# 國立交通大學

## 統計學研究所

## 碩 士 論 文

# 在因子關聯結構模型下用高效率權重取樣估計聯合違約機率

# Efficient Importance Sampling in Estimation of Default Probability under Factor Copula Models

研 究 生：李侃學

指導教授：韓傳祥 教授

洪慧念 教授

中 華 民 國 九 十 八 年 六 月

# 在因子關聯結構模型下用高效率權重取樣估計聯合違約機率
# Efficient Importance Sampling in Estimation of Default Probability under Factor Copula Models


研 究 生：李侃學　　Student：Kan Hsueh Lee

指導教授：韓傳祥　　Advisor：Chuan-Hsiang Han

洪慧念　　Hui-Nien Hung


國 立 交 通 大 學

統計學研究所

碩 士 論 文

A Thesis

Submitted to Institute of Statistics

College of Science

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

# 在因子關聯結構模型下用
# 高效率權重取樣估計聯合違約機率

研究生：李侃學　　　　　　　　　　　　　　　　　　　　指導教授：韓傳祥
洪慧念

國立交通大學統計學研究所

## 摘　　　要

權重取樣是常見的蒙地卡羅方法之一，可用於有效的估計罕見事件。 此方法加重罕見事件發生之機率，所以適用於估計一些罕見之違約相關事件。 此方法也同時容易應用於處理不同的違約事件上，但主要的問題在於選擇有效的權重取樣，不但加重罕見事件機率，也同時縮小估計值變異數。 在多變量的架構下，縮小變異數的問題相當複雜，也多數無直接解答。此論文提出有效的權重取樣演算法，同時的增加罕見事件機率並縮小估計量變易數，在 Large Deviation Theory 下縮小變異數，然後將此演算法用於常見的信用商品上，並提出數據說明此演算法有效的增進速度及估計準確度。

# Efficient Importance Sampling in Estimation of Default Probability under Factor Copula Models

Student: Kan Hsueh Lee

Advisor: Chuan-Hsiang Han
Hui-Nien Hung

Institute of Statistics

National Chiao Tung University

## Abstract

Importance sampling is a commonly used technique to improve Monte Carlo methods, especially in working with rare events. It is designed to increase the probability of sampling from rare events and is therefore well-suited for estimating default related items in various products given the rarity of default events. It is also simple to implement and versatile in that in can be easily extended to estimate different items. But the main challenge is selecting an importance sampling scheme that not only increases the probability of rare events but also effectively reduces the variance of the estimate. Under the multivariate framework when multiple entities are involved, variance reduction becomes even more challenging as there is no closed form solution for such optimization problem. In this study, we propose an effective importance sampling algorithm that both increases the probability of rare events and reduce variance of estimates. We consider the problem of variance reduction under the framework of Large Deviation Theory, and establish an efficient importance sampling estimator that can be applied to evaluating default events. Then we extend this importance sampling scheme to another popular type of default event and incorporate it into a conditional importance sampling scheme. Our numerical results confirm that the proposed algorithms for direct importance sampling and conditional importance sampling are more efficient in terms of variance reduction. Our algorithms are overall more robust under different specified initial conditions.

# 誌謝

I am thankful for the opportunity to study at NCTU these past two years. I want to first thank Dr. Chung-Hsiang Han for being my advisor and very thoroughly guiding me on my thesis. Without his help, I could not have finished this research project. He has been very patient in teaching me the basics of quantitative finance step by step all the way to helping me understand current research topics in the field of credit risk modeling. He also provided me the opportunity to intern at the futures trading department in ChinaTrust, which turned out to be an extremely valuable experience. Now I not only understand the basics of continuous-time finance but also appreciate how such tools are used in industry. It is my pleasure and privilege to have studied under Dr. Han. I also want to thank my co-advisor Dr. Hui-Nien Hung for giving me valuable research tips. As an experienced practitioner and researcher in many fields of statistics, Dr. Hung gave me valuable career advice. Finally, I want to thank my church for giving me the opportunity to serve in college ministry here in NCTU and NTHU. I thoroughly enjoyed the experience and wish I can stay here longer. Lastly, I want to thank all my friends and family for supporting me and always encouraging me. I could not have survived these past two years without these precious people in my life.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction and Motivation

Effectively estimating default probability in credit derivatives has been an area of ongoing research. The problem starts with the characterization of default time. There are two main approaches: structural form and reduced form. Structural form models the asset and debt value of the company. It treats them as a first passage time problem and considers default to occur when asset value of the firm falls below the debt value. Merton [17] (1974) first proposed this model and later Black and Cox [1] (1976), Geske [9] (1977), Leland [16] (1994), Longstaff and Schwartz [15] (1995), and Zhou [18] (2001)also followed this line of thinking. They mainly worked with default of a single firm. Only Zhou [18] (2001) modeled default of two firms with two correlated brownian motions and derived closed form solution to the joint default probability of the two firms. His results, however cannot be easily extended to more firms. Hull and White [11] (2001) shows how multiple firms can be dealt with using Monte Carlo methods. Typically, evaluating such models requires computationally intensive numerical procedures.

Reduced form models on the other hand treats default as exogenous. It bypasses the particular firms captical structure and uses available market information to model defaults. Schönbucher [19] (2003), Duffie [7] (2003), and Lando [12] (2003) are a few main proponents of this approach. Up to now, the industry standard follows the reduced form's line of thinking. Li [13] (2000) first developed the copula approach to develop the correlation structure among default times and Laurent and Gregory [10] (2005) later extended it and represented the correlation structure in factor form, also known as factor-copula approach. We will follow the industry standard and adopt the factor copula approach in this study.

Once default time is characterized, we can utilize the model to now evaluate joint default probability and various credit derivatives. Basket Default Swaps (BDS) is a common type of multi-name credit derivatives along with Collaterized Debt Obligations (CDO). Chiang, Yueh, and Hsieh [5] (2007) proposed an efficient algorithm for valuing BDS (hereafter as CYH). Their study proposes an efficient algorithm to evaluate $k^{th}$ to default BDS. It adopts the factor copula approach to modeling default time and mainly works with Gaussian copula. In this study, we generalize their study and consider joint default probability along with $k^{th}$ to default BDS under both Gaussian and Student-T copula. We propose another algorithm that is more efficient and approach the problem of variance reduction from the perspective of Large Deviation Theory. Large Deviation Theory is an active area of applied probability that mainly focuses on behavior of extremal events. We use results from Large Deviation Theory to solve the problem of variance reduction for our estimation algorithm.

# Chapter 2

# Characterization of Default Time

## 2.1 Default Time of a Single Firm

We characterize distribution of a firm's default time $\tau$ in terms of its hazard rate function $h(.)$. Here we briefly review the definition and relationship between survival function, hazard function and CDF.

**Definition 2.1** (survival function)**.** If $\tau$ is the default time of a firm with CDF $F$, then $S(t) = \mathbb{P}(\tau > t) = 1 - F(t)$ is the *survival function* of $\tau$.

**Definition 2.2** (hazard rate function)**.** Suppose the default time $\tau$ has density function $f(t)$ and survival function $S(t)$, then the hazard rate function $h(t)$ is defined as

$$h_\tau(t) = \frac{f_\tau(t)}{S_\tau(t)}.$$

Hazard rate function is useful for understanding probability of a firm's default immediately after time $t$, given that it has survived up to time $t$. We can understand hazard rate function in the following conditional probability.

Given a firm has survived $t$ years, the probability it will default in the coming time interval $\Delta t$ can be written as follows:

$$\mathbb{P}(\tau \in (t, t + \Delta t) | \tau > t) = \frac{\mathbb{P}(\tau \in (t, t + \Delta t))}{\mathbb{P}(\tau > t)} \approx \frac{f(t)\Delta t}{1 - F(t)} = h(t)\Delta t. \quad (2.1)$$

We can see that this probability can be approximated by the value of hazard rate function at time $t$ times a small time increment $\Delta t$.

According to definition 2.2, we can write the distribution function in terms of the hazard rate function:

$$F(t) = \mathbb{P}(\tau \le t) = 1 - \exp\left\{-\int_0^t h(s)ds\right\} \tag{2.2}$$

Equation (2.2) shows that with effective estimation of hazard rate, we can model distribution of default time. According to Cherubini, Luciano and Vecchiato [4] (2004), the hazard rate function can be obtained in several ways:

- From historical default rates provided by rating agencies.
- By using the Merton approach according to Delianedis and Geske [6] (1998).
- Extracting default probabilities by using market observable information, such as asset swap spread, CDS spread or corporate bond prices according to Li [14] (1998).

We will not focus on methods of extracting hazard rate function in this study. A typical assumption is that the hazard rate is a constant $\lambda$. In this case, default time $\tau$ follows an exponential distribution with intensity $\lambda$. Here and on, we will use this assumption. This implies:

$$F(t) = \mathbb{P}(\tau \le t) = 1 - exp(-\lambda t) \tag{2.3}$$

## 2.2 Brief Introduction to Copula Function

We have characterized default time of a single firm, and now proceed to combine default time distributions of different firms into a joint distribution through copula functions. First, we give a brief introduction to copula functions.

**Definition 2.3** (copula function)**.** A n-dimensional copula $C$ is a real-value function with range $I = [0, 1]$ and domain $I^n$ such that

- $C(\mathbf{u})$ is increasing in every component $u_k$, $k = 1, 2 \cdots n$, called n-increasing.
- For every $\mathbf{u} \in I^n$, $C(\mathbf{u}) = 0$ if $u_k = 0$ for some $k$, and $C(\mathbf{u}) = u_k$ if $u_i = 1$, $i = 1, 2 \cdots n$ except the $k$-th component.

- For all **a**,**b** $\in I^n$ with **a** $\leq$ **b** in every component, then the n-box **B** $= [a_1, b_1] \cdot [a_2, b_2] \cdots [a_n, b_n]$ satisfies $V_n(\mathbf{B}) \geq 0$, where $V_n$ is the $n$-volume.

Copula function, $C$, can be intrinsically understood as a multivariate distribution function with uniform marginal distributions.

$$C(u_1, ..., u_n) = \mathbb{P}\left(U_1 \leq u_1, \cdots, U_n \leq u_n\right)$$

Let $F$ be n-dimensional distribution with $F_1, ..., F_n$ as the univariate marginal distributions. Note that $u_i = F_i(x_i)$ is uniform on $[0, 1]$ for $i = 1, ..., n$. The copula function can combine these uniform marginals $u_1, ..., u_n$ into a multivariate distribution function. By probability integral transformation, we can write,

$$\begin{aligned}
C(F_1(x_1), \cdots, F_n(x_n)) &= \mathbb{P}(U_1 \leq F_1(x_1), \cdots, U_n \leq F_n(x_n)) \\
&= \mathbb{P}(F_1^{-1}(U_1) \leq x_1, \cdots, F_n^{-1}(U_n) \leq x_n) \\
&= \mathbb{P}(X_1 \leq x_1, \cdots, X_n \leq x_n) \\
&= F(x_1, \cdots, x_n)
\end{aligned}$$

, where *Sklar theorem* guarantees the converse.

**Theorem 2.1** (Sklar 1959). *Let $F$ be an n-dimensional multivariate distribution function with marginal distributions $F_1(\cdot), \cdots, F_n(\cdot)$. Then there exists an n-dimensional copula function $C$ such that*

$$F(x_1, \cdots, x_n) = C(F_1(x_1), \cdots, F_n(x_n)).$$

*Furthermore, if $F_1(\cdot), \cdots, F_n(\cdot)$ are continuous, then $C$ is unique.*

Therefore, if we know marginal distributions $F_1, ..., F_n$ then we can specify the copula function and joint distribution.

## 2.3 Joint Default Time Under Gaussian and Student-T Copula

Now we have characterized default time of a single firm. We will combine default time distributions of several firms into a joint distribution through copula functions. We will focus on Gaussian and Student T copula functions and provide algorithms for generating joint default time according to these two copula functions.

Now suppose there are $n$ firms, and as mentioned before, we assume the default time $\tau_i$ of each each firm follows an exponential distribution with intensity $\lambda_i$ for $i = 1, ..., n$. Let $F_i(.)$ be the distribution function of default time $\tau_i$ for firm i, $i = 1, ..., n$. The main purpose of using copula function here is to generate a set of correlated uniform variates $(U_1, ..., U_n)$ according to the specified copula correlation structure. Next, we use the correlated uniform variates generated from the copula function to compute default time for each individual firm through inverse mapping of the firm's default time distribution i.e. $\tau_i = F_i^{-1}(U_i)$ for $i = 1, ..., n$.

**General Form of Gaussian Copula**

$$C(u_1, u_2, \cdots, u_n; \Sigma) = \Phi_\Sigma(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \cdots, \Phi^{-1}(u_n))$$

where $\Phi_\Sigma(.)$ is the standardized multivariate normal distribution with covariance matrix $\Sigma$ and $\Phi(.)$ is CDF of $N(0, 1)$. In this case, given $u_i = F_i(\tau_i)$, where $F_i(.)$ is default time distribution of firm i, we can rewrite the Gaussian copula function as follows,

$$C(F_1(\tau_1), \cdots, F_n(\tau_n); \Sigma) = \Phi_\Sigma(\Phi^{-1}(F_1(\tau_1)), \cdots, \Phi^{-1}(F_n(\tau_n)))$$

We can sample joint default time from a Gaussian copula as follows:

**Algorithm 2.1** (Joint default time under Gaussian copula).

1. Given $\Sigma$, generate correlated uniform variates $(U_1, ..., U_n)$ from the Gaussian Copula function.

   (1) Find the Cholesky decomposition A of $\Sigma$ such that: $\Sigma = AA^T$.

   (2) Generate n independent random variables $Z = (Z_1, \cdots, Z_n)^T$ from $N(0, 1)$.

   (3) Let $X = (X_1, ..., X_n)^T = AZ$. Now we know $W \sim \Phi_\Sigma(.)$

   (4) Let $U_i = \Phi(X_i)$ for $i = 1, ..., n$.

2. Use correlated uniform variates $(U_1, ..., U_n)$ generated from Gaussian Copula to compute default time through inverse mapping of the firm's default time distribution.

   (1) Let $(U_1, ..., U_n)^T = (F_1(\tau_1), ..., F_n(\tau_n))^T$.

   (2) Let $\tau_i = F_i^{-1}(U_i)$, $i = 1, \cdots n$.

**General Form of Student-T Copula**

$$C(u_1, u_2, \cdots, u_n; \Sigma, \nu) = T_{\Sigma,\nu}\left(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \cdots, t_\nu^{-1}(u_n)\right)$$

where $T_{\Sigma,\nu}(.)$ is the standardized multivariate Student-T distribution with covariance matrix $\Sigma$ and degrees of freedom, $\nu$. $t_v(.)$ is CDF of univariate Student-T with $\nu$ degrees of freedom. In this case, given $u_i = F_i(\tau_i)$, where $F_i(.)$ is default time distribution of firm i, we can rewrite the Student-T copula function as follows,

$$C(F_1(\tau_1), \cdots, F_1(\tau_n); \Sigma, \nu) = T_{\Sigma,\nu}\left(t_\nu^{-1}(F_1(\tau_1)), \cdots, t_\nu^{-1}(F_n(\tau_n))\right)$$

We can sample joint default time from a Student-T copula as follows:

**Algorithm 2.2** (Joint default time under Student-T copula)**.**

1. Given $\Sigma$, generate correlated uniform variates $(U_1, ..., U_n)$ from Student-T copula function.

   (1) Find the Cholesky decomposition A of $\Sigma$ such that $\Sigma = AA^T$.

   (2) Generate $n$ independent random variables $Z = (Z_1, \cdots, Z_n)^T$ from $N(0, 1)$.

   (3) Generate $\chi^2_\nu$, a Chi-square variable with d.f.$=\nu$.

   (4) Let $X = AZ$. Then $X \sim N(0, \Sigma)$.

   (5) Let $S = (S_1, ..., S_n)^T = X/\sqrt{\chi^2_\nu/\nu}$. Then $S \sim T_{\Sigma,\nu}(.)$

   (6) Let $U_i = t_\nu^{-1}(S_i)$ for $i = 1, ..., n$.

2. Use correlated uniform variates $(U_1, ..., U_n)$ generated from Student-T copula to compute default time through inverse mapping of the firm's default time distribution.

   (1) Let $(U_1, ..., U_n)^T = (F_1(\tau_1), ..., F_n(\tau_n))^T$.

   (2) Let $\tau_i = F_i^{-1}(U_i)$, $i = 1, \cdots, n$.

## 2.4 Covariance Matrix under Gaussian Factor Copula Model

We have shown how to determine joint default time using Gaussian and Student-T copula functions. But we haven't discussed the exact structure of correlation which is mainly determined by the covariance matrix $\Sigma$. To determine $\Sigma$, $n(n-1)/2$ variables need to be estimated, which is extremely difficult when $n$ is large. Laurent and Gregory [10] (2004) proposes the factor model which greatly reduces the number of variables by using two types of factors to intuitively explains firms' behavior in terms of economic trend and idiosyncratic movements. It reduces complexity from $O(n^2)$ to $O(n)$ Therefore, we adopt the factor model proposed by Laurent and Gregory for determining $\Sigma$. Under this model,

$$X_i = \rho_i Z_0 + \sqrt{1-\rho_i^2} Z_i, \; i = 1, 2, \cdots, n, \tag{2.4}$$

where $Z_0, \cdots, Z_n$ are i.i.d. $N(0,1)$ and $\rho_1, ..., \rho_n \epsilon [0,1]$ Let $X = [X_1, ..., X_n]^T$, then

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \rho_1 & \sqrt{1-\rho_1^2} & & & \\ \rho_2 & & \sqrt{1-\rho_2^2} & & \\ \vdots & & & \ddots & \\ \rho_n & & & & \sqrt{1-\rho_n^2} \end{bmatrix} \begin{bmatrix} Z_0 \\ Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}$$

The factor loading $\rho_i$ determines how strongly the $i$th factor is correlated to the common factor $Z_0$. Both $Z_0$ and $Z_i$ are $N(0,1)$, and the constraints posed on on factor loadings ensure that every factor $X_i$ is $N(0,1)$.

