}\right) \\ &= \Phi_2\left(-\Phi^{-1}(u_i), -\Phi^{-1}\left(u_j\right), \rho_{ij}\right) \\ &= \Phi_2\left(\Phi^{-1}(u_i), \Phi^{-1}\left(u_j\right), \rho_{ij}\right) \text{ by symmetry of the Gaussian distribution} \\ &= C^{Def}_{\rho_{i,j}}\left(u_i, u_j\right) \end{aligned} \tag{44}$$

In the d-factor Gaussian Model, we assume that the correlation matrix $\Sigma = (\rho_{ij})_{i=1,...,N;j=1,...,N}$ has a "factorial" structure, which means that the $\{A_i\}_{i=1,...,N}$ can be written as

- $A_i = \sum_{k=1}^{d} \beta_{ik} Z_k + \left(\sqrt{1 - \sum_{k=1}^{d} \beta_{ik}^2}\right) Y_i$

- $Z = (Z_1, ..., Z_d)$ vector of independent $N(0,1)$ random variables

- $Y = (Y_1, ..., Y_n)$ vector of independent $N(0,1)$ random variables independent from Z

Under these assumptions, the default probability of the issuer i by time T conditionally on Z is given by

$$\begin{aligned} Pr(\tau_i < T \mid Z) &= Pr(A_i < \Phi^{-1}(p_i(T)) \\ &= Pr\left(Y_i < \frac{\Phi^{-1}(p_i(T)) - \sum_{k=1}^{d} \beta_{ik} Z_k}{\sqrt{1 - \sum_{k=1}^{d} \beta_{ik}^2}} \mid Z\right) \\ &= \Phi\left(\frac{\Phi^{-1}(p_i(T)) - \sum_{k=1}^{d} \beta_{ik} Z_k}{\sqrt{1 - \sum_{k=1}^{d} \beta_{ik}^2}}\right) \end{aligned} \tag{45}$$

We can then easily derive the formula for the probability of default of the basket past a time T as the random variables $\{A_i\}_{i=1,...,N}$ are independent conditionally on $Z = \{Z_k\}_{k=1,...,d}$ which enables us to write

$$\begin{aligned} Pr(\text{No default in the basket by time T} \mid M) &= Pr(\forall i,\ \tau_i > T \mid Z) \\ &= \prod_{i=1}^{N} (1 - Pr(A_i < \Phi^{-1}(p_i(T))) \\ &= \prod_{i=1}^{N} \left(1 - \Phi\left(\frac{\Phi^{-1}(p_i(T)) - \sum_{k=1}^{d} \beta_{ik} Z_k}{\sqrt{1 - \sum_{k=1}^{d} \beta_{ik}^2}}\right)\right) \end{aligned} \tag{46}$$

Finally, in order to obtain the First-to-default survival probability of the basket, we simply have to integrate over all the possible values of the vector Z. If we denote $\{\tau_i^*\}_{i=1,...,n}$ the sorted vector of default times such that $\tau^{(*)}{}_i$ is the i-th time-to-default, we have :

$$\begin{aligned} Pr(\tau_1^* > T) &= E\left[\prod_{i=1}^{N} Pr(\tau_i > T \mid Z)\right] \\ &= \int_{\mathbb{R}^d} \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{\Phi^{-1}(p_i(T)) - \sum_{k=1}^{d} \beta_{ik} Z_k}{\sqrt{1 - \sum_{k=1}^{d} \beta_{ik}^2}}\right)\right] \times \prod_{i=1}^{N} \phi(Z_k) dZ_k \end{aligned} \tag{47}$$

In the One-Factor Gaussian Model, we make the assumption that the correlation matrix Σ has a 1 factor structure. In this case, A_i satisfies the relation

$$A_i = \beta_i Z + \sqrt{1-\beta_i^2}\, Y_i.$$

Z and Y_i being standard normal random variables. Z is called a systematic factor (or the system factor) and is common for every issuer in the basket, Y_i is called a idiosyncratic factor which depends on the issuer i and β_i is the factor loading associated to A_i.
In particular, we have $\mathrm{corr}\left(A_i, A_j\right) = \sqrt{\beta_i \beta_j}$. Assuming, furthermore, that the basket's correlation matrix is flat , e.g. $\beta_i = \beta$ for all i, these relations become

$$A_i = \beta Z + \sqrt{1-\beta^2}\, Y_i \text{ and } \mathrm{corr}\left(A_i, A_j\right) = \beta^2 = \rho \text{ for all } i.$$

Under these hypothesis, we can derive the formula of the probability of survival of the basket past a time T, conditionally on Z as a particular case of the d-factor Gaussian Model with $d = 1$.

First, let's find the probability of survival of the basket past T. Remembering that a default occurs at time T if $A_i \leq C_i(T) = \Phi^{-1}(p_i(T))$, we have

$$\begin{aligned} Pr(\text{No default in the basket by time T} \mid Z) &= Pr(\forall i,\ \tau_i > T \mid M) \\ &= \prod_{i=1}^{N} [1 - Pr(A_i < \Phi^{-1}(p_i(T) \mid Z)] \\ &= \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{\Phi^{-1}(p_i(T)) - \beta Z}{\sqrt{1-\beta^2}}\right)\right] \end{aligned} \tag{48}$$

And we can get the first-to-default survival probability of our basket by integrating over all the possible values of Z. As $Z \sim N(0,1)$, we have

$$Pr(\tau_1^* > T) = \int_Z \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{\Phi^{-1}(p_i(T)) - \beta Z}{\sqrt{1-\beta^2}}\right)\right] \phi(Z) dZ \tag{49}$$

The One-Factor approach requires a simple numerical integration which speeds up the computations a lot. However, there are limitations to this model as it doesn't capture the whole correlation structure of the issuers, in particular when there is a high correlation between the issuers but a low correlation between the sectors.

4.1.3 Conditional Hazard Rates

A popular way of modelling default times is to use Default Intensity Models. In Intensity Models, every default time τ_i is assumed to follow an exponential law. Recall that specifying the marginal distributions of the default times is enough to determine

their joint distribution as we only have to choose a copula dependence structure to link them.
A hazard rate (or intensity) is a strictly positive stochastic process $\lambda : t \rightarrow \lambda_t$. We define the corresponding hazard function (or cumulated intensity) as $\Lambda : t \rightarrow \int_0^t \lambda_s ds$. The default time τ we want to model is then defined as $\tau = \Lambda^{-1}(\xi)$ with ξ an exponential random variable independent from λ. Note that Λ^{-1} is defined as $\forall t \in \mathbb{R}^+$, $\lambda(t) > 0$ so Λ is a strictly increasing function of t. In practice, the random variable ξ can be built from a uniform random variable $U \sim Unif[0,1]$ on which we apply the inverse cumulative distribution function of an exponential distribution, i.e. by taking $\xi = -ln(1-U)$. In particular, the probability of survival past a time t reads

$$\mathbb{Q}(\tau > T) = \mathbb{Q}(\Lambda^{-1}(\xi) > t) = \mathbb{Q}(\xi > \Lambda(t)) \tag{50}$$

By the tower property, the independence of ξ and Λ and the fact that $\mathbb{Q}(\xi > t) = e^{-t}$ for any $t \in \mathbb{R}$ we have

$$\mathbb{Q}(\xi > \Lambda(t)) = \mathbb{E}[\mathbb{Q}(\xi > \Lambda(t)) \mid \Lambda(t)] = \mathbb{E}[e^{-\Lambda(t)}] = \mathbb{E}[e^{-\int_0^t \lambda_s ds}] \tag{51}$$

Let's get back to our One-factor model. Recall that the default probability by time T of the i-th issuer conditionally on Z reads

$$p_i(T \mid Z) = 1 - Q_i(T \mid Z) = \Phi\left(\frac{C_i(T) - \beta_i Z}{\sqrt{1-\beta_i^2}}\right) \tag{52}$$

We assume a flat and deterministic conditional hazard rate for the default times $\{\tau_i\}_{i=1,\dots,n}$. As we have seen above, this assumption implies

$$\text{for any } i = 1,\dots,n,\;\; Q_i(T \mid Z) = exp(-\lambda_i(T \mid Z)T) \tag{53}$$

And using (48) we have

$$\lambda_i(T \mid Z) = -\frac{1}{T}\ln\Phi\left(\frac{\beta_i Z - C_i(T)}{\sqrt{1-\beta_i^2}}\right) \tag{54}$$

The value of $\lambda_i(T \mid Z)$ is directly linked to the i-th issuer's probability of default (from (49)) as large values of $\lambda_i(T \mid Z)$ means a high probability of default and low values a low probability of default. It is therefore interesting to take a look at the distribution of the conditional hazard rates implied by the One-factor Gaussian copula model and its changes with respect to the factor loading β_i. We plot the conditional hazard rate distribution below :