We can see that $X$ has multivariate normal distribution $N(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} 1 & \rho_1\rho_2 & \rho_1\rho_3 & \cdots & \rho_1\rho_n \\ & 1 & \rho_2\rho_3 & \cdots & \rho_2\rho_n \\ & & & \ddots & \vdots \\ & & & 1 & \rho_{n-1}\rho_n \\ & & & & 1 \end{bmatrix}$$

Suppose $X_i$ represents firm $i$, one can intuitively understand $Z_0$ as the common factor such as certain macroeconomic or industry condition that affects all firms in a similar fashion, and $Z_i$ as the firm specific factors that affect only the particular firm. Hence we call $Z_0$ the *common factor*, and each $Z_i$ the *marginal factor*.

# Chapter 3

# Estimating Joint Default Probability

Now that we have characterized joint default time distribution, we proceed to formulate the problem of estimating joint default probability. Given time T, we wish to know the probability of all the firms defaulting sometime before T. i.e. we wish to evaluate the following,

$$p = \mathbb{P}\left(\tau_1 \leq T, ..., \tau_n \leq T\right) = \mathbb{E}\left\{\mathbb{I}_{(\tau_1 \leq T, ..., \tau_n \leq T)}\right\} = \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(\tau_i \leq T)}\right\} \tag{3.1}$$

Direct computation of this probability is equivalent of evaluating the CDF of default time $(\tau_1, ..., \tau_n)$ at $(T, ..., T)$

$$p = F(T, ..., T)$$

But this often involves evaluation of a complex multiple integral that has no closed form solution. One can then resort to numerical methods. But this integration suffers from curse of dimensionality, causing the accuracy of numerical integration to decrease significantly as dimension $n$ increases. The integral takes the following forms under Gaussian and Student-T copula. As mentioned above, here we denote $F_i(.)$ for $i = 1, ..., n$ as default time distribution for firm i.

**Under Gaussian Copula**

Recall in Algorithm 2.1, given $X = (X_1, ..., X_n) \sim N(0, \Sigma)$ and $X_i \sim N(0,1)$ for $i = 1, ..., n$, we can generate default time as $\tau_i = F_i^{-1}(\Phi(X_i))$ for $i = 1, ..., n$. Then we can also write,

$$\{\tau_i \leq T\} = \left\{F_i^{-1}(\Phi(X_i)) \leq T\right\} = \left\{X_i \leq \Phi^{-1}(F_i(T))\right\}$$

This means joint default probability becomes a problem of calculating the multivariate normal CDF, $\Phi_\Sigma(.)$,

$$\begin{aligned} p &= \Phi_\Sigma(\phi^{-1}(F_1(T)), ..., \phi^{-1}(F_n(T))) \qquad (3.2) \\ &= \int_{-\infty}^{\phi^{-1}(F_1(T))} \cdots \int_{-\infty}^{\phi^{-1}(F_n(T))} \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} exp\left(-\frac{1}{2}x^T\Sigma^{-1}x\right) dx_1...dx_n \end{aligned}$$

**Under Student-T Copula**

Recall in Algorithm 2.2, given $S = (S_1, ..., S_n) \sim T_{\Sigma,\nu}$ and $S_i \sim t_\nu$ for $i = 1, ..., n$, we can generate default time as $\tau_i = F_i^{-1}(t_\nu(S_i))$ for $i = 1, ..., n$. Then we can also write,

$$\{\tau_i \leq T\} = \left\{F_i^{-1}(t_\nu(S_i)) \leq T\right\} = \left\{S_i \leq t_\nu^{-1}(F_i(T))\right\}$$

This means joint default probability becomes a problem of calculating the multivariate Student-T CDF, $T_{\Sigma,\nu}(.)$,

$$\begin{aligned} p &= T_{\Sigma,\nu}(t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T))) \qquad (3.3) \\ &= \int_{-\infty}^{t_\nu^{-1}(F_1(T))} \cdots \int_{-\infty}^{t_\nu^{-1}(F_n(T))} \frac{\Gamma\left(\frac{\nu+n}{2}\right)|\Sigma|^{\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)(v\pi)^{\frac{n}{2}}} \left(1 + \frac{1}{v}x^T\Sigma^{-1}x\right)^{-\frac{\nu+n}{2}} dx_1...dx_n \end{aligned}$$

One can easily observe that evaluation of these two integrals is very difficult, especially when n is large. For our purpose, n is usually larger than 5, which rules out numerical integration as a practical approach. Therefore, we approach the problem with Monte Carlo methods. In this study, we focus on Basic Monte Carlo (Basic MC)and importance sampling methods. In literature, Genz and Bretz [8] (1999) has proposed Quasi Monte Carlo methods (Quasi MC), which is widely adopted as the main numerical method when $n$ is large.

## 3.1 Basic Monte Carlo Method

Basic MC method is simple to implement. Based on the law of large numbers, the joint default probability can be approximated by its sample mean when $m$ is sufficiently large,

$$p = \mathbb{E}\left\{\prod_{i=1}^{n}\mathbb{I}_{(\tau_i \le T)}\right\} \approx \frac{1}{m}\sum_{j=1}^{m}\prod_{i=1}^{n}\mathbb{I}_{(\tau_{i,j} \le T)}$$

where $\tau_{1,j}, \ldots, \tau_{n,j}$ for $j = 1, \ldots, m$ are m samples of $\tau_1, \ldots, \tau_n$. Based on the above, we just need to generate m samples for $\tau_1, \ldots, \tau_n$ according to Algorithm 2.1 for Gaussian copula and Algorithm 2.2 for Student-T copula and then evaluate the indicator function for our $m$ samples. However, default event for highly ranked firms is typically very rare, causing default times of $n$ firms to all be less than or equal to T an extremely rare event. This makes Basic MC method inaccurate. Therefore we modify the Basic MC method with importance sampling techniques to improve accuracy of our estimation.

## 3.2 Brief Review of Importance Sampling

Importance sampling is a commonly used tool for rare event simulation. The basic idea is to change the original probability measure $\mathbb{P}$ to a new probability measure $\tilde{\mathbb{P}}$ that puts more weight on the rare event we want to sample. With a good choice of $\tilde{\mathbb{P}}$, we can increase the simulation efficiency by generating more desired samples as well as reducing variance.

Suppose we want to estimate

$$\theta = \mathbb{E}[h(X)] = \int_{\mathbb{R}^n} h(x)f(x)dx \tag{3.4}$$

where $X = [X_1, \cdots, X_n]^T$ is a random vector in $\mathbb{R}^n$ with a joint density $f(x) = f(x_1, \cdots, x_n)$. Basic Monte Carlo Method gives us the following estimate

$$\hat{\theta} = \frac{1}{m}\sum_{i=1}^{m} h\left(X^{(i)}\right), \tag{3.5}$$

where every $X^{(i)}$ for $i = 1, \ldots, m$ are i.i.d. samples from probability measure $\mathbb{P}_f(.)$ with density $f(.)$.

If we find it ineffective to sample from $\mathbb{P}_f(.)$, we can sample from another probability measure, call it $\mathbb{P}_g(.)$ with density $g(.)$, which is absolute continuous with respect to $f(.)$. Then we can write the following,

$$\theta = \int_{\mathbb{R}^n} h(x)f(x)dx = \int_{\mathbb{R}^n} \frac{h(x)f(x)}{g(x)} g(x)dx = \mathbb{E}_g \left[ \frac{h(Y)f(Y)}{g(Y)} \right], \tag{3.6}$$

where the random vector $Y$ has the density $g(\cdot)$. Now our original estimator $\hat{\theta}$ is replaced by

$$\hat{\theta} = \frac{1}{m} \sum_{i=1}^{m} \frac{h\left(Y^{(i)}\right) f\left(Y^{(i)}\right)}{g\left(Y^{(i)}\right)}, \tag{3.7}$$

where the random samples are taken from $g(\cdot)$. The weight $\frac{f(Y^{(i)})}{g(Y^{(i)})}$ is called the likelihood ratio or Radon-Nikodym derivative evaluated at $Y^{(i)}$. For notational simplicity, here and on, we will denote Radon-Nikodym derivative in the following fashion $\frac{d\mathbb{P}_f}{d\mathbb{P}_g}$,

$$\frac{d\mathbb{P}_f}{d\mathbb{P}_g}(Y) = \frac{f(Y)}{g(Y)}$$

Our goal is twofold: 1)Find an appropriate measure $g(.)$ such that $g(Y) >> f(Y)$ on important regions. This will increase probability of generating rare samples from those regions we are interested in, 2) Find $g(.)$ that minimizes variance of importance sampling estimator, $Var(\hat{\theta})$.

## 3.3 Importance Sampling Problem Description

Having established the basic notion of importance sampling, we now proceed to formulate our importance sampling scheme. We will consider importance sampling under Gaussian copula first and then apply the results to Student T copula. Under Gaussian copula, we know that for $i = 1, ..., n$,

$$\{\tau_i \le T\} = \left\{F_i^{-1}(\Phi(X_i)) \le T\right\} = \left\{X_i \le \Phi^{-1}(F_i(T))\right\}$$

Then we can write,

$$p = \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(\tau_i \le T)}\right\} = \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(X_i \le \Phi^{-1}(F_i(T)))}\right\}$$

Recall that $X = [X_1, ..., X_n]^T$ is multivariate normal $N(0, \Sigma)$. We need to employ importance sampling when $\Phi^{-1}(F_i(T))$ is very negative, i.e. when T is very small or when intensity $\lambda_i$ of default time distribution for firm $i$ is very small, which is often the case. We can generalize the problem in the following way. Given some $D = [d_1, ..., d_n]^T$ where $d_1, ..., d_n \in \mathbb{R}$, we wish to instead evaluate,

$$p = \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(X_i \le d_i)}\right\} = \mathbb{E}\left\{\mathbb{I}(X \le D)\right\}$$

Here $X \le D$ means elements of $X$ are less than elements of $D$ i.e. $X_1 \le d_1, ..., X_n \le d_n$. We will use this notation from here and on. Recall that we can use the Basic MC to approximate $p$. However, Basic MC simulation with random variable $X \sim N(0, \Sigma)$ under $\mathbb{P}(.)$ is inaccurate when $D$ is small. Therefore, we consider importance sampling with $Y \sim N(u, \Sigma)$ under new measure $\mathbb{P}_u(.)$ in approximating $p$. Our goal is to find u that will minimize variance of our importance sampling estimator.

Under new measure $\mathbb{P}_u(.)$

$$p = \mathbb{E}^u\left\{\mathbb{I}_{(Y \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u}\right\} \tag{3.8}$$

With $Y^{(j)} \sim N(u, \Sigma)$ as i.i.d. samples from $\mathbb{P}_u$, the importance sampling estimator will be

$$\hat{p} = \frac{1}{m}\sum_{j=1}^{m} \mathbb{I}_{(Y^{(j)} \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u} \tag{3.9}$$

Note that $\hat{p}$ is an unbiased estimator because $\mathbb{E}^u\{\hat{p}\} = p$. Since $\hat{p}$ is the average of $m$ i.i.d samples, its variance will be $\frac{1}{m}$ times the variance of $\left(\mathbb{I}_{(Y \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u}\right)$,

$$Var(\hat{p}) = \frac{1}{m}\left(\mathbb{E}^u\left\{\left(\mathbb{I}_{(Y \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u}\right)^2\right\} - p^2\right) \tag{3.10}$$

To minimize $Var(\hat{p})$, we just need to optimize with respect to $u$ the second moment of $\mathbb{I}_{(Y \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u}$.

$$F(u) := \mathbb{E}^u\left\{\left(\mathbb{I}_{(Y \le D)} \frac{d\mathbb{P}}{d\mathbb{P}_u}\right)^2\right\} \tag{3.11}$$

Such optimization problem is extremely complicated because $Y$ is multivariate, and currently there is no closed form solution for $F(u)$. But we can approach the problem of variance reduction from the framework of Large Deviation Theory, and find a asymptotic minimizer that can reduce the second moment, $F(u)$, and hence reduce variance. This will produce *efficient* importance sampling asymptotically. We will define this in the next chapter.

# Chapter 4

# Importance Sampling in Large Deviation Theory

## 4.1 Efficient Importance Sampling

First we need to establish the notion of efficient importance sampling in Large Deviation Theory . We reformulate the problem by introducing a scaling factor $-\sqrt{L}$ that will later allow us to apply results from Large Deviation Theory. Let $D = -\sqrt{L}C = -\sqrt{L}[c_1, \cdots, c_n]^T$, where $c_1, \cdots, c_n$ are positive constants. Now, we rewrite $p$ and $F(u)$ as $p_L$ and $F_L(u)$. Under $\mathbb{P}(.)$ Let $X^{(i)} \sim N(0, \Sigma)$ for $i = 1, ..., L$ be i.i.d samples from $\mathbb{P}(.)$, then we know $\frac{1}{\sqrt{L}}X$ and $\frac{1}{L}\sum_{i=1}^{L} X^{(i)}$ are equal in distribution. Thus,

$$p_L = \mathbb{E}\left\{\mathbb{I}(X \le -\sqrt{L}C)\right\} = \mathbb{E}\left\{\mathbb{I}(\frac{1}{\sqrt{L}}X \le -C)\right\} = \mathbb{E}\left\{\mathbb{I}(\frac{1}{L}\sum_{i=1}^{L} X^{(i)} \le -C)\right\} \tag{4.1}$$

Under $\mathbb{P}_u(.)$ Let $Y_i \sim N(u, \Sigma)$ for $i = 1, ..., L$ be i.i.d samples from $\mathbb{P}_u(.)$,

$$F_L(u) = \mathbb{E}^u\left\{\left(\mathbb{I}(Y \le -\sqrt{L}C)\frac{d\mathbb{P}}{d\mathbb{P}_u}\right)^2\right\} = \mathbb{E}^u\left\{\left(\mathbb{I}(\frac{1}{\sqrt{L}}Y \le -C)\frac{d\mathbb{P}}{d\mathbb{P}_u}\right)^2\right\} \tag{4.2}$$

Then based on (3.10) and (3.11) we know,

$$Var(\hat{p_L}) = \frac{1}{k}(F_L(u) - p_L^2) \tag{4.3}$$

As mentioned earlier, we do not know how to minimize (4.3) directly with respect to $u$, but Large Deviation Theory helps us to understand (4.3) when

$L$ is very large. More precisely, it has results that allow us to evaluate the following limits,

$$\begin{aligned} H &= \lim_{L\to\infty} \tfrac{1}{L} log\, p_L \\ R &= \lim_{L\to\infty} \tfrac{1}{L} log\, F_L(u) \end{aligned} \tag{4.4}$$

First, we observe that (4.3) will always be greater or equal to zero for all $L$. i.e. $F_L(u) \geq p_L^2$ for all $L$. Now with (4.4), we can conclude

$$R = \lim_{L\to\infty} \frac{1}{L} log\, F_L(u) \geq \lim_{L\to\infty} \frac{1}{L} log\, p_L^2 = 2 \lim_{L\to\infty} \frac{1}{L} log\, p_L = 2H \tag{4.5}$$

When $R = 2H$, we say that our importance sampling estimator is *efficient*. This means $F_L \approx p_L^2$ or equivalently $Var(\hat{p_L}) \approx 0$ when $L$ is sufficiently large. According to Bucklew [3] (2004), in the framework of Large Deviation Theory and rare event simulation, if a family of simulation distributions is efficient, then it is a good choice.

## 4.2 Applying Results of Large Deviation Theory

Having established the notion of efficient important sampling, we now wish to show important sampling from some $\mathbb{P}_u(.)$ is indeed *efficient* (i.e. $R = 2H$.) We will accomplish this by using the Gartner-Ellis Theorem and Bucklew's (1990) calculation in Large Deviation Theory[2].

**Theorem 4.1.** *(special case of Gärtner-Ellis Theorem) Let $\{S_L\}$ be a sequence of $\mathbb{R}^n$ valued random varianbles and let $\theta, a \in \mathbb{R}^n$. $\langle \cdot, \cdot \rangle$ denotes the usual inner or dot product between vectors. Let $C = [c_1, ..., c_n]^T$, where $c_1, \cdots, c_n$ are positive constants. Define*

$$\begin{aligned} \varphi(\theta) &= \lim_{L\to\infty} \frac{1}{L} log\mathbb{E}\{exp[\langle \theta, S_L \rangle]\} \\ I(a) &= \sup_{\theta} [\langle \theta, a \rangle - \varphi(\theta)] \end{aligned}$$

*Then,*

$$\lim_{m\to\infty} \frac{1}{L} log\mathbb{P}\{\frac{S_L}{L} \leq -C\} = -\inf_{a \leq -C} I(a) \tag{4.6}$$

Now we can use the above result to evaluate $H$. Here, we use the calculation provided by Bucklew [2]. Let $S_L = \sum_{i=1}^{L} X^{(i)}$ where $X^{(i)}$ are i.i.d $N(0,\Sigma)$ for $i = 1, ..., L$. Then,

$$\begin{aligned}
\varphi(\theta) &= \lim_{L\to\infty} \frac{1}{L} log\, \mathbb{E}\left\{ exp\left( \theta \sum_{i=1}^{L} X^{(i)} \right) \right\} \\
&= \lim_{L\to\infty} log\, \mathbb{E}\{exp(\theta X^{(1)})\} = log\, \mathbb{E}\{exp(\theta X^{(1)})\}
\end{aligned}$$

The moment generating function Gaussian random vector $X^{(1)}$ is well known to be $\mathbb{E}\{exp(\theta X^{(1)})\} = exp\{\langle\theta, 0\rangle + \frac{1}{2}\theta^T\Sigma\theta\}$. This implies that $I(a) = \sup_\theta[\langle\theta, a\rangle - \frac{1}{2}\theta^T\Sigma\theta]$ Setting the gradient with respect to $\theta$ in $\langle\theta, a\rangle - \frac{1}{2}\theta^T\Sigma\theta$ to zero results in $a - \Sigma\theta = 0$. This implies that $\theta_{opt} = \Sigma^{-1}a$. Substituting this value of $\theta$ back into the supremum expression yields

$$I(a) = \frac{1}{2}a^T\Sigma^{-1}a \tag{4.7}$$

Then,

$$\inf_{a\le -C} I(a) = \frac{1}{2}C^T\Sigma C \tag{4.8}$$

Based on (4.1), we can use the above theorem to conclude,

$$\begin{aligned}
H &= \lim_{L\to\infty} \frac{1}{L} log\, p_L \\
&= \lim_{L\to\infty} \frac{1}{L} log\, \mathbb{E}\left\{ \frac{1}{L}\sum_{i=1}^{L} X^{(i)} \le -C \right\} \\
&= \lim_{L\to\infty} log\, \mathbb{P}\left\{ \frac{S_L}{L} \le -C \right\} \\
&= -\inf_{a\le -C} I(a) \\
&= -\frac{1}{2}C^T\Sigma^{-1}C
\end{aligned} \tag{4.9}$$

We can see from here that $\exp\{-\frac{1}{2}LC^T\Sigma^{-1}C\}$ is a good approximate for $p_L$ when $L$ is large, i.e. when $-\sqrt{L}c_1 \cdots -\sqrt{L}c_n$ are very small. This intuitively suggest the application of Large Deviation Theory in understanding small tail-end probability, $p_L$, where the default threshold, $-\sqrt{L}C$, is small. But our main goal is not to use (4.9) to approximate $p_L$ because it is accurate only when $L$ is very large which isn't necessarily true in actual cases. The main goal is to ensure our importance sampling estimator's variance is minimized when $L$ is very large. We use this to justify our choice of importance sampling estimator. Now we will show

**Theorem 4.2.** *Let $u = -\sqrt{L}C$, then importance sampling from $\mathbb{P}_u(.)$ is efficient, i.e.*

$$R = \lim_{L\to\infty} \frac{1}{L} log\, F_L(-\sqrt{L}C) = -C^T\Sigma^{-1}C = 2H \tag{4.10}$$

*Proof.* Let $u = -\sqrt{L}C$. We start with the expression of $F_L(-\sqrt{L}C)$ in (4.2), and wish to use results from Theorem 4.1 with (4.6) and (4.8) to calculate $R$ but we cannot directly evaluate $\frac{1}{\sqrt{L}}Y$ and $(\frac{d\mathbb{P}}{d\mathbb{P}_u})^2$ in (4.2). Therefore in order to apply the Theorem 4.1 to calculate $R$, we change measure and then change variable in the calculations below. This allows us to rewrite $\frac{1}{\sqrt{L}}Y$ in the form of $\frac{1}{L}\sum_{i=1}^{L} X^{(i)}$ and reduce $(\frac{d\mathbb{P}}{d\mathbb{P}_u})^2$, so that we can have an expression in the form of (4.6) where we can apply Theorem 4.1 above. Based on (4.2) and expansion of $\frac{d\mathbb{P}}{d\mathbb{P}_u}$, we have the following:

$$\begin{aligned}
F_L(-\sqrt{L}C) &= \mathbb{E}^u\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y\leq -C)}(\frac{d\mathbb{P}}{d\mathbb{P}_u})^2\right\} \\
&= \mathbb{E}^u\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y\leq -C)}\left(\frac{exp\left\{-\frac{Y^T\Sigma^{-1}Y}{2}\right\}}{exp\left\{-\frac{(Y-u)^T\Sigma^{-1}(Y-u)}{2}\right\}}\right)^2\right\} \\
&= \mathbb{E}^u\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y\leq C)}exp\left\{2\sqrt{L}C^T\Sigma^{-1}Y + LC^T\Sigma^{-1}C\right\}\right\} \\
&= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^u\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y\leq -C)}exp\{2\sqrt{L}C^T\Sigma^{-1}Y\}\right\}
\end{aligned}$$

Change measure to $Y' \sim N(-u,\Sigma)$ under $\mathbb{P}_{-u}$. Now $F_L(-\sqrt{L}C) =$

$$\begin{aligned}
&= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^{-u}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y'\leq -C)}exp\{2\sqrt{L}C^T\Sigma^{-1}Y'\}\frac{d\mathbb{P}_u}{d\mathbb{P}_{-u}}\right\} \\
&= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^{-u}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y'\leq -C)}exp\{2\sqrt{L}C^T\Sigma^{-1}Y'\}\frac{exp\left\{\frac{-(Y'-u)^T\Sigma^{-1}(Y'-u)}{2}\right\}}{exp\left\{\frac{-(Y'+u)^T\Sigma^{-1}(Y'+u)}{2}\right\}}\right\} \\
&= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^{-u}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y'\leq -C)}exp\{2\sqrt{L}C^T\Sigma^{-1}Y'\}exp\{-2\sqrt{L}C^T\Sigma^{-1}Y'\}\right\} \\
&= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^{-u}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y'\leq -C)}\right\}
\end{aligned}$$

Change variable to $Y'' = Y' + u \sim N(0,\Sigma)$, then $\{\frac{1}{\sqrt{L}}Y' \leq -C\} = \{\frac{1}{\sqrt{L}}(Y'' - u) \leq -C\} = \{\frac{1}{\sqrt{L}}Y'' \leq -2C\}$ Then,

$$F_L(-\sqrt{L}C) = exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}^{(0)}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}Y''\leq -2C)}\right\} \tag{4.11}$$

Note that $X$ and $Y''$ are equal in distribution, so we can rewrite (4.11) as,

$$\begin{aligned} F_L(-\sqrt{L}C) &= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}X\leq -2C)}\right\} \\ &= exp\left\{LC^T\Sigma^{-1}C\right\}\mathbb{E}\left\{\mathbb{I}_{(\frac{1}{L}\sum_{i=1}^{L}X^{(i)}\leq -2C)}\right\} \end{aligned}$$

Now we can apply the above theorem with (4.6) and (4.8) to calculate $R$.

$$\begin{aligned} R &= \lim_{L\to\infty}\frac{1}{L}log\, F_L(-\sqrt{L}C) \\ &= \frac{1}{L}log\, exp\{LC^T\Sigma^{-1}C\} + \frac{1}{L}log\,\mathbb{E}\left\{\mathbb{I}_{(\frac{1}{L}\sum_{i=1}^{L}X^{(i)}\leq -2C)}\right\} \\ &= C^T\Sigma^{-1}C + -\frac{4}{2}C^T\Sigma^{-1}C \\ &= -C^T\Sigma^{-1}C \end{aligned} \tag{4.12}$$

Finally, with (4.9) and (4.12), we can conclude $R = 2H$, i.e. when $L$ is very large, $var(\hat{p_L})$,variance of our choice of importance sampling estimator (sampling from $\mathbb{P}_u \sim N(u = -\sqrt{L}C, \Sigma)$ instead of original measure $\mathbb{P} \sim N(0, \Sigma)$ is minimized and very close to zero, implying that it is efficient in the framework of Large Deviation Theory.

□

## 4.3 Efficiency under Gaussian Factor Copula

We have only considered efficiency of our important sampling estimator after changing measure to a another multivariate normal distribution with a new mean but the same covariance matrix. We formulated the problem using multivariate normal distribution instead of factors. Now we revisit our original Gaussian factor copula model and wish to change measure on each factor to produce the same results. Changing measure on the factors individually is trickier but will allow us to sample from univariate normal distributions and examine more closely each company as a factor.

*Proof.* We start with the same problem. First, let $Z_0, \cdots, Z_n$ be i.i.d. standard normal random variables $N(0,1)$. Let $\rho_1, ..., \rho_n \epsilon [0,1]$ Let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} \rho_1 & \sqrt{1-\rho_1^2} & & & \\ \rho_2 & & \sqrt{1-\rho_2^2} & & \\ \vdots & & & \ddots & \\ \rho_n & & & & \sqrt{1-\rho_n^2} \end{bmatrix} \begin{bmatrix} Z_0 \\ Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}$$

Then $X = \begin{bmatrix} X_1, \cdots, X_n \end{bmatrix}^T$ has multivariate normal distribution $N(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} 1 & \rho_1\rho_2 & \rho_1\rho_3 & \cdots & \rho_1\rho_n \\ & 1 & \rho_2\rho_3 & \cdots & \rho_2\rho_n \\ & & & \ddots & \vdots \\ & & & 1 & \rho_{n-1}\rho_n \\ & & & & 1 \end{bmatrix}$$

Our goal is to approximate

$$p = \mathbb{P}(X \leq D) = \mathbb{E}\{\mathbb{I}(X \leq D)\} \tag{4.13}$$

But instead of working with X as a multivariate distribution. We can perform Monte Carlo method on $Z_0, ..., Z_n$ as individual factors and rewrite $p$ as follows,

$$p = \mathbb{P}(X \leq D) = \mathbb{E}\left\{\prod_{i=0}^{n} \mathbb{I}_{(Z_i \leq b_i)}\right\} \tag{4.14}$$

for some $b_0, ..., b_n$. Solving for $b_i$ can be difficult, so we utilize Cholesky decomposition to simplify the problem and reduce the number of factors. We proceed as follows. Let $W_1, ..., W_n$ be i.i.d standard normal variables. Let $W = [W_1, ..., W_n]^T$. We wish to find $A$ such that $AW$ and $X$ are equal. We can perform Cholesky decomposition on variance-covarance matrix of $X \sim N(0, \Sigma)$ to get $\Sigma = AA^T$ where $A$ is an lower triangular matrix. This guarantees that $AW$ and $X$ are equal. Notice that we started with $n+1$ factors $Z_0, ..., Z_n$ and now only have to work with $n$ factors $W_1, ..., W_n$. Also, we know that $A$ is invertible. This will simplify our problem later. Thus, here and on, we work with $W_1, ..., W_n$ instead of $Z_0, ..., Z_n$. In this formulation, we can write $p$ as follows,

$$p \;=\; \mathbb{P}(X \leq D) = \mathbb{P}(AW \leq D) = \mathbb{E}\left\{\Pi_{i=1}^{n} \mathbb{I}_{(AW[i] \leq d_i)}\right\} \tag{4.15}$$

where $AW[i]$ is the i-th component of $AW$.
However, we run into the same issue when $d_1, ..., d_n$ are very negative. We turn to importance sampling once again. Because $W_1, ..., W_n$ are i.i.d. We can easily change measure on each factor individually. Suppose we change to $W_1', ..., W_n'$, which correspond to new measure $\mathbb{P}_{u_1}(.), ..., \mathbb{P}_{u_n}(.)$ respectively. Suppose $W_i' \sim N(u_i, 1)$ for $i = 1, ..., n$ Then we consider variance of this importance sampling estimator $\hat{p}$.

$$
\begin{aligned}
p &= \mathbb{E}^{(u_1,...,u_n)}\left\{\prod_{i=1}^{n}\left(\mathbb{I}_{(AW'[i]\leq d_i)}\frac{d\mathbb{P}}{d\mathbb{P}_{u_i}}\right)\right\} && (4.16)\\
\hat{p} &= \frac{1}{m}\sum_{j=1}^{L}\left(\prod_{i=1}^{n}\mathbb{I}_{(AW'[i]^{(j)}\leq d_i)}\frac{d\mathbb{P}}{d\mathbb{P}_{u_i}}\right) && (4.17)\\
var(\hat{p}) &= \frac{1}{m}\left(\mathbb{E}^{(u_1,...,u_n)}\left\{\left(\prod_{i=1}^{n}\mathbb{I}_{(AW'[i]^{(1)}\leq d_i')}\frac{d\mathbb{P}}{d\mathbb{P}_{u_i}}\right)^2\right\} - p^2\right) && (4.18)
\end{aligned}
$$

where $W'[i]^{(j)}$ for $j = 1, ..., m$ are i.i.d samples from $N(u_i, 1)$
To minimize variance, we just need to minimize

$$
\mathbb{E}^{(u_1,...,u_n)}\left\{\left(\prod_{i=1}^{n}\mathbb{I}_{(AW'[i]^{(1)}\leq d_i)}\frac{d\mathbb{P}}{d\mathbb{P}_{u_i}}\right)^2\right\} \quad (4.19)
$$

But such optimization is extremely complicated and has no closed-form solution, so we go back to the Large Deviation Theory framework and try to find $u_1, ..., u_n$ that makes our importance sampling estimator *efficient*.
Recall $D' = -\sqrt{L}A^{-1}C$. Intuitively, $D' = [d_1', ..., d_n']^T$ seems like a reasonable choice for $u_1, ..., u_n$. Let $W_1', ..., W_n'$ be normal random variables $N(u_1 = d_1', 1), ..., N(u_n = d_n', 1)$ under new measure respectively. Recall $X$ and $X'$ in. Note that in the multivariate framework,

$$
\begin{aligned}
X &= AW \sim N(0, \Sigma)\\
X' &= AW' \sim N(D = -\sqrt{L}C, \Sigma)
\end{aligned}
$$

Our goal is confirm that such choice will make our importance sampling estimator *efficient*. Based on (4.9), we already know that

$$
H = \lim_{L\to\infty}\frac{1}{L}log\, p_L = -\frac{1}{2}C^T\Sigma^{-1}C
$$

It remains for us to show that $R = \lim_{L\to\infty} \frac{1}{L} log F_L(-\sqrt{L}C) = 2H$ in this factor framework. Recall that,

$$F_L(-\sqrt{L}C) = \mathbb{E}\left\{\left(\mathbb{I}_{(X\leq-\sqrt{L}C)}\right)^2\right\}$$

After changing measure on individual factors to $\mathbb{P}_{d'_1}, ..., \mathbb{P}_{d'_n}$ with $W'_i \sim N(d'_i, 1)$ for $i = 1, ..., n$

$$\begin{aligned}
F_L(-\sqrt{L}C) &= \mathbb{E}^{(d'_1,...,d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'[i]\leq d_i)}\left(\frac{d\mathbb{P}}{d\mathbb{P}_{d'_i}}\right)^2\right\} \\
&= \mathbb{E}^{(d'_1,...,d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'[i]\leq d'_i)}\left(\frac{exp(\frac{1}{2}W_i'^2)}{exp(\frac{1}{2}(W'_i - d'_i)^2}\right)^2\right\} \\
&= \mathbb{E}^{(d'_1,...,d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'[i]\leq d'_i)} exp(-2W'_i d'_i + d_i'^2)\right\} \\
&= \left(\prod_{i=1}^{n} exp(d_i'^2)\right)\mathbb{E}^{(d'_1,...,d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'[i]\leq d'_i)} exp(-2W'_i d'_i)\right\}
\end{aligned}$$

We change measure again on individual factors to $\mathbb{P}_{-d'_1}, ..., \mathbb{P}_{-d'_n}$ with $W''_i \sim N(-d'_i, 1)$ for $i = 1, ..., n$. Then,

$$\begin{aligned}
F_L(-\sqrt{L}C) &= \left(\prod_{i=1}^{n} exp(d_i'^2)\right)\mathbb{E}^{(-d'_1,...,-d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW''[i]\leq d_i)} exp(-2W''_i u_i)\frac{d\mathbb{P}_{d'_i}}{d\mathbb{P}_{-d'_i}}\right\} \\
&= \left(\prod_{i=1}^{n} exp(d_i'^2)\right)\mathbb{E}^{(-d'_1,...,-d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW''[i]\leq d_i)} exp(-2W''_i d'_i)\frac{exp(\frac{-(W''_i-d'_i)^2}{2})}{exp(\frac{-(W''_i+d'_i)^2}{2})}\right\} \\
&= \left(\prod_{i=1}^{n} exp(d_i'^2)\right)\mathbb{E}^{(-d'_1,...,-d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW''[i]\leq d_i)} exp(-2W''_i d'_i) exp(2W''_i d'_i)\right\} \\
&= \left(\prod_{i=1}^{n} exp(d_i'^2)\right)\mathbb{E}^{(-d'_1,...,-d'_n)}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW''[i]\leq d_i)}\right\}
\end{aligned}$$

We now change variable to $W'''_i = W''_i + d'_i \sim N(0, 1)$ for $i = 1, ..., n$. Note that

$W_i$ and $W_i'''$ are equal in distribution for $i = 1, ..., n$. Then

$$\begin{aligned} F_L(-\sqrt{L}C) &= \left(\prod_{i=1}^{n} exp({d_i'}^2)\right) \mathbb{E}^{(0,\ldots,0)} \left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'''[i]-d_i \le d_i)}\right\} \\ &= \left(\prod_{i=1}^{n} exp({d_i'}^2)\right) \mathbb{E}^{(0,\ldots,0)} \left\{\prod_{i=1}^{n} \mathbb{I}_{(AW'''[i] \le 2d_i)}\right\} \\ &= \left(\prod_{i=1}^{n} exp({d_i'}^2)\right) \mathbb{E} \left\{\prod_{i=1}^{n} \mathbb{I}_{(AW[i] \le 2d_i)}\right\} \end{aligned}$$

In multivariate framework, we can rewrite $F_L(-\sqrt{L}C)$ as follows

$$\begin{aligned} F_L(-\sqrt{L}C) &= \left(\prod_{i=1}^{n} exp({d_i'}^2)\right) \mathbb{E} \left\{\mathbb{I}_{(AW \le -2\sqrt{L}C)}\right\} \\ &= \left(\prod_{i=1}^{n} exp({d_i'}^2)\right) \mathbb{E} \left\{\mathbb{I}_{(X \le -2\sqrt{L}C)}\right\} \end{aligned}$$

Now we can apply results in (4.6) and (4.8) to calculate $R$. Let $X^{(i)} \sim N(0, \Sigma)$ for $i = 1, ..., L$ be i.i.d random variables. Recall that $\frac{1}{\sqrt{L}}X$ and $\frac{1}{L}\sum_{i=1}^{L} X^{(i)}$ are equal in distribution. Thus,

$$
\begin{aligned}
R &= \lim_{L\to\infty} \frac{1}{L} log\ F_L(-\sqrt{L}C) \\
&= \frac{1}{L}\sum_{i=1}^{n} {d'_i}^2 + \frac{1}{L} log \mathbb{E}\left\{\mathbb{I}_{(X\le -2\sqrt{L}C)}\right\} \\
&= \frac{1}{L}\sum_{i=1}^{n} {d'_i}^2 + \frac{1}{L} log \mathbb{E}\left\{\mathbb{I}_{(\frac{1}{\sqrt{L}}X\le -2C)}\right\} \\
&= \frac{1}{L}\sum_{i=1}^{n} {d'_i}^2 + \frac{1}{L} log \mathbb{E}\left\{\mathbb{I}_{(\frac{1}{L}\sum_{i=1}^{n} X^{(i)}\le -2C)}\right\} \\
&= \frac{1}{L} {D'}^T D' + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= \frac{1}{L}\left(-\sqrt{L}C^T(A^{-1})^t\right)\left(-\sqrt{L}A^{-1}C\right) + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= C^T(A^{-1})^t A^{-1}C + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= C^T(A^T)^{-1} A^{-1}C + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= C^T(AA^T)^{-1}C + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= C^T\Sigma^{-1}C + -\frac{4}{2} C^T\Sigma^{-1}C \\
&= -C^T\Sigma^{-1}C
\end{aligned}
$$

Now we can conclude $R = 2H$, i.e. our important sampling estimator under the factor model is *efficient*.