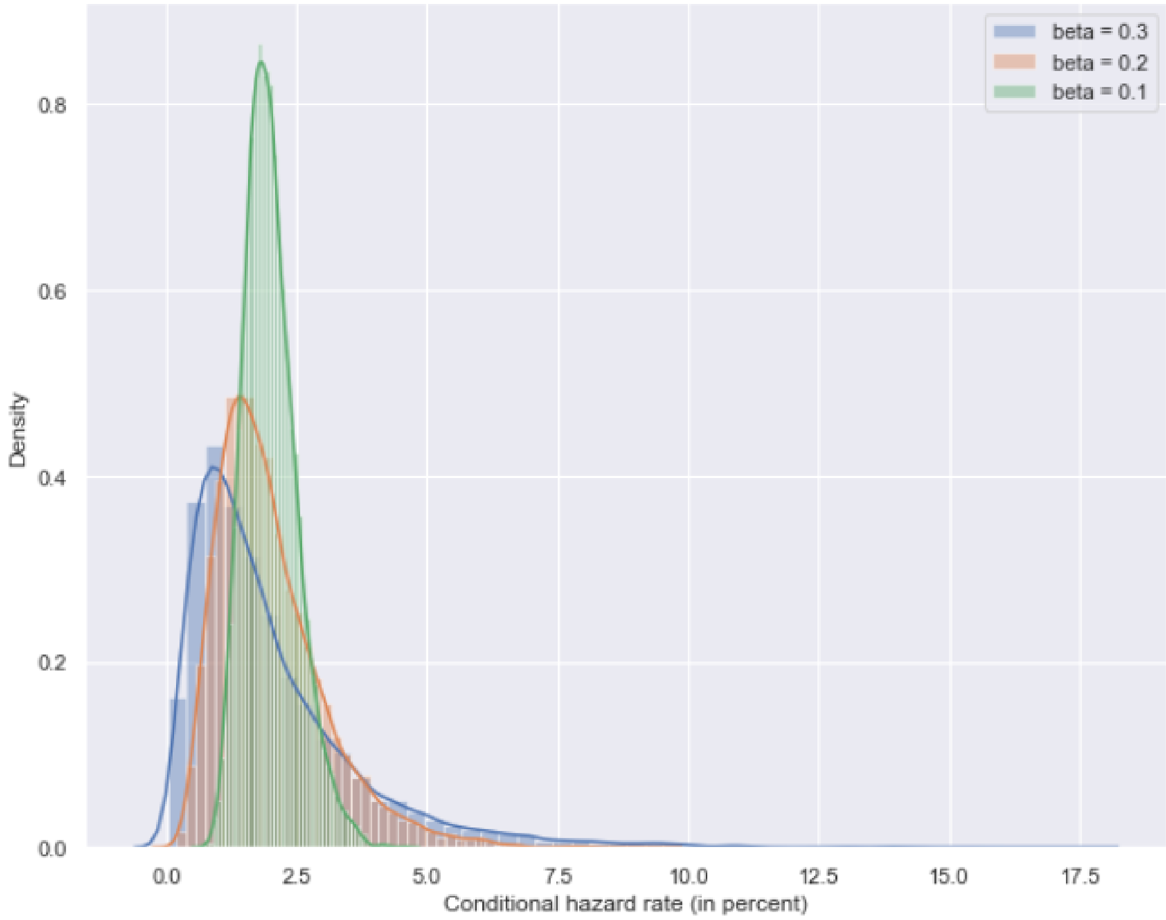

Figure 3: Conditional hazard rate distribution for $\beta = 0.3, 0.2, 0.1$ and a 1-year time horizon

We observe that both the skewness and the variance of the density falls as β decreases and when β tends to zero, the conditional hazard rate becomes an unconditional hazard rate and the defaults are independent.

4.2 Student-t copula model

In the previous scenario, we assumed a Gaussian copula between the credits. However, another copula commonly used in this context is the Student-t copula which has the advantage of taking into account tail dependence between the credits.
Let's first present the Student-t distribution with more details. A random variable X follows a Student-t distribution with $\nu \geq 0$ degrees of freedom if it can be written as

$$X = \sqrt{\frac{\nu}{\xi_\nu}} Z$$

with $Z \sim N(0,1)$ and $\xi_\nu \sim \chi^2_\nu$ independent from Z.
The probability density function of the Student-t distribution is given by

$$f_\nu(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \tag{55}$$

Γ denoting the Gamma function $\Gamma : x \mapsto \int_0^{+\infty} t^{x-1}e^{-t}\mathrm{d}t$. Let's plot the Student-t probability density function for different values of ν :

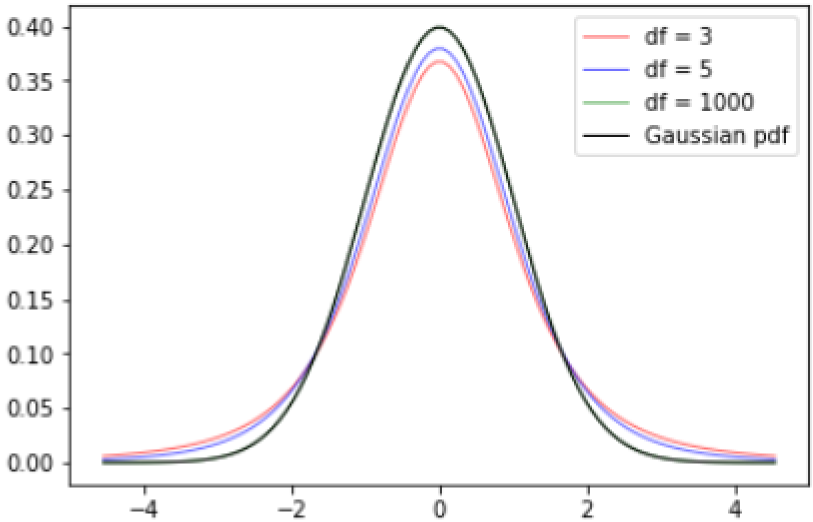

Figure 4: Probability density function of a Student-t distribution for $\nu = 3, 5, 1000$ and Gaussian probability density function

The Student-t distribution has the following properties :

- $\mathrm{E}(X)$ is not defined if $\nu = 1$ and is equal to 0 otherwise, which can be observed in the figure above by symmetry about zero.
- $\mathrm{Var}(X) = \frac{\nu}{\nu-2}$ for $\nu > 2$
- As $\nu \to \infty$ the Student-t distribution converges to a standard normal distribution, which we can see on the graph by observing the pdf of the Student-t distribution with $\nu = 1000$ which coincides with the Standard normal pdf.

We can simulate the default times of the credits in our basket using the following algorithms :

Algorithm 4.2.1 : Generating correlated gaussian variables
Assume $\Sigma \in \mathbb{R}^{nxn}$ is the correlation matrix of our basket.

- Find the Cholesky decomposition of Σ, e.g. the lower diagonal matrix A such that $\Sigma = AA^T$
- Simulate $z_1, z_2, ..., z_n \sim N(0,1)$ using for instance the Box-Muller method described in the appendix of the section.
- Set $x = Az$ with $z = (z_1, \ldots, z_n)^T$. Thus, the x_i are correlated gaussian variables.

Algorithm 4.2.2 : Simulating default times with Monte-Carlo

- Generate M vectors of correlated gaussian variables x^k for $k = 1, 2, ..., M$ with $z^k = \left(x_1^k, \ldots, x_n^k\right)^T$.
- Generate independently M random variables $\xi^k \sim \chi_\nu^2$ for $k = 1, 2, ..., M$ from uniform random numbers using the inverse cumulative distribution function method (if $U \sim Unif[0,1]$ and F is a continuous CDF, then the random variable $X := F^{-1}(U)$ has the distribution F).
- For each $k = 1, ..., M$ and for each $i = 1, ..., n$, generate the Student-t distributed random variable $y_i^k = \sqrt{\frac{\nu}{\xi_\nu^k}} x_i^k$.
- Map each correlated random variables y_i^k to the uniform random variable $u_i^k = t_\nu(y_i^k)$.
- Solve $\tau_i^k = Q^{-1}(u_i^k)$ for each $k = 1, ..., M$ and for each $i = 1, ..., n$
- For each $i = 1, ..., n$, average over all the simulation to return $\tau_i = \frac{1}{M} \sum_{k=1}^{M} \tau_i^k$

Let's use the same notations as in the previous section. The latent variable A_i associated to the issuer i verifies in this model the relation :

- $A_i = \left(\beta_i Z + \sqrt{1-\beta_i^2} Y_i\right)\sqrt{\frac{v}{\xi_v}}$
- $M \sim N(0,1)$
- $Y_i \sim N(0,1)$ independent from Z
- $\xi_v \sim \chi_v^2$

As before, the i-th basket defaults before time T if $A_i \leq C_i(T)$, with $C_i(T) = t_v^{-1}(1-Q_i(T))$, t_v^{-1} being the inverse cumulative distribution function of the Student-t distribution and $Q_i(T)$ the probability of survival of issuer i past time T. Thus, this condition reads

$$\left(\beta_i Z + \sqrt{1-\beta_i^2} Y_i\right)\sqrt{\frac{v}{\xi_v}} \leq C_i(t) \;\Leftrightarrow\; Y_i \leq \frac{C_i(t)\sqrt{\xi_v/v} - \beta_i Z}{\sqrt{1-\beta_i^2}}$$

Thus, the probability of default of the issuer i by time T conditionally on Z and ξ_v is given by

$$p_i(t \mid Z, \xi_v) = \Phi\left(\frac{C_i(t)\sqrt{\xi_v/v} - \beta_i Z}{\sqrt{1-\beta_i^2}}\right).$$

By independence of the events $\{A_i \leq C_i(T)\}$ conditionally on M and ξ_v, we have

$$\begin{aligned} Pr(\text{ No default in the basket by time } T \mid Z, \xi_v) &= Pr(\forall i,\, \tau_i > T \mid Z, \xi_v \\ &= \prod_{i=1}^{N} \left[1 - p_i(T \mid Z, \xi_v)\right] \\ &= \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{C_i(T)\sqrt{\xi_v/v} - \beta_i Z}{\sqrt{1-\beta_i^2}}\right)\right] \end{aligned} \tag{56}$$

In order to get the first-to-default survival probability of the basket, we integrate over all the possible values of Z and ξ_v. Using the fact that M and ξ_v are independent, and thus their joint density function is equal to the product of their marginal density function, we obtain that

$$Pr\left(\tau_1^* > T\right) = \int_Z \int_{\xi_v} \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{C_i(T)\sqrt{\xi_v/v} - \beta_i Z}{\sqrt{1-\beta_i^2}}\right)\right] \phi(Z) f_v(\xi_v) dZ d\xi_v \tag{57}$$

Under the assumption of a flat correlation matrix for the basket, e.g. if $\forall i,\ \rho_i = \rho$, the first-to-default probability at time T reads

$$Pr(\tau_1^*) > T = \int_Z \int_{\xi_\nu} \prod_{i=1}^{N} \left[1 - \Phi\left(\frac{C_i(T)\sqrt{\xi_\nu/v} - \beta Z}{\sqrt{1-\beta^2}} \right) \right] \phi(Z) f_\nu(\xi_\nu) dZ d\xi_\nu \quad (58)$$

with $\beta = \sqrt{\rho}$.
Even though pricing First-to-defaults under the assumption of a Student-t Copula has the advantage of taking into account tail dependence between the credits, the major issue is the computational efficiency of this approach, which requires the inversion of a Student-t cumulative distribution function as well as integrating over both the values of a standard normal distributed variable and a chi-2 distributed variable, instead of a simple integral in the One-Factor Gaussian Copula model. Hence, the Student-t copula model is not recommended in the case of a basket containing an important number of names.

4.3 Monte Carlo pricing of a First-to-default basket

Although the main purpose of this project is to study the correlation structure of First-to-default baskets rather than pricing them, we will introduce in this section a simple Monte Carlo approach to price an FTD from the simulation of the default times $\{\tau_i\}_{i=1,...,n}$ of the issuers, for which we will assume a One-Factor Gaussian copula dependence structure. The Monte Carlo pricing approach require a fast implementation, as it scales in time as $O(M)$, M being the number of simulations, but its convergence rate is only $O(1/\sqrt{M})$. In order to increase the accuracy of the Monte-Carlo estimate, a classic procedure is to use variance reduction techniques, which reduce the constant factor of the $O(1/\sqrt{(M)})$ (see [5]). Amongst these methods, we find Importance sampling and Antithetic variables, discussed in the appendix, which can be used easily in a one-dimensional approach.
First, we need to simulate default times $\{\tau_i\}_{i=1,...,n}$ for which we have assumed a One-Factor Gaussian copula. Using the same notations as in section 4.1, we proceed the following way :

Algorithm 4.3.1 : Simulating default times in the One-Factor Gaussian copula model

- Compute every thresholds $C_i(T) = \Phi^{-1}(p_i(T))$
- Use Box-Muller algorithm to generate M standard normal random variables Z^k and MN standard normal random variables Y_i^k, for $k = 1,\dots,M$, for $i = 1,\dots,N$
- Compute $A_i^k = \beta_i Z^k + \sqrt{1-\beta_i^2} Y_i^k$ for $k = 1,\dots,M$, for $i = 1,\dots,N$
- Compute $u_i^k = 1 - \Phi(A_i^k)$ for $k = 1,\dots,M$, for $i = 1,\dots,N$
- Compute $\tau_i^k = Q_i^{-1}(u_i^p)$ for $k = 1,\dots,M$, for $i = 1,\dots,N$

The remaining steps of the Monte Carlo pricing of the First-to-default require to compute the discounted payoff for each of the M sets of default times $\tau_1^k, \ldots, \tau_N^k$
As we have seen in section 3.1, the First-to-default discounted payoff from the point of view of the protection seller at time t reads

$$\begin{aligned} &\mathrm{PV}_{\mathrm{Prem}}(t) - \mathrm{PV}_{\mathrm{Def}}(t) = \\ &(\sum_{i=1}^{m} S_i \alpha_i D(t, t_i) + S_m D(t, \tau_1^*)(\tau_1^* - t_m))\mathbb{1}_{t_m \le \tau_1^* < t_{m+1}} - D(t, \tau_1^*)\left(1 - R_{i(1)}\right)\mathbb{1}_{\tau_1^* < T} \end{aligned} \tag{59}$$

using the same notations as in section 3.1.
We only have to average the discounted payoffs for each sets of default times to have an estimate of the First-to-default price.

4.4 Section 4 Appendix

4.4.1 Box-Muller

An efficient way to sample from a given distribution is to use the inverse CDF method when the inverse cumulative distribution function is known. However, there isn't a close form expression for the inverse cumulative distribution function of a normal distribution. A fast way of sampling from a standard normal distribution is the Box-Muller method, which provides a procedure for generating two independent standard normal random variables from independent uniform random variables.
Before jumping to the formula, let's describe a bit the motivation behind this method. Recall that the probability density function of a standard normal distribution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{60}$$

So if (X, Y) is a pair of independent standard normal distributed random variables, their joint density is given by

$$f(x, y) = f(x)f(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \tag{61}$$

The idea is to convert the cartesian coordinates x and y into polar coordinates with radius r and angle θ, which can be done using the relations

$$\begin{aligned} r^2 &= x^2 + y^2 \\ x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \tag{62}$$

Therefore, we consider the vector of random variables (R, Θ) such that

$$\begin{aligned} X &= R\cos\Theta \\ Y &= R\sin\Theta \end{aligned} \tag{63}$$

with $\Theta \in [0, 2\pi]$.
Θ being uniformly distributed in $[0, 2\pi]$ we can sample it from

$$\Theta = 2\pi U_1 \tag{64}$$

with $U_1 \sim Unif[0,1]$.

Moreover, R^2 is the sum of the squares of two independent standard normal variables, which is known to follow a χ^2 distribution with 2 degrees of freedom. It can be shown that the χ^2_2 distribution is equivalent to an exponential distribution with parameter $\lambda = \frac{1}{2}$. It is therefore possible to sample R^2 using the inverse CDF method for an exponential distribution, which gives

$$R^2 = -\frac{log(U_2)}{\lambda} = -2log(U_2) \tag{65}$$

with $U_2 \sim Unif[0,1]$.
Finally, the Muller-Box method can be described with the following steps :

Algorithm 4.4.1 : Box-Muller Method

- Draw independently $U_1, U_2 \sim Unif[0,1]$
- Compute $\Theta = 2\pi U_1$ and $R = \sqrt{-2log(U_2)}$
- The random variables X and Y defined as

$$\begin{aligned} X = R\cos\Theta &= \sqrt{-2log(U_2)}\cos(2\pi U_1) \\ Y = R\sin\Theta &= \sqrt{-2log(U_2)}\sin(2\pi U_1) \end{aligned} \tag{66}$$

  are independent standard normal random variables.

We plot below the results of a simulation using $N = 10000$ samples.

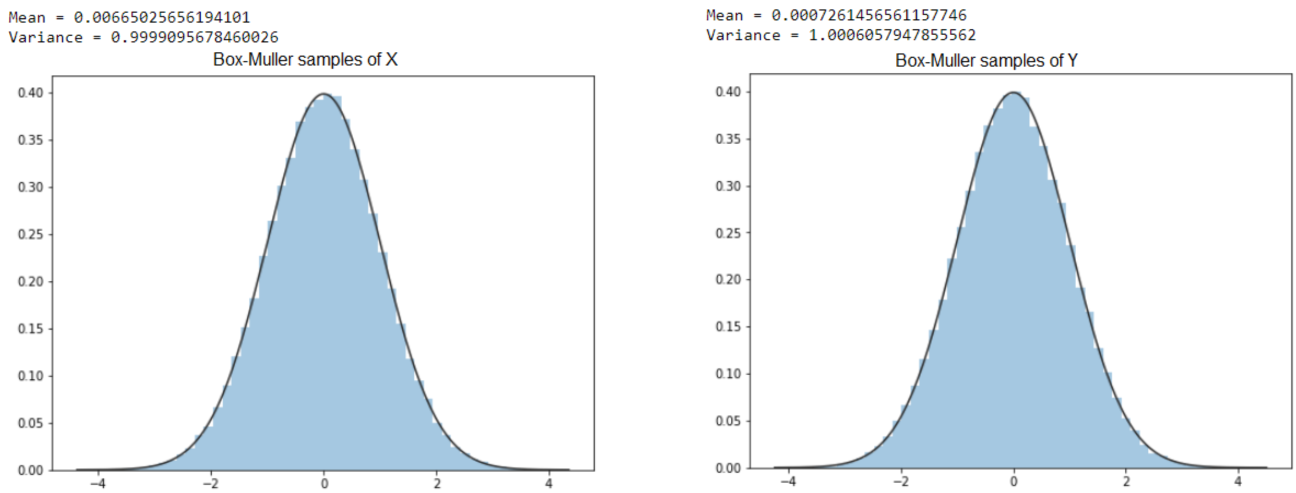

Figure 5: Box-Muller samples for X and Y with N = 10000 simulations

4.4.2 Importance sampling

The purpose of Importance sampling is to reduce the variance of a Monte-Carlo Method by changing the probability measure used to sample its paths.
Let X a d-dimensional random vector with probability density function $f_X^{\mathbb{P}}$ with respect to the probability measure $\mathbb{P}$ and assume we want to estimate

$$I(g) = \mathbb{E}^{\mathbb{P}}[g(X)] = \int g(x) f_X^{\mathbb{P}}(x) dx \tag{67}$$

with $g : \mathbb{R}^d \to \mathbb{R}$.
Let $X_1, \ldots, X_n$ be independent random variables sampled from $f_X^{\mathbb{P}}$. A classic Monte Carlo estimator of $I(g)$ is given by

$$\hat{I}_{\mathbb{P}}(g) = \frac{1}{n} \sum_{i=1}^{n} g(X_i) \tag{68}$$

Let $\mathbb{Q}$ be another probability measure equivalent to $\mathbb{P}$ and denote $f_X^{\mathbb{Q}}$ the strictly positive density function of X with respect to $\mathbb{Q}$ (it actually has to be strictly positive only on the support of $f_X^{\mathbb{P}}$, i.e. on the set $\{x \in \mathbb{R}^d : f_X^{\mathbb{P}}(x) \neq 0\}$). Then we have

$$\begin{aligned} I(g) = \mathbb{E}^{\mathbb{P}}[g(X)] &= \int g(x) f_X^{\mathbb{P}}(x) dx \\ &= \int g(x) \frac{f_X^{\mathbb{P}}(x)}{f_X^{\mathbb{Q}}(x)} f_X^{\mathbb{Q}}(x) dx \\ &= \mathbb{E}^{\mathbb{Q}}\left[g(X) \frac{f_X^{\mathbb{P}}(X)}{f_X^{\mathbb{Q}}(X)} \right] \end{aligned} \tag{69}$$