□

# Chapter 5

# Application I: Joint Default Probability

## 5.1 Under Gaussian Copula

We now apply this importance sampling scheme to evaluating joint default probability $p$. We will use the setup and results in Section 4.3 but specify the conditions, namely, $\Sigma$ and $D'$. First we will assume $\Sigma$ follows the factor copula model described in Section 2.4 and $\rho_i$ for $i = 1, ..., n$ are all equal. Let $D = [d_1, ..., d_n]^T = [\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. Now we proceed to specify $D'$. Let $X \sim N(0, \Sigma)$. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Then we know $X$ and $AW$ are equal in distribution, where $W = [W_1, ..., W_n]^T$ and $W_i$ for $i = 1, ..., n$ are i.i.d $N(0, 1)$. Let $D' = [d'_1, ..., d'_n]^T = A^{-1}[\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. This implies,

$$\{\tau_i \leq T\} = \{X_i \leq \Phi^{-1}(F_i(T))\} = \{AW[i] \leq d_i\}$$

$$
\begin{aligned}
p &= \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(\tau_i \le T)}\right\} \\
&= \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(X_i \le \Phi^{-1}(F_i(T)))}\right\} \\
&= \mathbb{E}\left\{\prod_{i=1}^{n} \mathbb{I}_{(AW[i] \le d_i)}\right\} \\
&= \mathbb{E}^{(d'_1,\ldots,d'_n)}\left\{\prod_{i=1}^{n} \left(\mathbb{I}_{(AW'[i] \le d_i)} \frac{d\mathbb{P}}{d\mathbb{P}_{d'_i}}\right)\right\} \qquad W'_i \sim N(d'_i, 1) \\
&= \mathbb{E}^{(d'_1,\ldots,d'_n)}\left\{\prod_{i=1}^{n} \left(\mathbb{I}_{(AW'[i] \le d_i)} exp\left(\frac{d'^2_i}{2} - W_i {d'}_i\right)\right)\right\}
\end{aligned}
$$

Then the importance sampling estimator will be

$$
\hat{p} = \frac{1}{m} \sum_{j=1}^{m} \prod_{i=1}^{n} \left(\mathbb{I}_{(AW'[i]^{(j)} \le d_i)} exp\left(\frac{d'^2_i}{2} - {W'}^{(j)}_i d'_i\right)\right)
$$

where for each $i$, ${W'}^{(j)}_i$ are i.i.d. samples of $W'_i$ for $j = 1, \ldots, m$

Let

$$
\hat{p}_k = \prod_{i=1}^{n} \left(\mathbb{I}_{(AW'[i]^{(j)} \le d_i)} exp\left(\frac{d'^2_i}{2} - {W'}^{(j)}_i d'_i\right)\right)
$$

Then standard error will be

$$
SE = \sqrt{var(\hat{p})}
$$

Recall in (3.2), calculating joint default probability is equivalent to evaluating the multivariate normal CDF.

$$
p \;=\; \Phi_{\Sigma}(\phi^{-1}(F_1(T)), \ldots, \phi^{-1}(F_n(T)))
$$

There is no close form solution but Genz and Bretz [8](1999) proposed Quasi MC to evaluate this CDF when $n$ is large. We will compare our importance sampling scheme to Quasi MC.

## 5.2 Under Student-T Copula

Result in Section 4.3 for efficient important sampling estimator is established under the Gaussian copula case. We will apply the same result to the Student-T copula case by performing conditional importance sampling. First let $X \sim$

$N(0,\Sigma)$. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Let $\chi^2_\nu$ be chi-square random variable with $\nu$ degrees of freedom. Then $S = [S_1, ..., S_n]^T = X\sqrt{\frac{\nu}{\chi^2_\nu}} \sim T_{\Sigma,\nu}$ is multivariate Student-T variable. This means $S_i$ are marginally Student-T variables, $S_i \sim t_\nu$. Here, we will use $S_i$ instead of $X_i$ for $i = 1, ..., n$ to represent the firms. Then for $i = 1, ..., n$,

$$\{\tau_i \leq T\} = \{S_i \leq t_\nu^{-1}(F_i(T))\}$$

We perform conditional importance sampling first by conditioning on $\chi^2_\nu$. Then

$$S\Big|\chi^2_\nu = X\sqrt{\frac{\nu}{\chi^2_\nu}}\Big|\chi^2_\nu \sim N\left(0, \frac{\nu}{\chi^2_\nu}\Sigma\right)$$

Let $A' = \sqrt{\frac{\nu}{\chi^2_\nu}}A$. Let $\Sigma' = \frac{\nu}{\chi^2_\nu}\Sigma$. Note that now, $S\Big|\chi^2_\nu = [S_1', ..., S_n'] \sim N(0,\Sigma')$ and $A'$ is Cholesky decomposition of $\Sigma'$ and $A'^{-1} = \sqrt{\frac{\chi^2}{\nu}}A^{-1}$. Let $W = [W_1, ..., W_n]^T$ where $W_1, ..., W_n$ are i.i.d. $N(0,1)$ Then $A'W$ and $S\Big|\chi^2_\nu$ are equal in distribution. Let $D = [d_1, ..., d_n]^T = [t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T))]^T$. Let $D' = [d_1', ..., d_n']^T = A'^{-1}[t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T))]^T$ Now, we are ready to formulate our double expectation and conditional importance sampling.

$$\begin{aligned}
p &= \mathbb{E}\left\{\prod_{i=1}^{n}\mathbb{I}_{(\tau_i \leq T)}\right\} \\
&= \mathbb{E}\left\{\prod_{i=1}^{n}\mathbb{I}_{\left(S_i \leq t_\nu^{-1}(F_i(T))\right)}\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\prod_{i=1}^{n}\mathbb{I}_{\left(S_i \leq t_\nu^{-1}(F_i(T))\right)}\Big|\chi^2_\nu\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\prod_{i=1}^{n}\mathbb{I}_{(AW[i] \leq d_i)}\Big|\chi^2_\nu\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{E}^{(d_1',...,d_n')}\left\{\prod_{i=1}^{n}\mathbb{I}_{(AW'[i] \leq d_i)}\frac{d\mathbb{P}}{d\mathbb{P}_{d_i'}}\Big|\chi^2_\nu\right\}\right\} && W_i' \sim N(d_i', 1) \\
&= \mathbb{E}\left\{\mathbb{E}^{(d_1',...,d_n')}\left\{\prod_{i=1}^{n}\mathbb{I}_{(AW'[i] \leq d_i)}exp\left(\frac{d_i'^2}{2} - W_i d_i'\right)\Big|\chi^2_\nu\right\}\right\}
\end{aligned}$$

Our importance sampling estimator will be,

$$\hat{p} = \frac{1}{m_2} \sum_{l=1}^{m_2} \frac{1}{m_1} \sum_{j=1}^{m_1} \left\{ \prod_{i=1}^{n} \mathbb{I}_{(AW'[i]^{(j)} \le d_i^{(l)})} exp \left( \frac{d'^{(l)2}_i}{2} - W_i^{(j)} d'^{(l)}_i \right) \Big| \chi_\nu^{2(l)} \right\}$$

where for each $i$, $W'^{(j)}_i$ are i.i.d. samples of $W'_i$ for $j = 1, ..., m_1$, and $d'^{(l)}_i$ for $l = 1, ..., m_2$ are samples of $d'_i$ which is dependent on samples of $\chi^2_\nu$ in the outer expectation. Let Standard error will be,

$$SE = \sqrt{var(\hat{p})}$$

Recall in (3.3), calculating joint default probability is equivalent to evaluating the multivariate student-T CDF.

$$p = T_{\Sigma,\nu}(t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T)))$$

There is no close form solution but Genz and Bretz [8] (1999) proposed Quasi MC to evaluate this CDF when n is large. We will compare our importance sampling scheme to Quasi MC.

## 5.3 Numerical Comparison

We compare the performance of Basic MC, Importance Sampling and Quasi MC under different scenarios for Gaussian copula and Stuent-T copula. First we take into consideration the rarity of the default event by testing different threshold D. Recall that $D = [d_1, ..., d_n]^T = [\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. But for simplicity, we let $d_1, ..., d_n$ be equal to one constant and call it D. For our purpose, we just need to test performance as D decreases, causing default to be more rare so we do not need to construct D from $\Phi(.)$ and $F_i(.)$ at the moment. Then we compare performance of different methods with different number of firms, $n$. For constructing covariance matrix $\Sigma$ we use the factor copula model and assume $\rho_1, ..., \rho_n$ are all equal and call it $\rho$.

In the following Tables,

- $D =$ default threshold for each firm.
- $\rho =$ correlation coefficient used to construct $\Sigma$ under factor copula model
- $n =$ number of firms

- $df$ = degree of freedom for chi-square variable under Student-T copula
- $m$ = number of iterations in Basic Monte Carlo and Importance Sampling method for Gaussian copula.
- $m_1$ = number of iterations in evaluating inner expectation under Conditional Importance Sampling for Student-T copula.
- $m_2$ = number of iterations in evaluating outer expectation under Conditional Importance Sampling for Student-T copula.

Under both Gaussian and Student-T copula, we can see importance sampling and Quasi MC both performed significantly better than Basic MC method. As $D$ gets smaller, Basic MC method is no longer capable of sampling from such rare events. Importance sampling and Quasi MC are comparable in terms of performance when $D$ gets very small and when $n$ increases. Quasi MC seems more accurate overall, but importance sampling is a simpler, easily modifiable and versatile approach. Given that the importance sampling method performs reasonably well under both Gaussian and Student-T copula, we take advantage of it's simplicity and later apply it to more complex problems such as calculating tail probability for some order statistics of default time, where Quasi MC method is not easily applicable.

Table 5.1: Estimating Joint Default Probability with Different Default Threshold under Gaussian Copula

| parameters | $n$ | $\rho$ | $m$ |
|---|---|---|---|
| | 5 | 0.5 | 25000 |

| | Basic MC | | Importance Sampling | | Quasi MC | |
|---|---|---|---|---|---|---|
| $D$ | $p$ | $SE$ | $p$ | $SE$ | $p$ | Error |
| 0.0 | 9.00E-02 | 1.81E-03 | 8.90E-02 | 1.80E-03 | 9.07E-02 | 6.91E-05 |
| -0.5 | 2.30E-02 | 9.47E-04 | 2.13E-02 | 4.61E-04 | 2.10E-02 | 5.61E-05 |
| -1.0 | 2.84E-03 | 3.37E-04 | 3.07E-03 | 8.25E-05 | 3.05E-03 | 1.36E-05 |
| -1.5 | 2.00E-04 | 8.94E-05 | 2.60E-04 | 9.01E-06 | 2.68E-04 | 1.76E-06 |
| -2.0 | 0.00E+00 | 0.00E+00 | 1.34E-05 | 5.53E-07 | 1.40E-05 | 1.05E-07 |
| -2.5 | 0.00E+00 | 0.00E+00 | 4.61E-07 | 2.65E-08 | 4.24E-07 | 4.51E-09 |
| -3.0 | 0.00E+00 | 0.00E+00 | 7.73E-09 | 5.39E-10 | 7.39E-09 | 2.37E-10 |
| -3.5 | 0.00E+00 | 0.00E+00 | 7.27E-11 | 6.26E-12 | 7.17E-11 | 9.68E-13 |
| -4.0 | 0.00E+00 | 0.00E+00 | 3.75E-13 | 4.05E-14 | 3.96E-13 | 5.64E-15 |
| -4.5 | 0.00E+00 | 0.00E+00 | 1.33E-15 | 1.46E-16 | 1.21E-15 | 1.41E-17 |
| -5.0 | 0.00E+00 | 0.00E+00 | 2.26E-18 | 3.86E-19 | 2.11E-18 | 4.19E-20 |
| -5.5 | 0.00E+00 | 0.00E+00 | 3.20E-21 | 6.69E-22 | 1.97E-21 | 5.82E-23 |
| -6.0 | 0.00E+00 | 0.00E+00 | 1.03E-24 | 1.85E-25 | 1.04E-24 | 4.29E-26 |
| -6.5 | 0.00E+00 | 0.00E+00 | 2.66E-28 | 4.20E-29 | 2.89E-28 | 1.09E-29 |
| -7.0 | 0.00E+00 | 0.00E+00 | 5.44E-32 | 1.28E-32 | 4.58E-32 | 1.48E-33 |
| -7.5 | 0.00E+00 | 0.00E+00 | 6.50E-36 | 2.29E-36 | 3.83E-36 | 1.52E-37 |
| -8.0 | 0.00E+00 | 0.00E+00 | 1.64E-40 | 5.34E-41 | 1.42E-40 | 1.35E-42 |
| -8.5 | 0.00E+00 | 0.00E+00 | 4.06E-45 | 9.79E-46 | 6.29E-46 | 8.76E-48 |
| -9.0 | 0.00E+00 | 0.00E+00 | 1.24E-49 | 7.58E-50 | 7.49E-52 | 9.33E-54 |
| -9.5 | 0.00E+00 | 0.00E+00 | 3.29E-55 | 9.99E-56 | 3.56E-58 | 1.89E-60 |
| -10.0 | 0.00E+00 | 0.00E+00 | 6.45E-61 | 2.46E-61 | 4.60E-65 | 2.34E-67 |

Table 5.2: Estimating Joint Default Probability with Different Number of Firms under Gaussian Copula

| parameters | $D$ | $\rho$ | $m$ |
|---|---|---|---|
| | -2 | 0.5 | 25000 |

| | Basic MC | | Importance Sampling | | Quasi MC | |
|---|---|---|---|---|---|---|
| $n$ | $p$ | $SE$ | $p$ | $SE$ | $p$ | Error |
| 5 | 4.00E-05 | 4.00E-05 | 1.35E-05 | 5.46E-07 | 1.40E-05 | 1.31E-07 |
| 6 | 0.00E+00 | 0.00E+00 | 4.74E-06 | 2.69E-07 | 4.77E-06 | 1.14E-07 |
| 7 | 0.00E+00 | 0.00E+00 | 1.78E-06 | 1.03E-07 | 1.85E-06 | 3.93E-08 |
| 8 | 0.00E+00 | 0.00E+00 | 7.50E-07 | 5.57E-08 | 8.12E-07 | 2.39E-08 |
| 9 | 0.00E+00 | 0.00E+00 | 3.96E-07 | 2.94E-08 | 3.81E-07 | 3.49E-08 |
| 10 | 0.00E+00 | 0.00E+00 | 2.13E-07 | 1.71E-08 | 2.01E-07 | 1.62E-08 |
| 11 | 0.00E+00 | 0.00E+00 | 1.02E-07 | 1.10E-08 | 1.07E-07 | 8.07E-09 |
| 12 | 0.00E+00 | 0.00E+00 | 8.55E-08 | 9.43E-09 | 6.30E-08 | 4.77E-09 |
| 13 | 0.00E+00 | 0.00E+00 | 3.43E-08 | 4.36E-09 | 3.65E-08 | 2.83E-09 |
| 14 | 0.00E+00 | 0.00E+00 | 2.09E-08 | 2.32E-09 | 2.24E-08 | 1.71E-09 |
| 15 | 0.00E+00 | 0.00E+00 | 1.67E-08 | 2.46E-09 | 1.52E-08 | 2.30E-09 |
| 16 | 0.00E+00 | 0.00E+00 | 8.83E-09 | 1.73E-09 | 9.77E-09 | 1.73E-09 |
| 17 | 0.00E+00 | 0.00E+00 | 6.17E-09 | 8.85E-10 | 7.61E-09 | 2.59E-09 |
| 18 | 0.00E+00 | 0.00E+00 | 4.41E-09 | 5.78E-10 | 4.61E-09 | 1.08E-09 |
| 19 | 0.00E+00 | 0.00E+00 | 2.94E-09 | 5.34E-10 | 3.80E-09 | 2.15E-09 |
| 20 | 0.00E+00 | 0.00E+00 | 2.54E-09 | 4.02E-10 | 2.56E-09 | 7.42E-10 |
| 21 | 0.00E+00 | 0.00E+00 | 1.47E-09 | 3.62E-10 | 1.64E-09 | 3.34E-10 |
| 22 | 0.00E+00 | 0.00E+00 | 1.45E-09 | 2.89E-10 | 1.35E-09 | 3.99E-10 |
| 23 | 0.00E+00 | 0.00E+00 | 1.28E-09 | 2.18E-10 | 1.07E-09 | 4.47E-10 |
| 24 | 0.00E+00 | 0.00E+00 | 1.05E-09 | 1.89E-10 | 6.23E-10 | 1.67E-10 |
| 25 | 0.00E+00 | 0.00E+00 | 4.42E-10 | 1.10E-10 | 7.26E-10 | 3.44E-10 |

Table 5.3: Estimating Joint Default Probability with Different Default Threshold under Student-T Copula

| parameters | $n$ | $\rho$ | $df$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|
| | 5 | 0.5 | 10 | 25000 | 10 | 2500 |

| | Basic MC | | Importance Sampling | | Quasi MC | |
|---|---|---|---|---|---|---|
| $D$ | $p$ | $SE$ | $p$ | $SE$ | $p$ | Error |
| 0.0 | 8.99E-02 | 1.81E-03 | 9.09E-02 | 1.80E-03 | 9.07E-02 | 6.43E-05 |
| -0.5 | 2.39E-02 | 9.66E-04 | 2.27E-02 | 5.38E-04 | 2.31E-02 | 9.38E-05 |
| -1.0 | 4.68E-03 | 4.32E-04 | 4.62E-03 | 1.70E-04 | 4.82E-03 | 8.65E-05 |
| -1.5 | 9.60E-04 | 1.96E-04 | 8.27E-04 | 4.67E-05 | 9.52E-04 | 4.39E-05 |
| -2.0 | 2.80E-04 | 1.06E-04 | 2.11E-04 | 2.27E-05 | 1.92E-04 | 1.19E-05 |
| -2.5 | 4.00E-05 | 4.00E-05 | 3.79E-05 | 7.12E-06 | 4.39E-05 | 6.97E-06 |
| -3.0 | 4.00E-05 | 4.00E-05 | 1.36E-05 | 3.25E-06 | 1.22E-05 | 3.72E-06 |
| -3.5 | 0.00E+00 | 0.00E+00 | 3.85E-06 | 1.23E-06 | 3.72E-06 | 2.51E-06 |
| -4.0 | 0.00E+00 | 0.00E+00 | 1.94E-06 | 1.13E-06 | 1.46E-06 | 1.06E-06 |
| -4.5 | 0.00E+00 | 0.00E+00 | 2.03E-07 | 6.64E-08 | 5.81E-07 | 5.18E-07 |
| -5.0 | 0.00E+00 | 0.00E+00 | 2.24E-08 | 1.79E-08 | 1.24E-07 | 1.76E-07 |
| -5.5 | 0.00E+00 | 0.00E+00 | 8.40E-08 | 8.02E-08 | 9.17E-09 | 9.45E-09 |
| -6.0 | 0.00E+00 | 0.00E+00 | 6.82E-09 | 6.36E-09 | 8.38E-08 | 2.49E-07 |
| -6.5 | 0.00E+00 | 0.00E+00 | 5.64E-09 | 5.04E-09 | 5.08E-09 | 8.56E-09 |
| -7.0 | 0.00E+00 | 0.00E+00 | 3.33E-11 | 3.05E-11 | 6.34E-10 | 8.39E-10 |
| -7.5 | 0.00E+00 | 0.00E+00 | 1.39E-11 | 1.33E-11 | 7.08E-09 | 2.45E-08 |
| -8.0 | 0.00E+00 | 0.00E+00 | 7.75E-09 | 7.75E-09 | 1.80E-11 | 2.68E-11 |
| -8.5 | 0.00E+00 | 0.00E+00 | 7.47E-14 | 7.45E-14 | 1.17E-10 | 3.65E-10 |
| -9.0 | 0.00E+00 | 0.00E+00 | 7.10E-15 | 5.42E-15 | 1.45E-11 | 2.58E-11 |
| -9.5 | 0.00E+00 | 0.00E+00 | 9.25E-13 | 9.25E-13 | 3.44E-12 | 8.76E-12 |
| -10.0 | 0.00E+00 | 0.00E+00 | 3.82E-13 | 3.80E-13 | 3.44E-10 | 1.18E-09 |

Table 5.4: Estimating Joint Default Probability with Different Number of Firms under Student-T Copula

| parameters | $D$ | $\rho$ | $df$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|
| | -2 | 0.5 | 10 | 25000 | 10 | 2500 |

| | Basic MC | | Importance Sampling | | Quasi MC | |
|---|---|---|---|---|---|---|
| $n$ | $p$ | $SE$ | $p$ | $SE$ | $p$ | Error |
| 5 | 3.20E-04 | 1.13E-04 | 2.01E-04 | 1.78E-05 | 2.00E-04 | 1.54E-05 |
| 6 | 0.00E+00 | 0.00E+00 | 9.36E-05 | 1.19E-05 | 1.00E-04 | 1.06E-05 |
| 7 | 0.00E+00 | 0.00E+00 | 5.30E-05 | 1.09E-05 | 5.64E-05 | 1.26E-05 |
| 8 | 0.00E+00 | 0.00E+00 | 3.46E-05 | 5.09E-06 | 2.91E-05 | 5.83E-06 |
| 9 | 0.00E+00 | 0.00E+00 | 1.72E-05 | 3.77E-06 | 2.11E-05 | 7.28E-06 |
| 10 | 0.00E+00 | 0.00E+00 | 1.74E-05 | 3.19E-06 | 1.17E-05 | 2.38E-06 |
| 11 | 0.00E+00 | 0.00E+00 | 1.08E-05 | 3.39E-06 | 8.91E-06 | 2.22E-06 |
| 12 | 0.00E+00 | 0.00E+00 | 8.29E-06 | 2.17E-06 | 6.01E-06 | 2.12E-06 |
| 13 | 0.00E+00 | 0.00E+00 | 3.79E-06 | 1.14E-06 | 4.56E-06 | 1.69E-06 |
| 14 | 0.00E+00 | 0.00E+00 | 1.96E-06 | 4.34E-07 | 3.37E-06 | 1.06E-06 |
| 15 | 0.00E+00 | 0.00E+00 | 2.42E-06 | 1.06E-06 | 2.21E-06 | 8.64E-07 |
| 16 | 0.00E+00 | 0.00E+00 | 9.02E-07 | 3.35E-07 | 3.41E-06 | 3.05E-06 |
| 17 | 0.00E+00 | 0.00E+00 | 1.14E-06 | 3.25E-07 | 1.58E-06 | 1.25E-06 |
| 18 | 0.00E+00 | 0.00E+00 | 1.82E-06 | 9.91E-07 | 1.09E-06 | 8.00E-07 |
| 19 | 0.00E+00 | 0.00E+00 | 9.04E-07 | 2.96E-07 | 6.47E-07 | 2.79E-07 |
| 20 | 0.00E+00 | 0.00E+00 | 6.33E-07 | 2.46E-07 | 6.01E-07 | 3.48E-07 |
| 21 | 0.00E+00 | 0.00E+00 | 4.14E-07 | 2.92E-07 | 5.63E-07 | 3.37E-07 |
| 22 | 0.00E+00 | 0.00E+00 | 3.53E-07 | 1.88E-07 | 5.44E-07 | 5.20E-07 |
| 23 | 0.00E+00 | 0.00E+00 | 5.44E-07 | 3.36E-07 | 3.54E-07 | 1.96E-07 |
| 24 | 0.00E+00 | 0.00E+00 | 2.51E-07 | 1.68E-07 | 2.62E-07 | 1.76E-07 |
| 25 | 0.00E+00 | 0.00E+00 | 4.75E-08 | 3.07E-08 | 3.58E-07 | 2.50E-07 |

# Chapter 6

# Application II: Basket Default Swap

## 6.1 Introduction to Basket Default Swaps

We now wish to apply our importance sampling scheme in evaluating joint default probability to evaluating multi-name credit derivatives. This is motivated by CYH's study [5] on BDS. CYH mainly employed conditional importance sampling under the Gaussian copula factor model. We will first correct his approach which seems to only consider the outer expectation. Then we suggest a different conditional importance sampling scheme which improves accuracy under specified conditions. In this case, we perform change of measure as suggest by results in Section 4.3. Finally, we introduce direct importance sampling based on the method used in evaluating joint default probability. Lastly, we will compare performance of these methods under Gaussian copula model. As an extension, we apply similar method to evaluating BDS under Student-T copula and compare it with Basic MC.

The mechanism of a credit default swap (CDS) is similar to that of an insurance. The protection buyer makes periodical premium payments (protection leg or PL) until some credit events happen. Then swap issuer compensates for the non-recovered part of the reference entities' notional amounts (default leg or DL). CDS provides credit protection only for a single underlying. Multi-name credit derivatives, such as BDS and CDO have gained increasing popularity in recent years because they extend credit protection to a pool of underlying. We focus on BDS in this study which provides protection to a pool of underlying until default of one underlying. A $k^{th}$-to-default BDS offers protection only against the event of the $k^{th}$ default on a pool of $n$ underlying.

Pricing BDS is equivalent to determining the fair premium a BDS buyer needs to pay. Under risk-neutral measure, the fair premium is determined by equating protection leg or PL to default leg or DL. Here we introduce some notations and assumptions of our pricing model.

- $n$: Number of names in one basket, usually 5 or 6.
- $T$: Terminal time of a BDS contract.
- $R$: Recovery rate.
- $M$: Notional amount.
- $\triangle_{j-1,j},\ j=1,2,\cdots,N$: The time increment $t_j - t_{j-1}$.
- $\tau_i$: Default time of the $i$th company.
- $\tau$: $k^{th}$ default time.
- $B(0,\tau)$: $\exp\left(-\int_o^\tau r(u)du\right)$, the discount factor, (or price of zero coupon bond) where $r(\cdot)$ denotes risk free interest rate.
- $prem$: Fair premium for protection buyer.

The default leg when $k^{th}$ default takes place is

$$DL = \mathbb{E}\left\{(1-R)\cdot M\cdot B(0,\tau)\cdot \mathbb{I}_{(\tau\le T)}\right\} \tag{6.1}$$

where $\mathbb{E}$ is the expectation under risk neutral measure. When $k^{th}$ default does not take place, then protection leg is

$$PL = \mathbb{E}\left\{\sum_{j=1}^{N}\triangle_{j-1,j}\cdot M\cdot prem\cdot B(0,t_j)\cdot \mathbb{I}_{(\tau\ge t_j)}\right\} \tag{6.2}$$

We equate DL (6.1) and PL (6.2) to derive premium:

$$prem = \frac{\mathbb{E}\left\{(1-R)\cdot B(0,\tau)\cdot \mathbb{I}(\tau\le T)\right\}}{\mathbb{E}\left\{\sum_{j=1}^{N}\triangle_{j-1,j}\cdot B(0,t_j)\cdot \mathbb{I}(\tau > t_j)\right\}}. \tag{6.3}$$

To be more precise, we take accrued interests into considerations, modifying (6.2) by:

$$PL^{acc} = \mathbb{E}\left\{\sum_{j=1}^{N}\left(\frac{\tau - t_{j-1}}{t_j - t_{j-1}}\triangle_{j-1,\,j}\right)\cdot M \cdot prem \cdot B(0,\tau)\cdot \mathbb{I}(t_{j-1}\leq \tau \leq t_j)\right\}, \tag{6.4}$$

when defaults happen between two payment dates.

To derive fair premium (or spread), we need to evaluate both PL and DL. As suggested by CYH, we will mainly focus on evaluation of DL, which is what mainly affects accurate evaluation of fair premium because default events are extremely rare.

## 6.2 Algorithms under Gaussian Copula

### 6.2.1 Basic Monte Carlo Method

Here we quickly present the Basic MC to evaluate DL. This method utilizes Cholesky decomposition on $\Sigma$ to decompose $X \sim N(0,\Sigma)$ into $W = [W_1, ..., W_n]^T$ where $W_1, ..., W_n$ are i.i.d. $N(0,1)$. Let $D = [d_1, ..., d_n]^T = [\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. We know $X \sim N(0,\Sigma)$. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Then we know $X$ and $AW$ are equal in distribution, where $W = [W_1, ..., W_n]^T$ and $W_i$ for $i = 1, ..., n$ are i.i.d $N(0,1)$.

Then we transform $W_i$ into default time $\tau_i$ and find $\tau$ as the $k^{th}$ order statistics on $\{\tau_1, ..., \tau_n\}$. Since, $\tau$ is dependent on $W_1, ..., W_n$ in that it is the $k^{th}$ order statistics of $\tau_1, ..., \tau_n$ and $\tau_i = F_i^{-1}\Phi(AW[i])$, where $AW[i]$ is the i-th element of $AW$ for $i = 1, ..., n$. We denote $\tau$ as $\tau(W)$ below. We can write DL as,

$$DL \;=\; \mathbb{E}\left\{(1-R)B(0,\tau(W))\mathbb{I}_{(\tau(W)\leq T)}\right\} \tag{6.5}$$

Now, we are ready to propose our Basic MC algorithm.

**Algorithm 6.1** (Basic Monte Carlo Method).

1. Perform Cholesky decomposition on $\Sigma$ to find $A$ such that $\Sigma = AA^T$.

2. Find $A^{-1}$

3. Let $D' = A^{-1}D$. let $d_i'$ be the i-th element of $D'$ for $i = 1, ..., n$.

4. For $j = 1$ to $m$,

   (1) Generate $W_i^{(j)}$ from $N(0,1)$.

   (2) Let $U_i^{(j)} = \Phi(AW[i]^{(j)})$ for $i = 1, ..., n$.

   (3) Let $\tau_i^{(j)} = F_i^{-1}(U_i^{(j)})$ for $i = 1, ..., n$.

   (4) Let $\tau^{(j)}$ be the k-th order statistic of $\left\{\tau_1^{(j)}, ..., \tau_n^{(j)}\right\}$.

5. Approximate expectation

$$\widehat{DL} = (1-R)\frac{1}{m}\sum_{j=1}^{m}\left(e^{-\int_0^{\tau^{(j)}} r(t)dt}\right)\mathbb{I}\left(\tau^{(j)} \leq T\right)$$

6. Calculate $SE$,

$$SE = \sqrt{Var(\widehat{DL})}$$

### 6.2.2 Conditioning on All Marginal Factors

Recall under factor model, we used common factor $Z_0$, and marginal factors $Z_1, ..., Z_n$, which are all i.i.d $N(0,1)$, to construct covariance matrix $\Sigma$. In this section, we first propose the corrected CYH algorithm, which is conditional on all marginal factors, and then suggest another approach which is conditional on common factor. Recall the marginal default time of firm $i$ can be written as $\tau_i = F_i^{-1}(\Phi(X_i))$. That is, conditional on $\{Z_1 = z_1, \cdots, Z_n = z_n\}$, the default event is simply characterized by $Z_0$:

$$\begin{aligned}
\{\tau_i \leq T\} &= \left\{F_i^{-1}(\Phi(X_i)) \leq T\right\} \\
&= \{\Phi(X_i) \leq F_i(T)\} \\
&= \left\{X_i \leq \Phi^{-1}(F_i(T))\right\} \\
&= \left\{\rho_i Z_0 + \sqrt{1-\rho_i^2}z_i \leq \Phi^{-1}(F_i(T))\right\} \\
&= \left\{Z_0 \leq \frac{\Phi^{-1}(F_i(T)) - \sqrt{1-\rho_i^2}z_i}{\rho_i}\right\}, \qquad (6.6)
\end{aligned}$$

and we define

$$b_i \triangleq \frac{\Phi^{-1}(F_i(T)) - \sqrt{1-\rho_i^2}z_i}{\rho_i}.$$

Consider the time of the $k^{th}$ default event $\{\tau \le T\}$,

$$\mathbb{I}_{(\tau \le T)} = 1 \Leftrightarrow \sum_{i=1}^{n} \mathbb{I}_{(\tau_i \le T)} \ge k \Leftrightarrow \sum_{i=1}^{n} \mathbb{I}_{(Z_0 \le b_i)} \ge k.$$

Let $b$ be the $(n-k+1)^{th}$ order statistics of $\{b_1, \cdots, b_n\}$. Then,

$$\mathbb{I}(\tau \le T) = 1 \Leftrightarrow \mathbb{I}(Z_0 \le b) = 1,$$

Now, DL can be written as,

$$\begin{aligned}
\text{DL} &= \mathbb{E}\left\{(1-R)B(0,\tau)\mathbb{I}(\tau \le T)\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{(1-R)B(0,\tau)\mathbb{I}(\tau \le T)\middle| Z_1 = z_1, \cdots, Z_n = z_n\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{(1-R)B(0,\tau)\mathbb{I}(Z_0 \le b)\middle| Z_1 = z_1, \cdots, Z_n = z_n\right\}\right\}
\end{aligned}$$

In evaluating DL, the rare default event is in essence $\{\tau \le T\}$, or $\{Z_0 \le b\}$. CYH proposed an effective importance sampling scheme that reduces variance, namely changing measure on $Z_0$. We will quickly establish a lemma that guarantees variance reduction under specified conditions, and apply it to our importance sampling scheme.

**Lemma 6.1.** *If a random variable $X$ has density function $f(x)$ and distribution function $F(x)$. Suppose we are interested in estimating $p = \mathbb{E}\{\psi(X)\mathbb{I}_{(X \le a)}\}$, let $V = \psi(X)\mathbb{I}_{(X \le a)}$ and we define likelihood ratio*

$$Q(x) = \begin{cases} \frac{f(x)}{f^*(x)} & \text{if } x \le a, \\ 0 & \text{elsewhere.} \end{cases}$$

*where $f^*(.)$ is defined as*

$$f^*(x) = \begin{cases} \frac{f(x)}{F(a)} & \text{if } x \le a, \\ 0 & \text{otherwise.} \end{cases}$$

*Then $V^* = \psi(X)\mathbb{I}(X \le a)Q(X)$ is an unbiased estimator of $p$ under $\mathbb{E}^*(\cdot)$ and $Var^*(V^*) \le Var(V)$.*

*Proof.* By the definition of $Q(x)$ and $f^*(x)$, we have

$$\begin{aligned}
\mathbb{E}^*\{\psi(X)\mathbb{I}(X \le a)Q(X)\} &= \int_{-\infty}^{a} \psi(X)F(a)f^*(x)dx \\
&= F(a)\int_{-\infty}^{a} \psi(X)f^*(x)dx \\
&= \mathbb{E}\{\psi(X)\mathbb{I}(X \le a)\},
\end{aligned}$$

where $\mathbb{E}^*$ is the expectation operator under $f^*(\cdot)$. This shows that $V^*$ is an unbiased estimator of $V$.