Now let $X_1, \ldots, X_n$ be independent random variables sampled from $f_X^{\mathbb{Q}}$. We can define a new estimator of $I(g)$ as

$$\hat{I}_{\mathbb{Q}}(g) = \frac{1}{n} \sum_{i=1}^{n} g(X_i) \frac{f_X^{\mathbb{P}}(X_i)}{f_X^{\mathbb{Q}}(X_i)} \tag{70}$$

It follows from the linearity of E that both the estimators $\hat{I}_{\mathbb{P}}(g)$ and $\hat{I}_{\mathbb{Q}}(g)$ are unbiased, i.e. $\mathbb{E}^{\mathbb{P}}[\hat{I}_{\mathbb{P}}(g)] = I(g)$ and $\mathbb{E}^{\mathbb{Q}}[\hat{I}_{\mathbb{Q}}(g)] = I(g)$. Hence, comparing the variance of these two estimators is equivalent to comparing their second moments. Note that we have

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[\left(g(X) \frac{f_X^{\mathbb{P}}(X)}{f_X^{\mathbb{Q}}(X)} \right)^2 \right] &= \int g(x) \frac{f_X^{\mathbb{P}}(x)}{f_X^{\mathbb{Q}}(x)} f_X^{\mathbb{P}}(x) dx \\ &= \mathbb{E}^{\mathbb{P}}\left[g(X)^2 \frac{f_X^{\mathbb{P}}(X)}{f_X^{\mathbb{Q}}(X)} \right] \end{aligned} \tag{71}$$

Depending on the choice of $\mathbb{Q}$, this quantity can be set lower than $\mathbb{E}^{\mathbb{P}}[g(X)^2]$. Note that if g is a nonnegative function, then $\forall x \in \mathbb{R}^d$, $g(x) f_X^{\mathbb{P}}(x) \geq 0$ and we can normalize this product to make it a probability density. Suppose we choose $\mathbb{Q}$ such that $f_X^{\mathbb{Q}}$ is equal to this density, then $\exists K \in \mathbb{R}$ s.t. $\forall x \in \mathbb{R}^d$, $f_X^{\mathbb{Q}}(x) = K g(x) f_X^{\mathbb{P}}(x)$. Then the variance of $\hat{I}_{\mathbb{Q}}(g)$ is equal to 0 as $\forall i = 1, \ldots, n, g(X_i) \frac{f_X^{\mathbb{P}}(X_i)}{f_X^{\mathbb{Q}}(X_i)} = K$ and the $\{X_i\}_{i=1,\ldots,n}$ are independent. However, normalizing $g(x) f_X^{\mathbb{P}}(x)$ require to know the value of its integral, which is precisely what we want to estimate. However, we can try to approximate a density proportional to $g(x) f_X^{\mathbb{P}}(x)$. Several methods exist to do so and we will briefly discuss two popular procedures : the Likelihood approach and the variance minimization method. Both these methods are said to be parametric importance sampling, in the sense that that we want to obtain an approximation of the optimal sampler among a parametric family of functions

$$Q = \{q_\theta \ : \theta \in \Theta\} \tag{72}$$

with $\Theta \subset \mathbb{R}^d$ and q_θ is a density function in $\mathbb{R}^d$. Denote $q^*(x) = \frac{g(x) f_X^{\mathbb{P}}(x)}{I(g)}$.

Approximation of the likelihood We define the likelihood function as

$$L(\theta) = \int \ln\left(\frac{q_\theta(x)}{q^*(x)}\right) q^*(x)dx \tag{73}$$

which can be rewritten as

$$L(\theta) = \frac{1}{I(g)} \int \ln(\frac{q_\theta(x)}{|g(x)f_X^{\mathbb{P}}(x)|})|g(x)f_X^{\mathbb{P}}(x)|dx - \ln(\frac{1}{I(g)})$$

Therefore, maximizing the likelihood function L is equivalent to maximizing the integral $\int \ln(\frac{q_\theta(x)}{|g(x)f_X^{\mathbb{P}}(x)|})|g(x)f_X^{\mathbb{P}}(x)|dx$. This can be done using the following algorithm :

Algorithm 4.4.2 : Likelihood importance sampling

- Sample independently $1 < m < n$ random variables $X_1, \ldots, X_m$ from the density $f_X^{\mathbb{P}}$ and compute

$$\hat{\theta}_m = argmax_{\theta\in\Theta} \frac{1}{m} \sum_{i=1}^{m} ln\left(\frac{q_\theta(X_i)}{|g(X_i)f_X^{\mathbb{P}}(X_i)|}\right)|g(X_i| \tag{74}$$

- Let $k_m = n - m$. Sample $Z_1, \ldots, Z_{k_m}$ independently from the density $q_{\hat{\theta_m}}$ and compute

$$\frac{1}{k_m} \sum_{i=1}^{k_m} \frac{g(Z_i)f_X^{\mathbb{P}}(Z_i)}{q_{\hat{\theta_m}}(Z_i)} \tag{75}$$

Approximation of the variance Denote

$$\begin{aligned} \theta^* \in argmin_{\theta\in\Theta} \mathbb{E}^{\mathbb{Q}}\left[\left(g(X)\frac{f_X^{\mathbb{P}}(X)}{q_\theta(X)}\right)^2\right] &= argmin_{\theta\in\Theta} \mathbb{E}^{\mathbb{P}}\left[g(X)^2 \frac{f_X^{\mathbb{P}}(X)}{q_\theta(X)}\right] \\ &= argmin_{\theta\in\Theta} \int \frac{(g(x)f_X^{\mathbb{P}}(x))^2}{q_\theta(x)} dx \end{aligned} \tag{76}$$

In practice, computing $\int \frac{(g(x)f_X^{\mathbb{P}}(x))^2}{q_\theta(x)}dx$ cannot be done as it involves the quantity $g(x)f_X^{\mathbb{P}}(x)$. Therefore, our goal is to estimate θ^* to return a "nearly-optimal" estimator q_{θ^*}. This can be done with the following procedure :

Algorithm 4.4.2

- Sample independently $1 < m < n$ random variables $X_1, \ldots, X_m$ from the density $f_X^{\mathbb{P}}$ and compute

$$\hat{\theta}_m = argmin_{\theta \in \Theta} \frac{1}{m} \sum_{i=1}^{m} \frac{g(X_i)^2 f_X^{\mathbb{P}}(X_i)}{q_\theta(X_i)} \tag{77}$$

- Let $k_m = n - m$. Sample $Z_1, \ldots, Z_{k_m}$ independently from the density $q_{\hat{\theta_m}}$ and compute

$$\frac{1}{k_m} \sum_{i=1}^{k_m} \frac{g(Z_i) f_X^{\mathbb{P}}(Z_i)}{q_{\hat{\theta_m}}(Z_i)} \tag{78}$$

5 Calibration results

5.1 Model's results

In this section, we present the numerical results that we have obtained using the models previously described. To do so, we will work with a First-to-default basket with 10 names. The main characteristics of the basket, such as the probability of survival and the individual spreads are known by a time horizon of 10 years. For clarity purpose, the market dates used for the figures in this section is a list of 10 elements corresponding to only one date per year. We assume that the correlation for the basket is flat and equal to 0.42. The figure below shows the individual probabilities of default from the issuers :

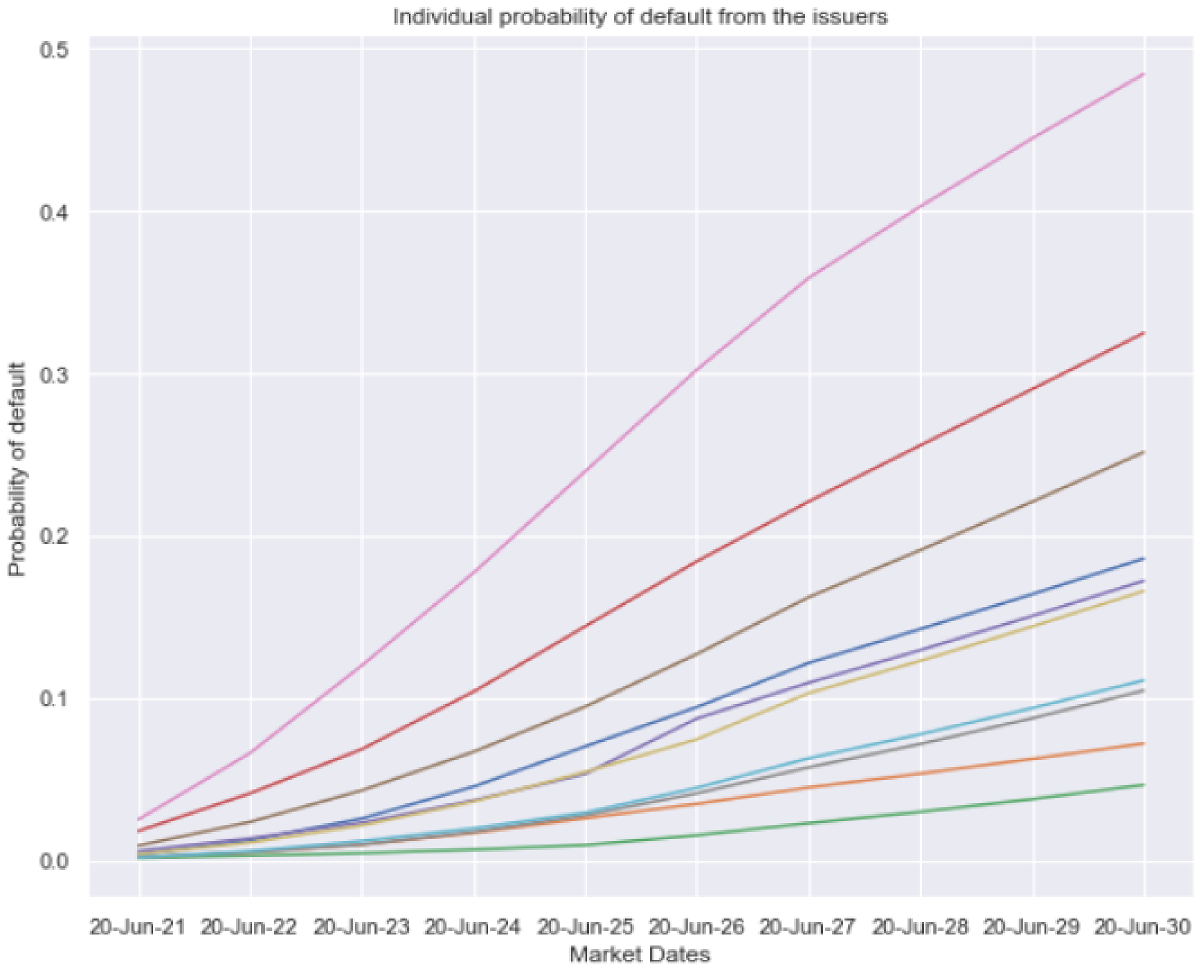

Figure 6: Individual probabilities of default from the issuers at each market date

5.1.1 First-to-default probability

The first thing we will present are the plots of the First-to-default probabilities computed in the Gaussian copula model and in the Student-t copula model. We will then analyze the sensitivity to First-to-default probabilities with respect to the correlation and see if we get the theoretical results we expect.
Let's focus on the First-to-default probabilities computed under the assumption of a One-Factor Gaussian copula for the basket. For the correlation value $\rho = 0.42$, the FTD probability's evolution with respect to time has the following behaviour :

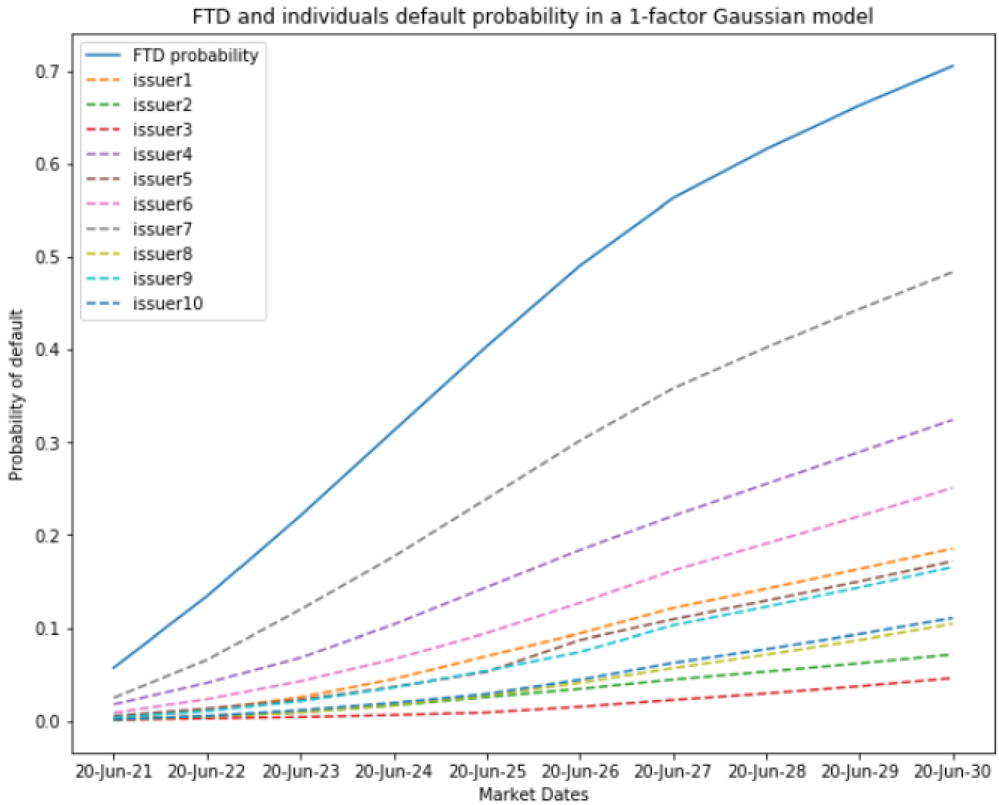

Figure 7: First-to-default and individual probabilities with $\rho = 0.42$ in the 1-Factor Gaussian Copula model

We can observe that the FTD probability is always higher than the maximum of the individual probabilities, as one may have expected. In order to assess the consistency of our implementation, recall what we have discussed in the section 3.2 for a basket containing 2 names. We have distinguished three different scenario concerning the dependence structure of the defaults in the basket : the minimum dependence scenario where $\rho = -1$, the independence scenario where $\rho = 0$ and the maximum dependence scenario where $\rho = 1$. Here, we assume that the correlation value is always positive, which eliminates the minimum dependence scenario. We have shown that the FTD probability tends to decrease when the dependence between the defaults (or in this case simply the correlation) increases, and in the 2 names basket case (A and B), the lower bound of the FTD probability is the maximum of the default probabilities between A and B (maximum dependence scenario) and its upper bound is equal to $P(A)+P(B)-P(A)P(B) = 1-(1-P(A))(1-P(B))$ (using the same notations as in the section 3.2). Then, by setting successively $\rho = 1$ and $\rho = 0$, we can expect that the FTD probability would be equal respectively to the maximum of all the individual default probabilities and to $1-\prod_{i=1}^{10}(1-p_i)$ (p_i being the probability of default from the issuer i at a given market date). We obtained the following results :

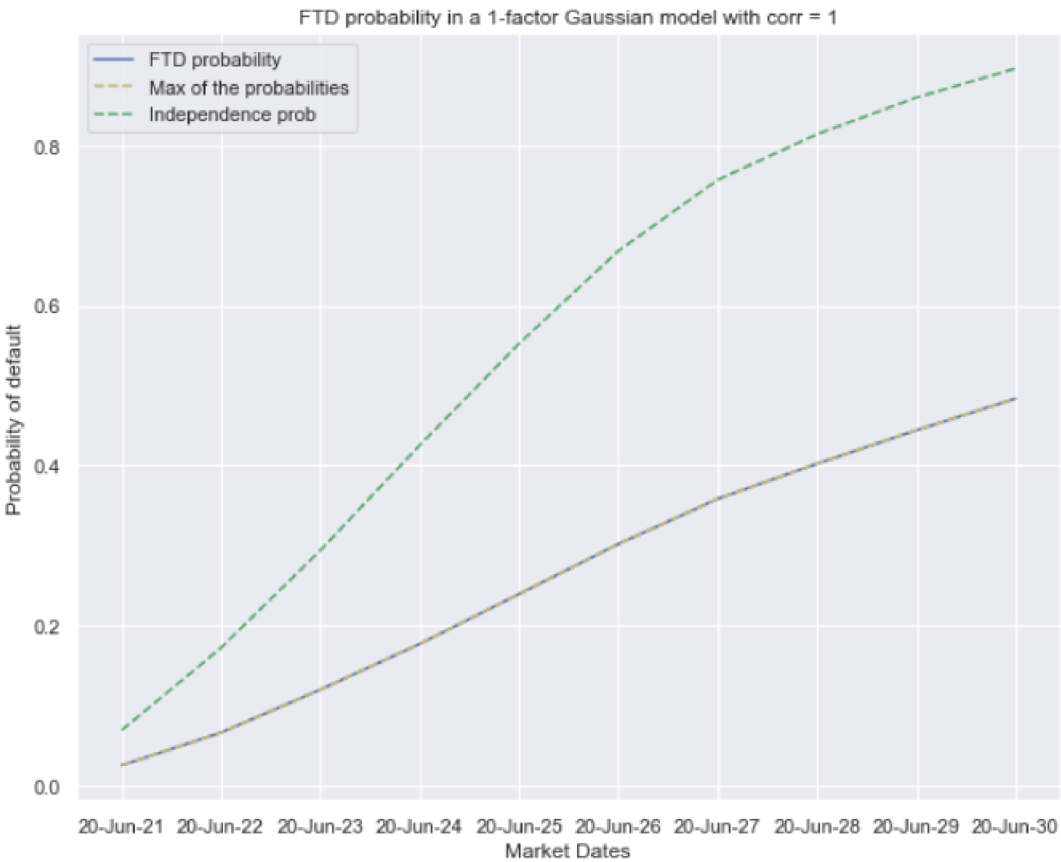

Figure 8: First-to-default probability in the maximum dependence scenario in the 1-Factor Gaussian Copula model