Its variance under the new measure, denoted by $\text{Var}^*(\cdot)$, is

$$\begin{aligned}
\text{Var}^*(V^*) &= \mathbb{E}^*\left\{\psi^2(X)\mathbb{I}^2(X \le a)Q^2(X)\right\} - \left(\mathbb{E}^*\left\{\psi(X)\mathbb{I}(X \le a)Q(X)\right\}\right)^2 \\
&= \int_{-\infty}^{a} \psi^2(X)F^2(a)f^*(x)dx - \left(\int_{-\infty}^{a} \psi(X)F(a)f^*(x)dx\right)^2 \\
&= \int_{-\infty}^{a} \psi^2(X)F^2(a)\frac{f(x)}{F(a)}dx - \left(\int_{-\infty}^{a} \psi(X)F(a)\frac{f(x)}{F(a)}dx\right)^2 \\
&= \int_{-\infty}^{a} \psi^2(X)F(a)f(x)dx - \left(\int_{-\infty}^{a} \psi(X)f(x)dx\right)^2 \\
&= F(a)\mathbb{E}\left\{\psi^2(X)\mathbb{I}_{(X\le a)}\right\} - \left(\mathbb{E}\left\{\psi(X)\mathbb{I}_{(X\le a)}\right\}\right)^2 \\
&\le F(a)\mathbb{E}\left\{\psi^2(X)\mathbb{I}_{(X\le a)}\right\} - F(a)\left(\mathbb{E}\left\{\psi(X)\mathbb{I}_{(X\le a)}\right\}\right)^2 \\
&= F(a)Var(V) \le Var(V)
\end{aligned}$$

The last inequalities hold because $F(a) \le 1$. When $F(a) \le 1$, then the last inequalities would all be strictly less than. □

By this lemma, an effective measure change of $Z_0$ would be:

$$\phi^*(z) = \begin{cases} \frac{\phi(z)}{\Phi(z)}, & \text{if } z \le b, \\ 0 & \text{otherwise.} \end{cases}$$

This means the likelihood ratio $Q(z)$ is

$$Q(z) = \begin{cases} \frac{\phi(z)}{\phi^*(z)} = \Phi(b), & \text{if } z \le b, \\ 0 & \text{otherwise.} \end{cases}$$

Condition on $Z_1 = z_1, ..., Z_n = z_n$, $\tau$ is dependent on $Z_0, z_1, ..., z_n$ in that $\tau$ is the $k^{th}$ order statistics of $\{\tau_1, ..., \tau_n\}$, and for $i = 1, ..., n$

$$\tau_i = F_i^{-1}(\Phi(\rho Z_0 + \sqrt{1-\rho^2}z_i))$$

. Thus we denote $\tau$ as $\tau(Z_0, z_1, ..., z_n)$ below. Now we can write DL as

$$\begin{aligned}
\text{DL} &= \mathbb{E}\left\{\mathbb{E}\left\{(1-R)B(0,\tau(Z_0,z_1,...,z_n))\mathbb{I}_{(Z_0\le b)}\middle| Z_1 = z_1,\cdots,Z_n = z_n\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{E}^*\left\{(1-R)B(0,\tau(Z_0,z_1,...,z_n)\Phi(b)\middle| Z_1 = z_1,\cdots,Z_n = z_n\right\}\right\} \quad (6.7)
\end{aligned}$$

Note that first $Z_1, ..., Z_n$ are i.i.d $N(0,1)$ samples. Then condition on $Z_1 = z_1, ..., Z_n = z_n$, $Z_0$ is randomly sampled from $\phi^*(\cdot)$. We can now estimate DL by conditional importance sampling estimator $\widehat{DL}$,

$$\frac{1}{m_2}\sum_{l=1}^{m_2}\frac{1}{m_1}\sum_{j=1}^{m_1}\left\{(1-R)B(0,\tau(Z_0^{(j,l)}, Z_1^{(l)}, ..., Z_n^{(l)})\Phi\left(b^{(l)}\right)\middle| Z_1^{(l)}, \cdots, Z_n^{(l)}\right\},$$

We now propose our importance sampling algorithm that is conditional on all marginal factors.

**Algorithm 6.2** (corrected CYH algorithm)**.**

1. Calculate $F_i(T)$. Recall that $F_i(.)$ is CDF of the i-th firm's default time, which is assumed to follow exponential distribution with intensity $\lambda_i$.

2. For $l = 1$ to $m_2$,

   (1) Generate $Z_i^{(l)}$ from $N(0,1)$, $i = 1, \cdots, n$.

   (2) Set
   $$b_i^{(l)} = \frac{\Phi^{-1}(F_i(T)) - \sqrt{1-\rho_i^2}Z_i^{(l)}}{\rho_i}$$

   (3) Set $b^{(l)}$ as the $(n-k+1)^{th}$ order statistics of $\left\{b_1^{(l)}, \cdots, b_n^{(l)}\right\}$.

   (4) Set $Q^{(l)} = \Phi\left(b^{(l)}\right)$. $\Phi(.)$ is CDF of $N(0,1)$.

   (5) For $j = 1$ to $m_1$,

      [1] Generate $U^{(j)}$ from $U(0,1)$.

      [2] Let $Z_0^{(j,l)} = \Phi^{-1}\left(Q^{(l)}U^{(j)}\right)$.

      [3] Let $W_i^{(j,l)} = \rho_i Z_0^{(j,l)} + \sqrt{1-\rho_i}Z_i^{(l)}$, for $i = 1, \cdots, n$.

      [4] Let $\tau_i^{(j,l)} = F_i^{-1}\left(\Phi\left(W_i^{(j,l)}\right)\right)$, for $i = 1, \cdots, n$.

      [5] Let $\tau^{(j,l)}$ as the $k^{th}$ order statistics of $\left\{\tau_1^{(j,l)}, \cdots, \tau_n^{(j,l)}\right\}$.

   (6) Approximate inner expectation
   $$\widehat{DL}^{(l)} = (1-R)Q^{(l)}\frac{1}{m_1}\sum_{j=1}^{m_1}\left(e^{-\int_0^{\tau^{(j,l)}} r(t)dt}\right).$$

3. Approximate outer expectation $\widehat{DL} = \frac{1}{m_2}\sum_{l=1}^{m_2}\widehat{DL}^{(l)}$.

4. Calculate
   $$SE = \sqrt{Var\left(\widehat{DL}\right)}$$

### 6.2.3 Conditioning on the Common Factor

Now we propose an alternative conditional importance sampling scheme to evaluating DL. We first condition on common factor $Z_0 = z_0$, and then change the probability measure of all marginal factors in the inner expectation. The default event of each underlying can be written as,

$$\begin{aligned}\{\tau_i \leq T\} &= \{F_i^{-1}(\Phi(X_i)) \leq T\} \\ &= \{X_i \leq \Phi^{-1}(F_i(T))\} \\ &= \left\{\rho_i z_0 + \sqrt{1-\rho_i^2} Z_i \leq \Phi^{-1}(F_i(T))\right\} \\ &= \left\{Z_i \leq \frac{\Phi^{-1}(F_i(T)) - \rho_i z_0}{\sqrt{1-\rho_i^2}}\right\}.\end{aligned}$$

Under factor model formulation, default events occur if

$$\tau_i = F_i^{-1}(\Phi(X_i)) \leq T \Leftrightarrow X_i \leq \Phi^{-1}(F_i(T)) \triangleq d_i.$$

If $d_i$ is very small, we shift mean of each $X_i$ to $d_i$ as suggested by Section 4.3. Condition on $Z_0 = z_0$, $X_1, ..., X_n$ are independent, and thus we can change measure on each $Z_i$ separately. Condition on $Z_0 = z_0$, $X_i \sim N(\rho z_0, 1 - \rho^2)$. Under new measure, condition on $Z_0 = z_0$, $X_i' \sim N(d_i, 1 - \rho_i^2)$ and $Z_i' \sim N(\rho_i z_0 - d_i, 1)$. For each $Z_i'$ likelihood ratio would be

$$Q_i(Z_i') = \frac{d\mathbb{P}}{d\mathbb{P}_{d_i}} \exp\left\{(\rho_i z_0 - d_i) Z_i' + (d_i - \rho_i z_0)^2/2\right\}.$$

Because $X_i'$ for $i = 1, ..., n$ are independent, we can take the product of the individual likelihood ratio.

$$Q(Z) = \prod_{i=1}^{n} \frac{d\mathbb{P}}{d\mathbb{P}_{d_i}} = \prod_{i=1}^{n} Q_i(Z_i').$$

Note that $\tau$ is dependent on $z_0, Z_1, ..., Z_N$ in that it is the $k^{th}$ order statistics of $\{\tau_1, ..., \tau_n\}$ and for $i = 1, ..., n$,

$$\tau_i = F_i^{-1}(\Phi(\rho z_0 + \sqrt{1-\rho^2} Z_i))$$

Thus we denote $\tau$ as $\tau(z_0, Z_1, ..., Z_n)$ below. Now we can rewrite DL as,

$$\begin{aligned}&\mathbb{E}\left\{\mathbb{E}\left\{(1-R)B(0, \tau(z_0, Z_1, ..., Z_n))\mathbb{I}_{(\tau(z_0, Z_1, ..., Z_n) \leq T)} \middle| Z_0 = z_0\right\}\right\} \\ =&\mathbb{E}\left\{\widetilde{\mathbb{E}}\left\{(1-R)B(0, \tau(z_0, Z_1, ..., Z_n))\mathbb{I}_{(\tau(z_0, Z_1, ..., Z_n) \leq T)} Q(Z) \middle| Z_0 = z_0\right\}\right\} \qquad (6.8)\end{aligned}$$

Now to estimate DL, the conditional importance sampling estimator $\widehat{DL}$ would be,

$$\frac{1}{m_2}\sum_{l=1}^{m_2}\frac{1}{m_1}\sum_{j=1}^{m_1}\left\{(1-R)B\left(0,\tau(Z_0^{(l)},Z_1^{(j,l)},...,Z_n^{(j,l)})\right)Q(Z)\mathbb{I}_{\left(\tau(Z_0^{(l)},Z_1^{(j,l)},...,Z_n^{(j,l)})\leq T\right)}\middle|Z_0^{(l)}\right\}$$

Here $Q(Z)=Q(Z_0^{(l)},Z_1^{(j,l)},...,Z_n^{(j,l)})$.

**Algorithm 6.3** (conditional on the common factor)**.**

1. Set $d_i=\Phi^{-1}(F_i(T))$, for $i=1,2,\cdots,n$.
2. For $l=1$ to $m_2$
   (1) Sample $Z_0^{(l)}$ from $N(0,1)$.
   (2) For $j=1$ to $m_1$
      [1] Sample $Z_i^{(j,l)}$ from $N(c_i-\rho_i Z_0^{(l)},1)$, for $i=1,2,...,n$.
      [2] Let
      $$Q_i^{(j,l)}=\exp\left\{\left(\rho_i Z_0^{(l)}-c_i\right)Z_i^{(j,l)}+\left(c_i-\rho_i Z_0^{(l)}\right)^2/2\right\},\ \text{and}Q^{(j,l)}=\prod_{i=1}^{n}Q_i^{(j,l)}.$$
      [3] Let $X_i^{(j,l)}=\rho_i Z_0^{(l)}+\sqrt{1-\rho_i}Z_i^{(j,l)}$, for $i=1,...,n$.
      [4] Let $\tau_i^{(j,l)}=F_i^{-1}\left(\Phi\left(W_i^{(j,l)}\right)\right)$, for $i=1,2,...,n$.
      [5] Let $\tau^{(j,l)}$ as the $k^{th}$ order statistics of $\left\{\tau_i^{(j,l)},...,\tau_n^{(j,l)}\right\}$.
   (3) Evaluate
   $$\widehat{DL}^{(l)}=(1-R)\frac{1}{m_1}\sum_{j=1}^{m_1}\left(e^{-\int_0^{\tau^{(j,l)}}r(t)dt}\right)\mathbb{I}\left(\tau^{(j,l)}\leq T\right)Q^{(j,l)}.$$
3. Evaluate $\widehat{DL}=\frac{1}{m_2}\sum_{l=1}^{m_2}\widehat{DL}^{(l)}$.
4. Calculate
$$SE=\sqrt{Var\left(\widehat{DL}\right)}$$

### 6.2.4 Direct Importance Sampling

Here we propose an algorithm for direct importance sampling to evaluate $DL$. This method utilizes Cholesky decomposition on $\Sigma$ to decompose $X \sim N(0, D)$ into $W = [W_1, ..., W_n]^T$ where $W_1, ..., W_n$ are i.i.d. $N(0, 1)$. Then we perform change of measure as suggested by result in Section 4.3.

Let $D = [d_1, ..., d_n]^T = [\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. We know $X \sim N(0, \Sigma)$. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Then we know $X$ and $AW$ are equal in distribution, where $W = [W_1, ..., W_n]^T$ and $W_i$ for $i = 1, ..., n$ are i.i.d $N(0, 1)$. Let $D' = [d'_1, ..., d'_n]^T = A^{-1}[\Phi^{-1}(F_1(T)), ..., \Phi^{-1}(F_n(T))]^T$. This implies,

$$\{\tau_i \leq T\} = \{X_i \leq \Phi^{-1}(F_i(T))\} = \{AW[i] \leq d_i\}$$

We know $\{\tau_i \leq T\}$, or $\{AW[i] \leq d_i\}$ is rare and therefore very hard to sample under original measure when $d_i$ is very negative. Thus we perform change of measure for $i = 1, ..., n$ and sample from $W'_i \sim N(d'_i, 1)$ instead. For each $W'_i$, likelihood ratio

$$Q_i = \frac{d\mathbb{P}}{d\mathbb{P}_{d_i}} = \frac{exp\left(-\frac{1}{2}W'^2_i\right)}{exp\left(-\frac{1}{2}(W'_i - d'_i)^2\right)}$$

Since $W_1, ..., W_n$ are independent. We can take the product of the individual likelihood ratios.

$$Q(W') = \prod_{i=1}^{n} \frac{d\mathbb{P}}{d\mathbb{P}_{d_i}} = \prod_{i=1}^{n} \frac{exp\left(-\frac{1}{2}W'^2_i\right)}{exp\left(-\frac{1}{2}(W'_i - d'_i)^2\right)}$$

Then we transform $W'_i$ into default time $\tau_i$ and find $\tau$ as the $k^{th}$ order statistics on $\{\tau_1, ..., \tau_n\}$. Since, $\tau$ is dependent on $W'_1, ..., W'_n$ in that it is the $k^{th}$ order statistics of $\tau_1, ..., \tau_n$ and $\tau_i = F_i^{-1}\Phi(AW'[i])$, where AW'[i] refers to the i-th element of AW' for $i = 1, ..., n$. We denote $\tau$ as $\tau(W'_1, ..., W'_n)$ below. We can write DL as,

$$\begin{aligned} DL &= \mathbb{E}\left\{(1-R)B(0, \tau(W_1, ..., W_n))\mathbb{I}_{(\tau(W_1, ..., W_n) \leq T)}\right\} \\ &= \mathbb{E}^{(d_1, ..., d_n)}\left\{(1-R)B(0, \tau(W'_1, ..., W'_n))\mathbb{I}_{\left(\tau(W'_1, ..., W'_n) \leq T\right)}Q(W)\right\} \end{aligned}$$

Now, we are ready to propose our direct change of measure algorithm.

**Algorithm 6.4** (Direct Change of Measure)**.**

1. Perform Cholesky decomposition on $\Sigma$ to find $A$ such that $\Sigma = AA^T$.

2. Find $A^{-1}$

3. Let $D' = A^{-1}D$. let $d'_i$ be the i-th element of $D'$ for $i = 1, ..., n$.

4. For $j = 1$ to $m$,

   (1) Generate $W_i'^{(j)}$ from $N(d'_i, 1)$.

   (2) Let
$$Q^{(j)} = \prod_{i=1}^{n} \frac{exp\left(-\frac{1}{2}W'^{(j)2}_i\right)}{exp\left(-\frac{1}{2}(W'^{(j)}_i - d'_i)^2\right)}$$

   (3) Let $\tau_i^{(j)} = F_i^{-1}(\Phi(W'^{(j)}_i))$ for $i = 1, ..., n$.

   (4) Let $\tau^{(j)}$ be the k-th order statistic of $\left\{\tau_1^{(j)}, ..., \tau_n^{(j)}\right\}$.

5. Approximate expectation
$$\widehat{DL} = (1-R)\frac{1}{m}\sum_{j=1}^{m}\left(e^{-\int_0^{\tau^{(j)}} r(t)dt}\right)\mathbb{I}\left(\tau^{(j)} \le T\right)Q^{(j)}$$
   Calculate $SE$,
$$SE = \sqrt{Var(\widehat{DL})}$$

### 6.2.5 Numerical Comparison

Here we present our main numerical result. We compare the DL estimates using the four different algorithms proposed in this study. We mainly consider the standard error ($SE$) of each estimate and calculate ratio by dividing it into the $SE$ of estimates under Basic MC. The goal is to use Basic MC's $SE$ as benchmark and examine each method's $SE$ reduction ratio. First we compare performance with different correlation strengths $\rho$ (Table 6.1), and then with different default intensity(Table 6.2). We limit each algorithm to the same number of total iterations, namely, 1,000,000. For conditional importance sampling, we take 10 iterations for the inner expectation and 100,000 for the outer expectation. Below are our results,

Table 6.1: Estimating DL under Different Correlation Strengths

| parameters | $n$ | $k$ | $T$ | $\lambda^{-1}$ | $r$ | $R$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|---|---|---|
| | 5 | 3 | 2 | 100 | 0.05 | 0.4 | 1000000 | 10 | 100000 |

| | Basic MC | | Condition on Marginal Factors | | | Condition on Common Factor | | | Direct Change of Measure | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho$ | $DL$ | $SE$ | $DL$ | $SE$ | Ratio | $DL$ | $SE$ | Ratio | $DL$ | $SE$ | Ratio |
| 0.05 | 4.18E-05 | 4.83E-06 | 3.94E-05 | 1.36E-05 | 0.35 | 4.25E-05 | 1.59E-06 | 3.04 | 4.65E-05 | 5.95E-06 | 0.81 |
| 0.10 | 5.23E-05 | 5.40E-06 | 3.58E-05 | 1.03E-05 | 0.53 | 4.94E-05 | 3.53E-06 | 1.53 | 4.96E-05 | 2.40E-06 | 2.25 |
| 0.15 | 6.64E-05 | 6.09E-06 | 5.36E-05 | 1.08E-05 | 0.56 | 5.78E-05 | 2.13E-06 | 2.85 | 6.16E-05 | 2.52E-06 | 2.42 |
| 0.20 | 7.70E-05 | 6.56E-06 | 8.16E-05 | 1.18E-05 | 0.56 | 7.29E-05 | 2.10E-06 | 3.12 | 7.42E-05 | 1.39E-06 | 4.71 |
| 0.25 | 1.03E-04 | 7.56E-06 | 1.28E-04 | 1.43E-05 | 0.53 | 1.08E-04 | 4.50E-06 | 1.68 | 1.06E-04 | 2.67E-06 | 2.83 |
| 0.30 | 1.51E-04 | 9.19E-06 | 1.55E-04 | 1.24E-05 | 0.74 | 1.50E-04 | 4.88E-06 | 1.88 | 1.56E-04 | 2.71E-06 | 3.39 |
| 0.35 | 2.26E-04 | 1.13E-05 | 2.31E-04 | 1.33E-05 | 0.85 | 2.31E-04 | 7.29E-06 | 1.54 | 2.29E-04 | 2.92E-06 | 3.85 |
| 0.40 | 3.34E-04 | 1.37E-05 | 3.53E-04 | 1.45E-05 | 0.95 | 3.31E-04 | 1.00E-05 | 1.37 | 3.38E-04 | 3.20E-06 | 4.27 |
| 0.45 | 4.98E-04 | 1.67E-05 | 5.13E-04 | 1.39E-05 | 1.20 | 5.04E-04 | 1.42E-05 | 1.18 | 4.98E-04 | 2.96E-06 | 5.64 |
| 0.50 | 7.22E-04 | 2.01E-05 | 7.27E-04 | 1.44E-05 | 1.40 | 7.47E-04 | 2.23E-05 | 0.90 | 7.26E-04 | 3.40E-06 | 5.92 |
| 0.55 | 1.07E-03 | 2.45E-05 | 1.06E-03 | 1.61E-05 | 1.52 | 1.06E-03 | 2.96E-05 | 0.83 | 1.05E-03 | 4.08E-06 | 6.02 |
| 0.60 | 1.54E-03 | 2.94E-05 | 1.51E-03 | 1.67E-05 | 1.76 | 1.51E-03 | 4.13E-05 | 0.71 | 1.50E-03 | 4.67E-06 | 6.29 |
| 0.65 | 2.07E-03 | 3.41E-05 | 2.06E-03 | 1.73E-05 | 1.97 | 1.99E-03 | 5.60E-05 | 0.61 | 2.07E-03 | 5.57E-06 | 6.13 |
| 0.70 | 2.76E-03 | 3.94E-05 | 2.81E-03 | 1.83E-05 | 2.15 | 2.80E-03 | 7.59E-05 | 0.52 | 2.80E-03 | 6.65E-06 | 5.93 |
| 0.75 | 3.67E-03 | 4.55E-05 | 3.69E-03 | 1.84E-05 | 2.46 | 3.66E-03 | 9.53E-05 | 0.48 | 3.71E-03 | 7.78E-06 | 5.84 |
| 0.80 | 4.89E-03 | 5.25E-05 | 4.84E-03 | 1.93E-05 | 2.72 | 4.72E-03 | 1.21E-04 | 0.43 | 4.81E-03 | 9.26E-06 | 5.66 |
| 0.85 | 6.15E-03 | 5.88E-05 | 6.11E-03 | 1.83E-05 | 3.21 | 6.25E-03 | 1.65E-04 | 0.36 | 6.13E-03 | 1.10E-05 | 5.37 |
| 0.90 | 7.73E-03 | 6.59E-05 | 7.65E-03 | 1.70E-05 | 3.88 | 7.42E-03 | 1.86E-04 | 0.35 | 7.64E-03 | 1.29E-05 | 5.13 |
| 0.95 | 9.29E-03 | 7.22E-05 | 9.30E-03 | 1.32E-05 | 5.45 | 1.00E-02 | 5.31E-04 | 0.14 | 9.40E-03 | 1.49E-05 | 4.84 |