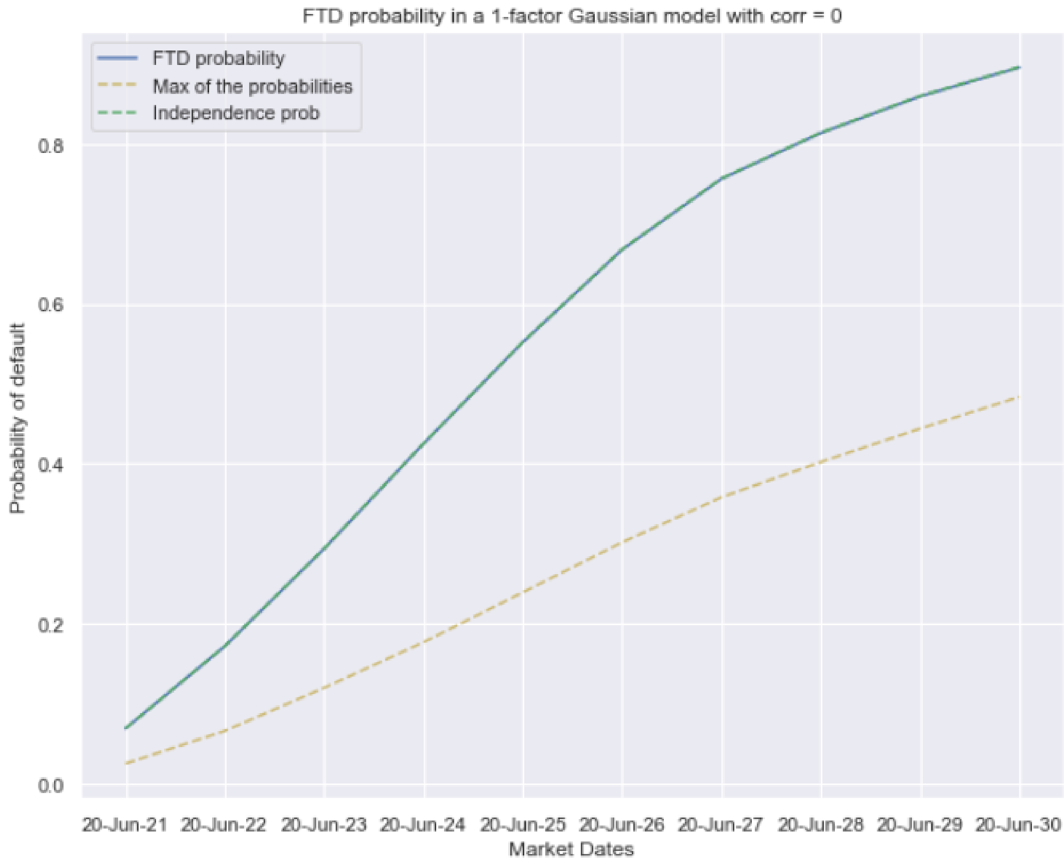

Figure 9: First-to-default probability in the independence scenario in the 1-Factor Gaussian Copula model

In the two graphs above, the FTD probability curve coincides with its boundaries as we were expecting.

Let's now see the results for the FTD probability computations when assuming a Student-t copula between the defaults. We are expecting the same behavior for the independence and the maximum dependence scenarios, but there is another parameter to vary in the Student-t copula model, which is the degrees of freedom. We will first plot the results for a correlation value $\rho = 0.42$ for the typical value of $\nu = 5$ and also for $\nu = 1000$. In this last case, we are expecting the FTD probabilities to be the same as in the Gaussian copula, as the Student-t distribution's limit when $\nu \to \infty$ is a standard normal distribution. The results are shown below :

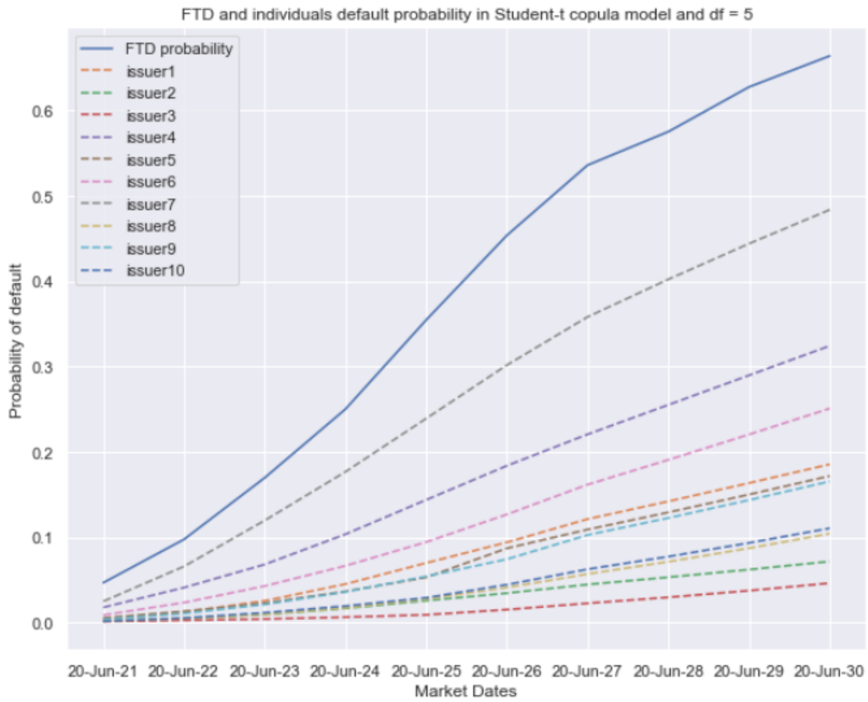

Figure 10: First-to-default and individual probabilities in the Student-t copula model with $\rho = 0.42$ and $\nu = 5$

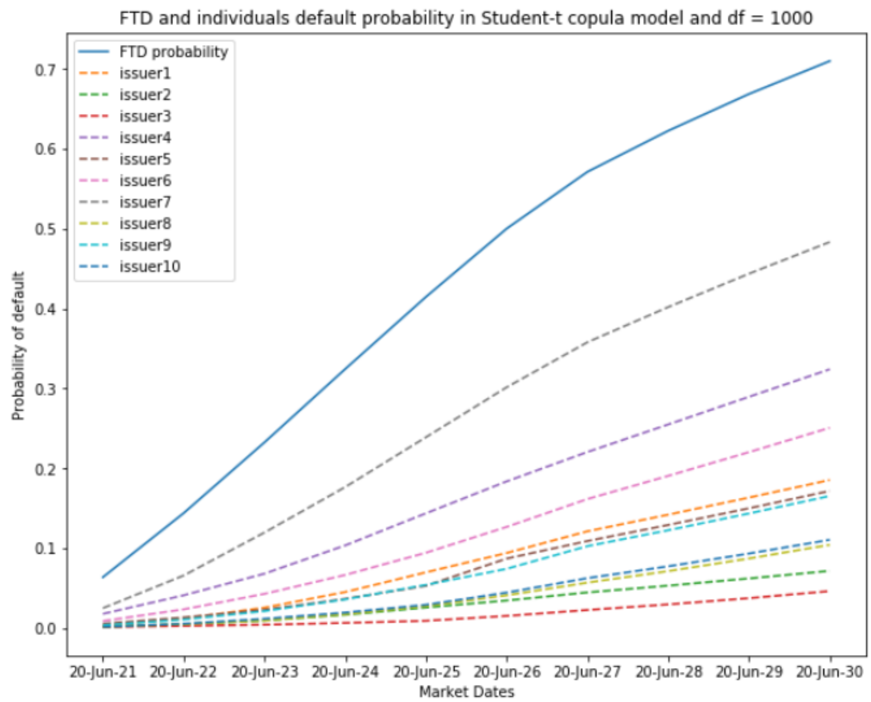

Figure 11: First-to-default and individual probabilities in the Student-t copula model with $\rho = 0.42$ and $\nu = 1000$

The second figure actually seem to coincide with the FTD probability in the One-Factor Gaussian copula model. In order to assess the similarity between the two graphs, we compute the mean absolute error (each time step corresponding to a market date). The latter being equal to 0.00886, we can conclude that this consistency check is valid.
Finally, let's check if we have the expected results for the maximum dependence and the independence scenarios. We do not have to check them in the case where $\nu = 1000$ as we have already shown that it was equivalent to a One-factor gaussian copula model, so we are going to choose $\nu = 5$ only.

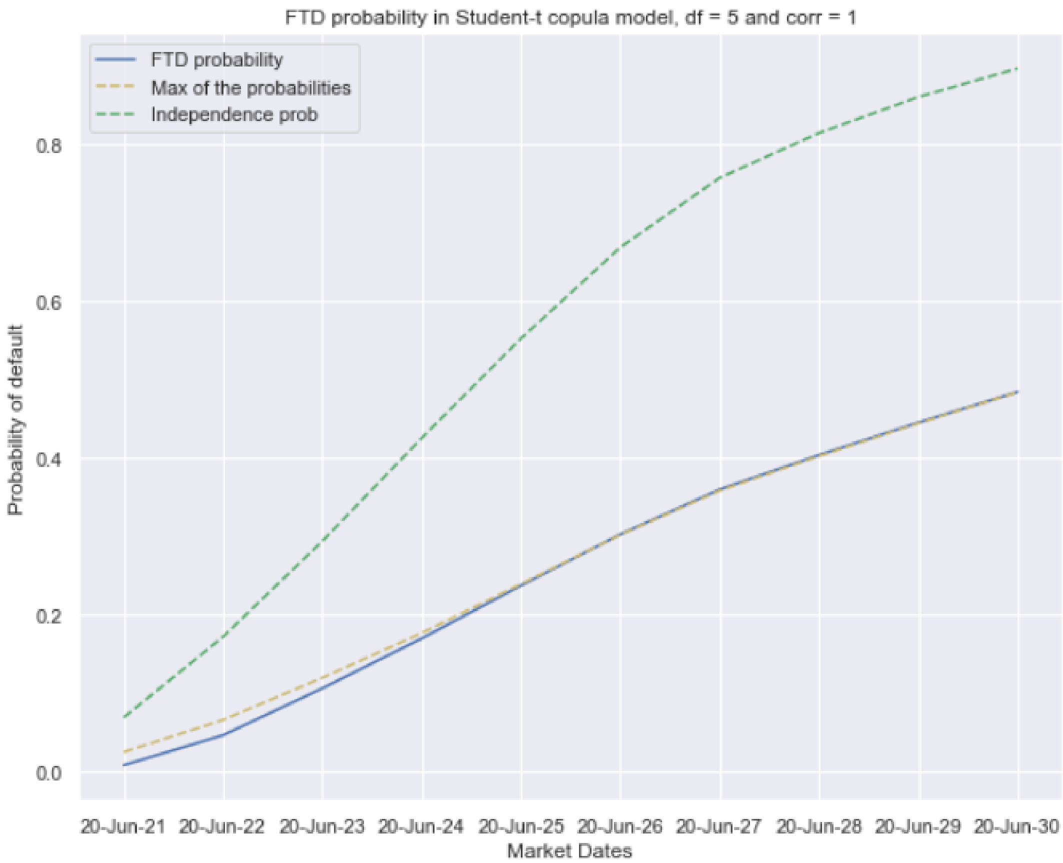

Figure 12: First-to-default probability in the maximum dependence scenario in the Student-t Copula model with $\nu = 5$

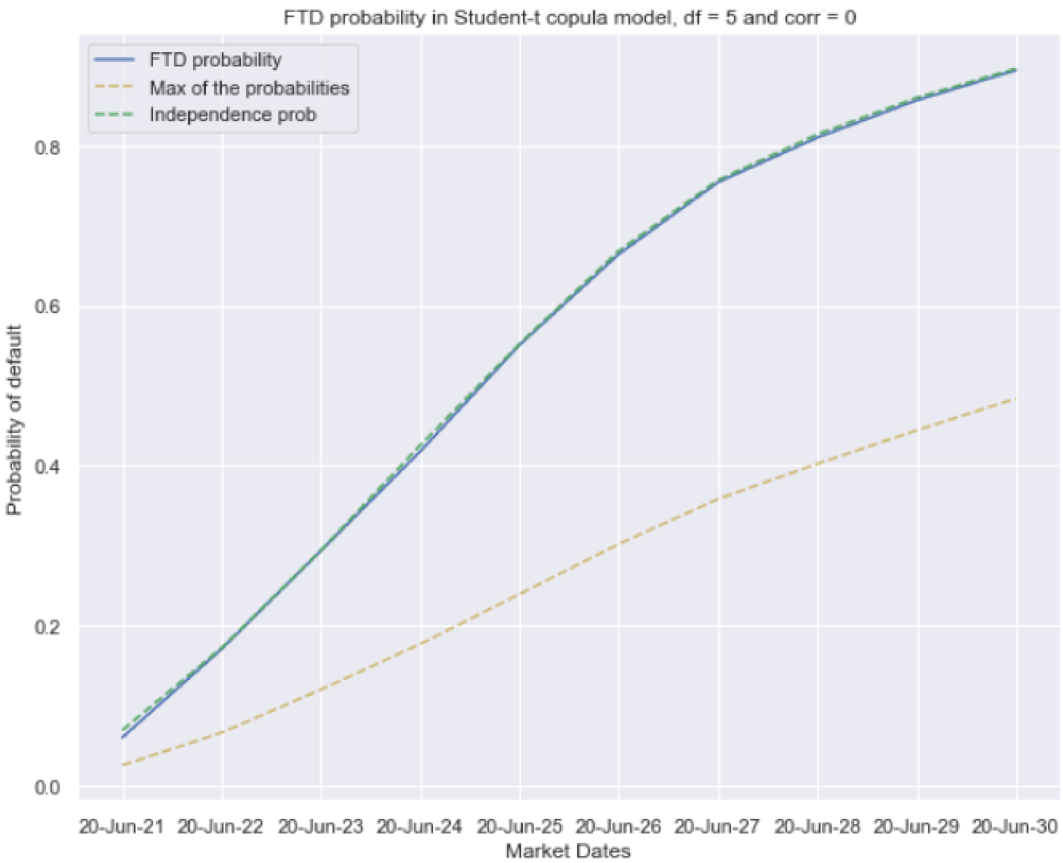

Figure 13: First-to-default probability in the independence scenario in the Student-t Copula model with $\nu = 5$

The behaviour is approximately what we were expecting, but the FTD probabilities do not coincide with the limit cases scenarios as accurately as in the One-Factor model. The slight errors between the curves may be due to the fact that the assumption of a Student-t model requires to compute numerically a double integral in order to return the FTD probability, which may cause additional approximations comparing to the simple integral in the One-Factor model.

5.1.2 FTD Spreads

The second quantity of interest we will study is the FTD spread, which we have to determine in order to price our First-to-default basket.
For clarity purposes, we show on the same figure the FTD spread computed under the assumption of a One-Factor Gaussian copula model and under the assumption of a Student-t copula model. We take the same correlation value $\rho = 0.42$ and we set $\nu = 5$. As the FTD default probabilities are higher in the Gaussian copula model, we also expect higher spread values (the more likely a default happens, the more the protection buyer has to pay the seller). We obtained the following result :

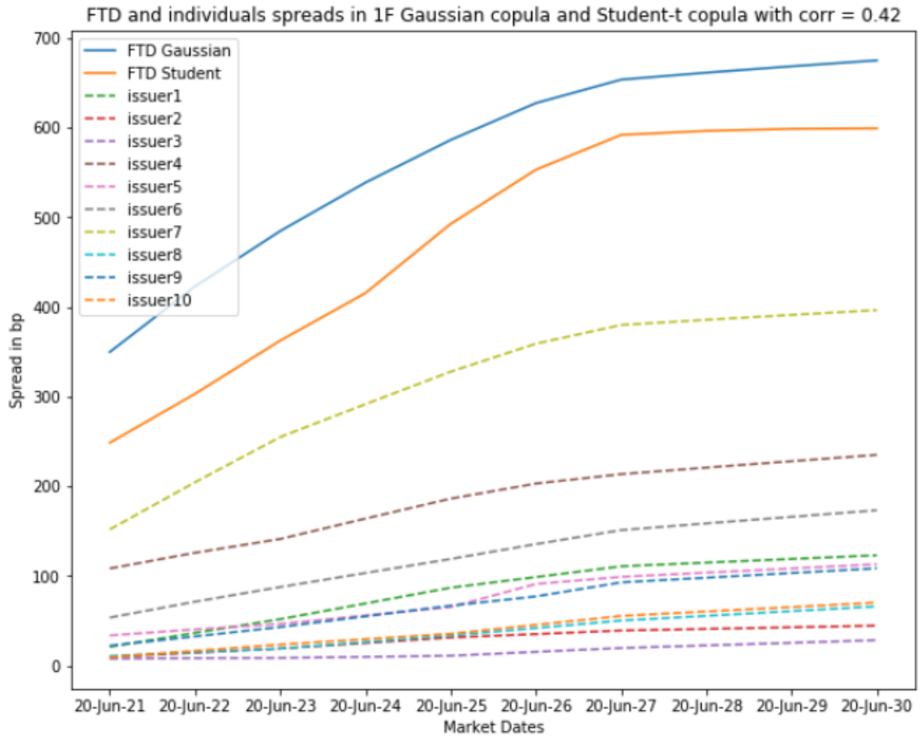

Figure 14: First-to-default spreads in 1F-Gaussian copula and Student-t copula models with $\nu = 5$ and $\rho = 0.42$

The only different input for the calculation of the FTD spread in the two models is the FTD default probabilities. As we have already shown the consistency of our implementation for the FTD default probabilities computations, and in particular that if we set $= 1000$ we had a mean absolute error of only 0.00886 compared to the Gaussian copula model, it is already proven that the FTD spread curve in the Student-t copula will coincide with the FTD spread curve in the Gaussian copula for large values of ν.

We will now evaluate the impact of the correlation over the FTD spread values. To do so, we make the assumption of a One Factor Gaussian copula model, and we will show the FTD spread curve evolution when we vary the correlation value ρ of our basket. As we have seen in section 3.2, the FTD default probability decreases when the dependence between the issuers increases. As the FTD spread is an increasing function of the FTD default probability, we also expect the FTD spread to be lower when the correlation increases. We chose to plot on a same figure the FTD spread curves for the values $\rho = 1, 0.8, 0.6, 0.42, 0.2, 0$:

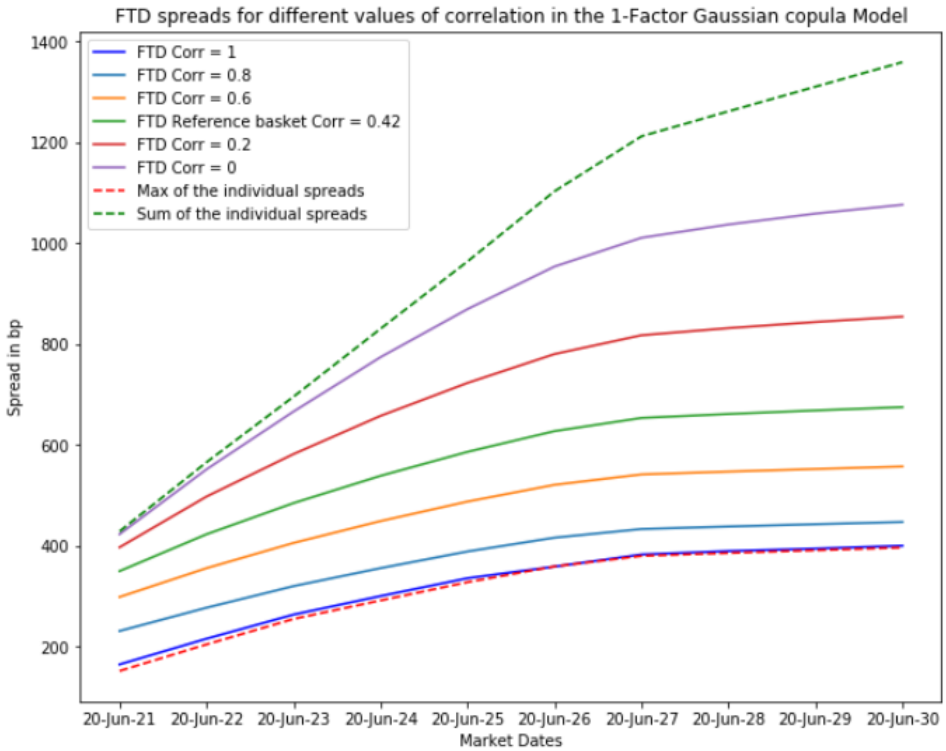

Figure 15: First-to-default spreads in the 1F-Gaussian copula model for $\rho = 1, 0.8, 0.6, 0.42, 0.2, 0$

Those results are in line with our expectations. Moreover, we can see that the same way the FTD default probability lies between the maximum of the individuals probabilities and the sum of all the probabilities, the FTD spread lies between the maximum of the individual spreads and the sum of the individual spreads.