In Table 6.1 and Figure 6.1, we can observe that under different correlation strengths, $\rho$ the conditional methods performed quite differently. The conditional on marginal factors method performs well when $\rho$ is high and the conditional on common factor method performs better when $\rho$ is low. This makes intuitive sense as well. For the condition on marginal factors method, when $\rho$ is high, more weight is placed on the common factor, which means the common factor plays a more significant role in determining default. Therefore, performing importance sampling on the common factor after conditioning on the marginal factors produces good error reduction results. On the other hand, when $\rho$ is low, more weight is placed on the marginal factors. Therefore, we perform importance sampling on them after conditioning on the common factor. Also we can see that direct importance sampling produced overall more accurate estimates and smaller $SE$. This method does not treat common and marginal factors separately, but performs importance sampling directly. Therefore, we can observe that it is not affected by changes in correlation. Table 6.1 shows that direct importance sampling is overall more

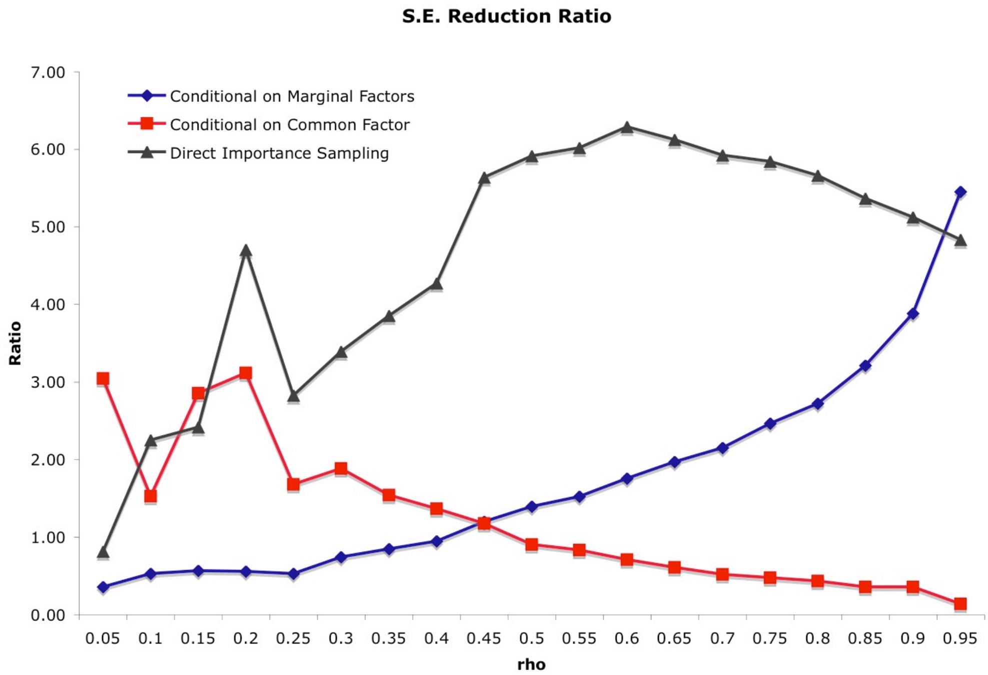


Figure 6.1: $SE$ Reduction Ratio under Different $\rho$

accurate and reliable while the other two conditional importance sampling methods perform well under more extreme conditions. Therefore, one can apply direct importance sampling in typical cases and reserve the conditional methods for more extreme cases in terms of correlation.

Table 6.2: Estimating DL under Different Default Intensity

| parameters | $n$ | $k$ | $T$ | $\rho$ | $r$ | $R$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|---|---|---|
| | 5 | 3 | 2 | 0.5 | 0.05 | 0.4 | 1000000 | 10 | 100000 |

| | Basic MC | | Condition on Marginal Factors | | | Condition on Common Factor | | | Direct Change of Measure | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda^{-1}$ | $DL$ | $SE$ | $DL$ | $SE$ | Ratio | $DL$ | $SE$ | Ratio | $DL$ | $SE$ | Ratio |
| 20.00 | 1.54E-02 | 9.20E-05 | 1.55E-02 | 1.13E-04 | 0.81 | 1.53E-02 | 1.41E-04 | 0.65 | 1.54E-02 | 3.50E-05 | 2.63 |
| 40.00 | 4.27E-03 | 4.88E-05 | 4.35E-03 | 4.99E-05 | 0.98 | 4.17E-03 | 6.26E-05 | 0.78 | 4.26E-03 | 1.30E-05 | 3.75 |
| 60.00 | 2.03E-03 | 3.37E-05 | 1.99E-03 | 2.81E-05 | 1.20 | 2.04E-03 | 4.17E-05 | 0.81 | 1.97E-03 | 7.74E-06 | 4.36 |
| 80.00 | 1.13E-03 | 2.52E-05 | 1.12E-03 | 1.91E-05 | 1.32 | 1.15E-03 | 2.80E-05 | 0.90 | 1.13E-03 | 4.90E-06 | 5.14 |
| 100.00 | 7.24E-04 | 2.01E-05 | 7.42E-04 | 1.53E-05 | 1.32 | 7.06E-04 | 1.96E-05 | 1.03 | 7.27E-04 | 3.31E-06 | 6.10 |
| 120.00 | 5.09E-04 | 1.69E-05 | 5.08E-04 | 1.14E-05 | 1.48 | 4.93E-04 | 1.67E-05 | 1.01 | 5.10E-04 | 2.64E-06 | 6.39 |
| 140.00 | 3.69E-04 | 1.44E-05 | 3.86E-04 | 9.40E-06 | 1.53 | 3.83E-04 | 1.46E-05 | 0.99 | 3.77E-04 | 2.07E-06 | 6.95 |
| 160.00 | 3.11E-04 | 1.32E-05 | 2.97E-04 | 9.10E-06 | 1.45 | 2.91E-04 | 1.17E-05 | 1.13 | 2.88E-04 | 1.83E-06 | 7.23 |
| 180.00 | 2.32E-04 | 1.14E-05 | 2.31E-04 | 6.33E-06 | 1.80 | 2.25E-04 | 1.02E-05 | 1.12 | 2.27E-04 | 1.47E-06 | 7.75 |
| 200.00 | 1.88E-04 | 1.03E-05 | 1.83E-04 | 5.59E-06 | 1.84 | 1.77E-04 | 8.06E-06 | 1.27 | 1.87E-04 | 1.24E-06 | 8.27 |
| 220.00 | 1.56E-04 | 9.37E-06 | 1.50E-04 | 4.54E-06 | 2.07 | 1.40E-04 | 7.04E-06 | 1.33 | 1.53E-04 | 1.07E-06 | 8.78 |
| 240.00 | 1.30E-04 | 8.53E-06 | 1.31E-04 | 4.51E-06 | 1.89 | 1.35E-04 | 7.05E-06 | 1.21 | 1.30E-04 | 8.79E-07 | 9.71 |
| 260.00 | 1.24E-04 | 8.36E-06 | 1.07E-04 | 3.57E-06 | 2.34 | 1.28E-04 | 1.48E-05 | 0.57 | 1.10E-04 | 8.28E-07 | 10.09 |
| 280.00 | 9.17E-05 | 7.18E-06 | 9.31E-05 | 3.89E-06 | 1.85 | 1.07E-04 | 6.74E-06 | 1.07 | 9.55E-05 | 1.08E-06 | 6.67 |
| 300.00 | 8.86E-05 | 7.05E-06 | 8.82E-05 | 3.53E-06 | 2.00 | 8.38E-05 | 5.30E-06 | 1.33 | 8.17E-05 | 6.03E-07 | 11.69 |
| 320.00 | 8.31E-05 | 6.83E-06 | 6.96E-05 | 2.38E-06 | 2.87 | 6.99E-05 | 5.06E-06 | 1.35 | 7.22E-05 | 5.74E-07 | 11.90 |
| 340.00 | 7.29E-05 | 6.39E-06 | 6.41E-05 | 2.71E-06 | 2.36 | 7.92E-05 | 8.11E-06 | 0.79 | 6.47E-05 | 5.27E-07 | 12.14 |
| 360.00 | 5.65E-05 | 5.63E-06 | 5.46E-05 | 2.13E-06 | 2.65 | 5.49E-05 | 4.59E-06 | 1.23 | 5.70E-05 | 4.91E-07 | 11.45 |
| 380.00 | 5.57E-05 | 5.59E-06 | 5.62E-05 | 2.93E-06 | 1.91 | 4.82E-05 | 3.76E-06 | 1.49 | 5.12E-05 | 4.53E-07 | 12.34 |
| 400.00 | 4.37E-05 | 4.95E-06 | 4.61E-05 | 2.01E-06 | 2.46 | 4.35E-05 | 3.62E-06 | 1.37 | 4.62E-05 | 3.92E-07 | 12.63 |
| 420.00 | 4.31E-05 | 4.92E-06 | 4.12E-05 | 1.97E-06 | 2.50 | 5.05E-05 | 4.24E-06 | 1.16 | 4.16E-05 | 3.65E-07 | 13.46 |
| 440.00 | 3.54E-05 | 4.46E-06 | 3.78E-05 | 1.90E-06 | 2.34 | 3.89E-05 | 3.09E-06 | 1.44 | 3.84E-05 | 5.21E-07 | 8.56 |
| 460.00 | 2.52E-05 | 3.76E-06 | 3.20E-05 | 1.34E-06 | 2.80 | 3.62E-05 | 2.86E-06 | 1.32 | 3.53E-05 | 4.03E-07 | 9.35 |
| 480.00 | 3.44E-05 | 4.41E-06 | 3.16E-05 | 1.68E-06 | 2.63 | 3.29E-05 | 3.28E-06 | 1.35 | 3.20E-05 | 2.79E-07 | 15.81 |
| 500.00 | 3.13E-05 | 4.19E-06 | 2.94E-05 | 1.78E-06 | 2.35 | 3.00E-05 | 2.89E-06 | 1.45 | 2.95E-05 | 2.66E-07 | 15.72 |
| 520.00 | 2.52E-05 | 3.76E-06 | 2.72E-05 | 1.46E-06 | 2.57 | 2.55E-05 | 2.43E-06 | 1.54 | 2.78E-05 | 2.55E-07 | 14.70 |
| 540.00 | 2.35E-05 | 3.63E-06 | 2.73E-05 | 1.83E-06 | 1.98 | 2.22E-05 | 2.14E-06 | 1.70 | 2.58E-05 | 4.39E-07 | 8.27 |
| 560.00 | 2.30E-05 | 3.59E-06 | 2.42E-05 | 1.28E-06 | 2.80 | 2.25E-05 | 1.91E-06 | 1.88 | 2.41E-05 | 2.58E-07 | 13.88 |
| 580.00 | 2.13E-05 | 3.46E-06 | 2.44E-05 | 1.95E-06 | 1.78 | 2.19E-05 | 2.55E-06 | 1.36 | 2.22E-05 | 2.45E-07 | 14.14 |
| 600.00 | 1.69E-05 | 3.08E-06 | 1.84E-05 | 8.91E-07 | 3.46 | 2.40E-05 | 2.99E-06 | 1.03 | 2.07E-05 | 2.41E-07 | 12.79 |
| 620.00 | 1.68E-05 | 3.07E-06 | 1.93E-05 | 1.05E-06 | 2.92 | 1.64E-05 | 1.46E-06 | 2.10 | 1.98E-05 | 2.84E-07 | 10.81 |
| 640.00 | 1.81E-05 | 3.20E-06 | 1.72E-05 | 1.01E-06 | 3.16 | 1.52E-05 | 1.32E-06 | 2.42 | 1.86E-05 | 4.07E-07 | 7.85 |
| 660.00 | 1.90E-05 | 3.26E-06 | 1.74E-05 | 8.43E-07 | 3.87 | 1.76E-05 | 2.34E-06 | 1.39 | 1.69E-05 | 1.75E-07 | 18.59 |
| 680.00 | 1.13E-05 | 2.52E-06 | 1.78E-05 | 1.34E-06 | 1.89 | 1.47E-05 | 1.62E-06 | 1.56 | 1.66E-05 | 2.55E-07 | 9.90 |
| 700.00 | 1.95E-05 | 3.30E-06 | 1.53E-05 | 9.86E-07 | 3.35 | 1.96E-05 | 2.56E-06 | 1.29 | 1.52E-05 | 2.26E-07 | 14.61 |
| 720.00 | 8.96E-06 | 2.24E-06 | 1.54E-05 | 1.21E-06 | 1.85 | 1.26E-05 | 1.27E-06 | 1.77 | 1.42E-05 | 1.76E-07 | 12.74 |
| 740.00 | 1.52E-05 | 2.92E-06 | 1.41E-05 | 9.66E-07 | 3.03 | 1.38E-05 | 1.93E-06 | 1.52 | 1.34E-05 | 1.46E-07 | 19.96 |
| 760.00 | 1.29E-05 | 2.68E-06 | 1.26E-05 | 6.38E-07 | 4.20 | 1.57E-05 | 3.07E-06 | 0.87 | 1.29E-05 | 1.74E-07 | 15.37 |
| 780.00 | 1.07E-05 | 2.46E-06 | 1.28E-05 | 1.09E-06 | 2.24 | 1.28E-05 | 1.56E-06 | 1.57 | 1.22E-05 | 1.44E-07 | 17.05 |
| 800.00 | 8.99E-06 | 2.25E-06 | 1.01E-05 | 6.04E-07 | 3.72 | 1.25E-05 | 1.60E-06 | 1.41 | 1.18E-05 | 1.61E-07 | 13.95 |

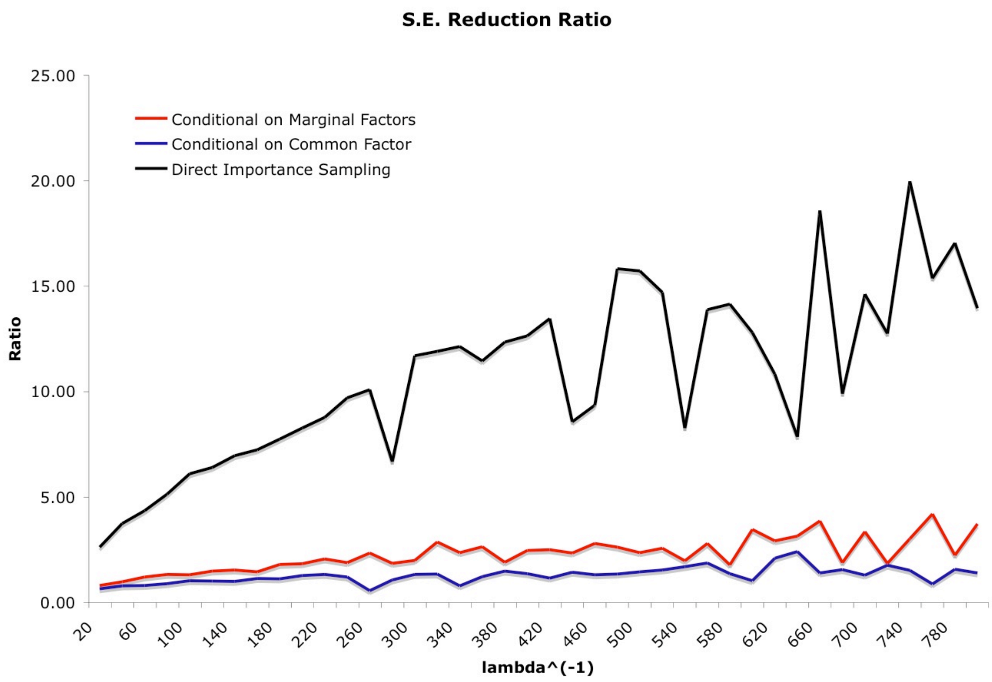


Figure 6.2: $SE$ Reduction Ratio under Different $\lambda^{-1}$

In Table 6.2 and Figure 6.2, we can observe that direct importance sampling's $SE$ reduction ratio increases as $\lambda^{-1}$ increases. First, $\lambda^{-1}$ increases means that default threshold, $D$ will be decrease. This is consistent with our proof in Section 4.3 which showed that as default threshold $D$ decreases, variance of our importance sampling estimator will approach zero, and thus making $SE$ smaller.