5.2 Implied correlation and historical correlation

In this section, we will first describe what the implied correlation is and in a second time try to establish a link between the historical correlations and the implied correlations. In order to understand what implied correlation is, we can first define it in the context of Collaterized debt obligations (CDO's) as the reader may be more familiar with this type of product, which is one of the most popular multi-name credit derivative, and we'll see how we can draw a parallel between the implied correlation in CDO's and the implied correlation in First-to-default baskets.
Assume we are given a portfolio of typically 125 issuers who can potentially default and thus cause losses to the issuers exposed to them. These losses can be divided into tranches, such that each tranche delimits a portion of the loss of the portfolio between two values. These tranches are called the CDO tranches and the index can be traded in terms of tranches. There are six standard tranches of the iTraxx Europe covering losses lying in the ranges :
.0-3% (Equity tranche)
.3-6% (Mezzanine tranche)

.6-9%
.9-12%
.12-22% (Senior tranche)
.22-100% (Super senior tranche)
Each CDO tranche involves a protection seller and a protection buyer in which the protection seller agrees to pay to the buyer an amount equal to all the default losses of the reference portfolio (minus the recoveries) when the loss lies between the two percentage bounds of the tranche, and the protection buyer pays a periodic amount to the seller corresponding to the tranche spread value.
The protection fees on each CDO tranche depend a lot on the dependency between the defaults. In practice, this dependency is simply called "correlation" because the correlation parameter is enough to specify a multivariate Gaussian joint distribution and the market assumes a Gaussian copula model for the defaults in the portfolio. In our hypothetical portfolio, the input correlation matrix of the copula has $125 \times \frac{124}{2} = 7750$ entries corresponding to the pairwise correlations between every issuer. However, the market assumes a single correlation value ρ for each tranche so the 7750 initial parameters are simplified to a single one. In order to determine the value of the correlation ρ corresponding to a standard tranche, we choose the correlation value such that the price of the tranche in the Gaussian copula model is the same as the price of the tranche given by the market. The parameter ρ is then called implied correlation as it is the value of the correlation implied by the market. Let's get back to the First-to-default basket case and see why the implied correlation is analogous to the CDO context.

A First-to-default situation actually corresponds to the Equity tranche for CDO's. The reason for that is that in the two situations, it is the first default which triggers the payment from the protection buyer, and there is no protection anymore for the upcoming defaults, the same way that the equity tranche covers the losses between 0 and 3%, so that a the protection seller has to pay the buyer after the first loss and there is no protection anymore after this event. Of course, in mosts cases, the capital tranche corresponding to a First-to-default is not the 0-3% tranche, because there is typically 5 issuers in a First-to-default basket against 125 in our previous example with CDO's which are much more granular, so the first tranche's bounds would be between 0% and 20%, but the reasoning remains valid. The same analogy could be established between the n-to-default baskets and the higher tranches, but we will stick with First-to-default baskets in this section.
In this project, we used a service named Totem, which provides market derivatives prices and in particular the First-to-default market spread values for some baskets of issuers. In order to calibrate the correlation value of the basket, we proceed as described above by choosing the correlation ρ which makes the spread value computed in the Gaussian copula model being equal to the spread market value. However, there are a large amount of issuers and lots of possibilities for the composition of a First-to-default basket, and Totem cannot provide a market spread value for each of these possibilities. Yet, we have access to all the historical pairwise correlations between the issuers at any date and in order to solve this issue, we may wonder if a link could be established between the average pairwise correlation of the issuers in

a First-to-default basket and the implied correlation for this basket, and more precisely if there is a transformation that could be applied to these historical pairwise correlations that would enable us to predict the implied correlation value.
To do so, we performed a linear regression using the OLS method described in the Preliminaries section. Our vector of observations x contained the historical pairwise correlation at a date T_i for a given quantile q of 43 First-to-default baskets and the vector y of labels contained the Totem calibrated correlation for each basket. We performed an OLS regression at several dates $\{T_i\}_{i=1,...,p}$ (1 month spread between each date). It turned out that the values of y were better aligned when increasing the quantile q but the regression didn't perform well in any cases. For very close values of historical correlations, the Totem implied correlation was very different and there were a lot of outliers. We cannot show the graphs of the regressions for confidentiality reasons, and it wouldn't be relevant as we couldn't come up with a clear pattern. One of the reasons might be that historical correlations do not move a lot from month to month whereas the implied correlation is driven by the market and is more sensitive to external events (such as the current COVID-19 pandemic for instance), which makes it difficult to establish a clear relationship between the two.

6 Summary and conclusions

In this thesis, we studied the default dependency between the names within a First-to-default basket by introducing different models and assumptions to describe their dependence structure.

We first discussed several properties about First-to-default baskets in a general case such as the general expression of the cash flows implied by a First-to-default contract, or the way the joint distribution of the issuer defaults affects a First-to-default basket characteristics.

Then, we moved on to a more specific framework by presenting different models, which require to introduce the concept of copulas, which was a key point of this thesis, as we used them in all of our models. As there needs to be a correlation factor to describe the defaults dependency of our First-to-default baskets, the copulas we use to connect the default times need to have a correlation parameter. This is why we focused on the Gaussian and the Student-t copulas. In both Gaussian and Student-t copula models, we were aiming at showing the computations of several quantities of interest, such as the default probability of an issuer in the basket, the First-to-default survival (and default) probability of the basket, the First-to-default spread and how to price them.

Most of our simulation algorithms were based on a Monte-Carlo approach. We insisted on the necessity of having a fast implementation to run such algorithms as a Monte-Carlo simulation convergence rate is $O(1/\sqrt(M))$ (M being the number of simulation) while it scales in $O(M)$ in time. In order to increase the accuracy of the estimate, we presented in an appendix the concept of Importance sampling and two algorithms to perform it.

We had already shown some results in the previous section (such as the conditional hazard rate distribution) but they were mainly for illustration purposes. In section 5.1, we present numerical results obtained by implementing the models in order to compare the Gaussian and the Student-t copula models, by showing the differences between quantities discussed in the previous section : the First-to-default default probabilities and the First-to-default spreads. In order to assess the consistency of our implementation, we check if our results are in line with the theoretical expectations. More specifically, we would expect the FTD default probability to lie between the two bounds mentioned in section 3.2, and to be equal to them in limit cases scenarios. The FTD default probability should also be the same in the case of a Gaussian copula and a Student-t copula with very high degrees of freedom as the Student-t distribution tends to a standard normal distribution in this case. As for the FTD spreads, we expect them to fall as correlation increases. We managed to have satisfying results as these expectations were all satisfied.

In the very last section, we introduced the important concept of implied correlation, and we tried to establish a link between the historical average pairwise correlation of the issuers and the implied correlation of the basket using linear regression techniques. However, there were a lot of outliers and not many basket implied correlation data to perform a satisfying regression.

As we have seen, the Gaussian copula has the disadvantage of not capturing the whole correlation structure of the defaults, and in particular, the Gaussian copula doesn't exhibit tail dependence. Then, the Student-t copula seems to be a good alternative as it also has a correlation parameter and has upper and lower tail dependencies, so extreme scenarios are taken into account. However, for baskets containing a lot of names or CDO's, it requires a very high computational time. A possible alternative, not based on the use of any copula, has been discussed by Chapovsky, Rennie and Tavares. In this approach, the dependence structure is described through the default intensities of the issuers.

References

[1] D. O'Kane *Modelling Single-name and Multi-name credit derivatives* Wiley Finance, 2008.

[2] John C. Hull *Options, Futures and other Derivatives* Pearson, 1995.

[3] D. Brigo, A. Pallavicini, R. Torresetti *Credit Models and the Crisis, or : How I learned to stop worrying and love the CDOs* Arxiv, 2010.

[4] D. Brigo, A. Pallavicini, R. Torresetti *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models* Wiley Finance, 2010.

[5] P. Glasserman *Monte Carlo Methods in Financial Engineering* Springer Science + Business Media New York, 2004.

[6] P. Jäckel *Monte Carlo Methods in Finance* John Wiley Sons, 2002.

[7] D. Brigo, K. Chourdakis *Counterparty Risk for Credit Default Swaps: Impact of spread volatility and default correlation* ssrn.com, 2008.

[8] K. Aas *Modelling the dependence structure of financial assets: A survey of four copulas, Norwegian Computing Center* nr.no, 2004.

[9] *First-to-Default Swaps* Mathworks.com, 2009.

[10] Loeffler, Gunter, Posch *Credit risk modeling using Excel and VBA* Wiley Finance, 2007.

[11] D. Brigo, J. Mai, M. Scherer *Consistent Iterated Simulation of Multi-Variate Default Times: A Markovian Indicators Characterization.* SSRN, 2014.

[12] P. Glasserman, Z. Chen *Fast Pricing of Basket Default Swaps* Operations Research, gsb.columbia.edu, 2007.

[13] P. Glasserman, J. Li *Importance sampling for portfolio credit risk* Management Sci., 2005.

[14] J.-P. Laurent, J. Gregory *Basket default swaps, CDOs, and factor copulas* University of Lyon BNP Paribas, 2003.

[15] D.X. Li *On default correlation: a copula function approach* Shanghai Advanced Institute of Finance, SJTU, 1999.

[16] A. Steiner *Valuation of nth to Default Swaps* University of Oxford, 2004.

[17] P. Embrechts, F. Lindskog, A. McNeil *Modelling dependence with copulas and applications to risk management* ETHZ, 2001.

[18] J.C. Hull, A. White *Valuation of a CDO and an n th to Default CDS Without Monte Carlo Simulation* Journal of derivatives, 2004.

[19] P. Demey, T. Roncalli *Interpretation and estimation of default correlations* Recherche opérationnelle Crédit Agricole SA, 2004.

[20] A. Chapovsky, A. Rennie, P. Tavares *Stochastic Intensity Modelling for Structured Credit Exotics* International Journal of Theoretical and Applied Finance, 2007.