## 6.3 Algorithms under Student-T Copula

### 6.3.1 Basic Monte Carlo Method

For pricing BDS, we worked under Gaussian copula. But one needs to take into consideration the fat-tail behavior of each firm which is modeled by $X \sim N(0, \Sigma)$ in the Gaussian case. To further model fat-tail behavior, one would naturally extend Gaussian copula to Student-T copula. According to Cherubini, Luciano, and Vecchiato [4] (2004), this is common practice in industry. We will extend our direct change of measure algorithm to incorporate evaluating BDS under Student-T copula by performing conditional importance sampling. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Let $\chi^2_\nu$ be chi-square random variable with $\nu$ degrees of freedom.

Then $S = [S_1, ..., S_n]^T = X\sqrt{\frac{\nu}{\chi^2_\nu}} \sim T_{\Sigma,\nu}$ is multivariate Student-T variable. This means $S_i$ are marginally Student-T variables, $S_i \sim t_\nu$. $\Sigma$ still has the same structure as before. Here, we will use $S_i$ instead of $X_i$ for $i = 1, ..., n$ to represent the firms. This allows us to apply our algorithm under Gaussian copula. Note that now for $i = 1, ..., n$,

$$\{\tau_i \leq T\} = \{S_i \leq t_\nu^{-1}(F_i(T))\}$$

Here we quickly present the Basic MC to evaluate DL. This method utilizes Cholesky decomposition on $\Sigma$ to decompose $X \sim N(0, \Sigma)$ into $W = [W_1, ..., W_n]^T$ where $W_1, ..., W_n$ are i.i.d. $N(0, 1)$. Let $D = [d_1, ..., d_n]^T = [t_v^{-1}(F_1(T)), ..., t_v^{-1}(F_n(T))]^T$. We know $X \sim N(0, \Sigma)$. Let $A$ be Cholesky decomposition of $\Sigma$ such that $\Sigma = AA^T$. Then we know $X$ and $AW$ are equal in distribution, where $W = [W_1, ..., W_n]^T$ and $W_i$ for $i = 1, ..., n$ are i.i.d $N(0, 1)$.

Then we transform $W_i$ into default time $\tau_i$ and find $\tau$ as the $k^{th}$ order statistics on $\{\tau_1, ..., \tau_n\}$. Since, $\tau$ is dependent on $W_1, ..., W_n$ in that it is the $k^{th}$ order statistics of $\tau_1, ..., \tau_n$ and $\tau_i = F_i^{-1}(t_v(S_i))$, where $S \sim AW\sqrt{\frac{\nu}{\chi^2_\nu}}$ We denote $\tau$ as $\tau(W)$ below. We can write DL as,

$$DL = \mathbb{E}\left\{(1-R)B(0, \tau(W))\mathbb{I}_{(\tau(W) \leq T)}\right\} \tag{6.9}$$

Now, we are ready to propose our basic Monte Carlo algorithm.

**Algorithm 6.5** (Basic Monte Carlo Method).

1. Perform Cholesky decomposition on $\Sigma$ to find $A$ such that $\Sigma = AA^T$.

2. Find $A^{-1}$

3. Let $D' = A^{-1}D$. let $d'_i$ be the i-th element of $D'$ for $i = 1, ..., n$.

4. For $j = 1$ to $m$,

   (1) Generate $W_i^{(j)}$ from $N(0,1)$.

   (2) Generate Chi-Square variable with $\nu$ degrees of freedom, $\chi_\nu^{(j)2}$

   (3) Let $S^{(j)} = [S_1^{(j)}, ..., S_n^{(j)}]^T = AW\sqrt{\frac{\nu}{\chi_\nu^{(j)2}}}$

   (4) Let $U_i^{(j)} = t_v(S_i^{(j)})$ for $i = 1, ..., n$.

   (5) Let $\tau_i^{(j)} = F_i^{-1}(U_i^{(j)})$ for $i = 1, ..., n$.

   (6) Let $\tau^{(j)}$ be the k-th order statistic of $\left\{\tau_1^{(j)}, ..., \tau_n^{(j)}\right\}$.

5. Approximate expectation

$$\widehat{DL} = (1-R)\frac{1}{m}\sum_{j=1}^{m}\left(e^{-\int_0^{\tau^{(j)}} r(t)dt}\right)\mathbb{I}\left(\tau^{(j)} \leq T\right)$$

6. Let $\widehat{DL}^{(j)} = (1-R)\left(e^{-\int_0^{\tau^{(j)}} r(t)dt}\right)\mathbb{I}\left(\tau^{(j)} \leq T\right)$. Calculate $SE$,

$$SE = \sqrt{Var(\widehat{DL})}$$

### 6.3.2 Conditional Importance Sampling

Basic MC method is inaccurate when default is very rare, so we resort to importance sampling. We do not know how to perform change of measure efficiently with multivariate Student-T variables., but we can first condition on $\chi_\nu^2$, which then makes $S$ Gaussian. Then we can apply our results from the Gaussian case as before. This means we perform conditional importance sampling first by conditioning on $\chi_\nu^2$. Then

$$S\Big|\chi_\nu^2 = X\sqrt{\frac{\nu}{\chi_\nu^2}}\Big|\chi_\nu^2 \sim N\left(0, \frac{\nu}{\chi_\nu^2}\Sigma\right)$$

Let $A' = \sqrt{\frac{\nu}{\chi^2_\nu}}A$. Let $\Sigma' = \frac{\nu}{\chi^2_\nu}\Sigma$. Note that now, $S\Big|\chi^2_\nu = [S'_1, ..., S'_n] \sim N(0, \Sigma')$ and $A'$ is Cholesky decomposition of $\Sigma'$ and $A'^{-1} = \sqrt{\frac{\chi^2}{\nu}}A^{-1}$. Let $W = [W_1, ..., W_n]^T$ where $W_1, ..., W_n$ are i.i.d. $N(0,1)$ Then $A'W$ and $S\Big|\chi^2_\nu$ are equal in distribution. Let $D = [d_1, ..., d_n]^T = [t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T))]^T$. Let $D' = [d'_1, ..., d'_n]^T = A'^{-1}[t_\nu^{-1}(F_1(T)), ..., t_\nu^{-1}(F_n(T))]^T$ Now, we are ready to formulate our double expectation and conditional importance sampling. We can write DL as,

Condition on $\chi^2_\nu = \chi$, $\tau$ is dependent on $W$ in that it is the $k^{th}$ order statistics of $\{\tau_1, ..., \tau_n\}$ and $\tau_i = F_i^{-1}\Phi(A'W[i])$, where $A'W[i]$ is the i-th element of $A'W$ for $i = 1, ..., n$

$$\begin{aligned} DL &= \mathbb{E}\left\{(1-R)B(0,\tau)\mathbb{I}_{(\tau\le T)}\right\} \\ &= \mathbb{E}\left\{\mathbb{E}\left\{(1-R)B(0,\tau(W))\mathbb{I}_{(\tau(W)\le T)}\Big|\chi^2_\nu = \chi\right\}\right\} \\ &= \mathbb{E}\left\{\mathbb{E}^{(d'_1,...,d'_n)}\left\{(1-R)B(0,\tau(W'))\mathbb{I}_{(\tau(W')\le T)}Q(W')\Big|\chi^2_\nu = \chi\right\}\right\} \end{aligned} \tag{6.10}$$

Here, the likelihood ratio for change of measure is,

$$Q(W') = \prod_{i=1}^{n} \frac{exp\left(-\frac{1}{2}W'^2_i\right)}{exp\left(-\frac{1}{2}(W'_i - d'_i)^2\right)}$$

Now we are ready to present our conditional importance sampling algorithm,

**Algorithm 6.6** (conditional Student-T).

1. Perform Cholesky decomposition on $\Sigma$ to find $A$ such that $\Sigma = AA^T$.
2. Find $A^{-1}$
3. For $l = 1$ to $m_2$
   (1) Generate $\chi^{2(l)}_\nu$
   (2) Let $A'^{(l)} = \sqrt{\frac{\nu}{\chi^{2(l)}_\nu}}A$. Let $A'^{(l)-1} = \sqrt{\frac{\chi^{2(l)}_\nu}{\nu}}A^{-1}$.
   (3) Let $D'^{(l)} = A'^{(l)-1}D$ Let $d'^{(l)}_i$ be the i-th element of $D'^{(l)}$ for $i = 1, ..., n$.

(4) For $j = 1$ to $m_1$

[1] Generate $W'^{(j,l)}_i$ from $N(d'^{(l)}_i, 1)$.

[2] Let

$$Q^{(j,l)} = \prod_{i=1}^{n} \frac{exp\left(-\frac{1}{2}W'^{(j,l)2}_i\right)}{exp\left(-\frac{1}{2}(W'^{(j,l)}_i - d'^{(j)}_i)^2\right)}$$

[3] Let $\tau^{(j)}_i = F_i^{-1}(\Phi(W'^{(j,l)}_i))$ for $i = 1, ..., n$.

[4] Let $\tau^{(j)}$ be the k-th order statistic of $\left\{\tau^{(j)}_1, ..., \tau^{(j)}_n\right\}$.

(5) Approximate expectation

$$\widehat{DL}^{(l)} = (1-R)\frac{1}{m}\sum_{j=1}^{m}\left(e^{-\int_0^{\tau^{(j)}} r(t)dt}\right)\mathbb{I}\left(\tau^{(j)} \le T\right)Q^{(j)}$$

4. Let $\widehat{DL} = \frac{1}{m_2}\sum_{l=1}^{m}\widehat{DL}^{(l)}$

5. Calculate SE,

$$SE = \sqrt{Var(\widehat{DL})}$$

### 6.3.3 Numerical Comparison

We compare Basic MC and conditional importance sampling under Student-T copula. First we compare $SE$ of DL estimates under different correlation strengths, $\rho$, then under different default intensity $\lambda^{-1}$. We calculate $SE$ reduction ratio in the same way as before.

We can observe that conditional importance sampling is able to consistently reduce $SE$ in Table 6.3 and Table 6.4. We also discovered that $SE$ reduction ratio increases as degrees of freedom increased in Table 6.5. This implies that with very low degrees of freedom, it is better to use Basic MC while with higher degrees of freedom, it is better to use conditional importance sampling. However, in most cases, it is still better to use conditional importance sampling as $SE$ reduction ratio is greater than 1 for degrees of freedom greater than 3.

Table 6.3: Evaluating DL with Different Correlation Strengths

| parameters | $df$ | $n$ | $k$ | $T$ | $\lambda^{-1}$ | $r$ | $R$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 | 5 | 3 | 2 | 100 | 0.05 | 0.4 | 100000 | 10 | 10000 |

| | Basic MC | | Conditional Importance Sampling | | |
|---|---|---|---|---|---|
| $\rho$ | $DL$ | $SE$ | $DL$ | $SE$ | Ratio |
| 0.00 | 3.73E-04 | 4.59E-05 | 3.73E-04 | 3.01E-05 | 1.53 |
| 0.05 | 3.45E-04 | 4.41E-05 | 3.77E-04 | 3.65E-05 | 1.21 |
| 0.10 | 4.79E-04 | 5.20E-05 | 4.64E-04 | 3.51E-05 | 1.48 |
| 0.15 | 4.73E-04 | 5.16E-05 | 4.28E-04 | 3.64E-05 | 1.42 |
| 0.20 | 4.85E-04 | 5.23E-05 | 5.03E-04 | 3.67E-05 | 1.43 |
| 0.25 | 7.01E-04 | 6.29E-05 | 5.85E-04 | 3.39E-05 | 1.85 |
| 0.30 | 6.48E-04 | 6.05E-05 | 7.79E-04 | 5.00E-05 | 1.21 |
| 0.35 | 8.94E-04 | 7.11E-05 | 8.66E-04 | 4.05E-05 | 1.75 |
| 0.40 | 1.06E-03 | 7.75E-05 | 9.94E-04 | 3.96E-05 | 1.96 |
| 0.45 | 1.39E-03 | 8.86E-05 | 1.30E-03 | 4.76E-05 | 1.86 |
| 0.50 | 1.74E-03 | 9.91E-05 | 1.75E-03 | 5.90E-05 | 1.68 |
| 0.55 | 2.14E-03 | 1.10E-04 | 2.09E-03 | 5.84E-05 | 1.88 |
| 0.60 | 2.62E-03 | 1.21E-04 | 2.61E-03 | 6.82E-05 | 1.78 |
| 0.65 | 3.17E-03 | 1.34E-04 | 3.26E-03 | 7.78E-05 | 1.72 |
| 0.70 | 4.07E-03 | 1.52E-04 | 3.88E-03 | 8.45E-05 | 1.79 |
| 0.75 | 4.91E-03 | 1.66E-04 | 4.96E-03 | 1.02E-04 | 1.63 |
| 0.80 | 5.91E-03 | 1.83E-04 | 5.96E-03 | 1.10E-04 | 1.67 |
| 0.85 | 7.01E-03 | 1.99E-04 | 6.95E-03 | 1.18E-04 | 1.68 |
| 0.90 | 7.89E-03 | 2.11E-04 | 8.39E-03 | 1.37E-04 | 1.54 |
| 0.95 | 9.83E-03 | 2.35E-04 | 9.89E-03 | 1.56E-04 | 1.51 |

Table 6.4: Evaluating DL with Different Default Intensity

| parameters | $df$ | $n$ | $k$ | $T$ | $\rho$ | $r$ | $R$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 | 5 | 3 | 2 | 0.5 | 0.05 | 0.4 | 100000 | 10 | 10000 |

| | Basic MC | | Conditional Importance Sampling | | |
|---|---|---|---|---|---|
| $\lambda^{-1}$ | $DL$ | $SE$ | $DL$ | $SE$ | Ratio |
| 20.00 | 1.86E-02 | 3.19E-04 | 1.90E-02 | 2.31E-04 | 1.38 |
| 40.00 | 6.67E-03 | 1.93E-04 | 6.66E-03 | 1.27E-04 | 1.52 |
| 60.00 | 3.55E-03 | 1.41E-04 | 3.59E-03 | 9.20E-05 | 1.54 |
| 80.00 | 2.45E-03 | 1.18E-04 | 2.40E-03 | 7.12E-05 | 1.65 |
| 100.00 | 1.76E-03 | 9.98E-05 | 1.71E-03 | 5.65E-05 | 1.77 |
| 120.00 | 1.33E-03 | 8.69E-05 | 1.39E-03 | 5.30E-05 | 1.64 |
| 140.00 | 1.24E-03 | 8.40E-05 | 1.08E-03 | 4.17E-05 | 2.01 |
| 160.00 | 9.58E-04 | 7.37E-05 | 8.61E-04 | 3.44E-05 | 2.14 |
| 180.00 | 8.61E-04 | 6.98E-05 | 8.43E-04 | 4.06E-05 | 1.72 |
| 200.00 | 6.45E-04 | 6.04E-05 | 6.25E-04 | 2.99E-05 | 2.02 |
| 220.00 | 6.67E-04 | 6.14E-05 | 5.94E-04 | 2.97E-05 | 2.07 |
| 240.00 | 4.77E-04 | 5.21E-05 | 4.99E-04 | 2.73E-05 | 1.91 |
| 260.00 | 4.14E-04 | 4.85E-05 | 4.92E-04 | 2.86E-05 | 1.70 |
| 280.00 | 4.06E-04 | 4.78E-05 | 3.82E-04 | 2.16E-05 | 2.21 |
| 300.00 | 3.40E-04 | 4.39E-05 | 3.73E-04 | 2.60E-05 | 1.69 |
| 320.00 | 4.06E-04 | 4.79E-05 | 3.39E-04 | 2.28E-05 | 2.10 |
| 340.00 | 3.30E-04 | 4.33E-05 | 3.36E-04 | 2.38E-05 | 1.82 |
| 360.00 | 3.33E-04 | 4.34E-05 | 2.75E-04 | 1.90E-05 | 2.29 |
| 380.00 | 2.85E-04 | 4.04E-05 | 2.96E-04 | 2.04E-05 | 1.98 |
| 400.00 | 2.90E-04 | 4.06E-05 | 2.52E-04 | 1.70E-05 | 2.39 |
| 420.00 | 2.67E-04 | 3.90E-05 | 2.40E-04 | 1.84E-05 | 2.12 |
| 440.00 | 2.32E-04 | 3.63E-05 | 2.41E-04 | 1.92E-05 | 1.89 |
| 460.00 | 2.26E-04 | 3.58E-05 | 2.28E-04 | 1.56E-05 | 2.30 |
| 480.00 | 1.92E-04 | 3.30E-05 | 1.95E-04 | 1.51E-05 | 2.19 |
| 500.00 | 2.27E-04 | 3.59E-05 | 1.76E-04 | 1.20E-05 | 2.99 |
| 520.00 | 1.43E-04 | 2.87E-05 | 2.32E-04 | 2.36E-05 | 1.22 |
| 540.00 | 1.59E-04 | 3.00E-05 | 1.78E-04 | 1.55E-05 | 1.93 |
| 560.00 | 1.31E-04 | 2.74E-05 | 1.68E-04 | 1.48E-05 | 1.84 |
| 580.00 | 1.18E-04 | 2.58E-05 | 1.39E-04 | 1.23E-05 | 2.10 |
| 600.00 | 1.52E-04 | 2.93E-05 | 1.56E-04 | 1.32E-05 | 2.23 |
| 620.00 | 1.20E-04 | 2.63E-05 | 1.42E-04 | 1.67E-05 | 1.57 |
| 640.00 | 1.37E-04 | 2.80E-05 | 1.41E-04 | 1.22E-05 | 2.30 |
| 660.00 | 1.02E-04 | 2.41E-05 | 1.22E-04 | 9.85E-06 | 2.45 |
| 680.00 | 1.20E-04 | 2.62E-05 | 1.11E-04 | 1.17E-05 | 2.25 |
| 700.00 | 1.41E-04 | 2.83E-05 | 1.17E-04 | 1.17E-05 | 2.43 |
| 720.00 | 1.13E-04 | 2.52E-05 | 1.14E-04 | 9.82E-06 | 2.57 |
| 740.00 | 9.09E-05 | 2.27E-05 | 1.27E-04 | 1.32E-05 | 1.72 |
| 760.00 | 1.26E-04 | 2.69E-05 | 1.38E-04 | 1.23E-05 | 2.19 |
| 780.00 | 1.03E-04 | 2.43E-05 | 9.93E-05 | 9.34E-06 | 2.61 |
| 800.00 | 6.24E-05 | 1.88E-05 | 1.13E-04 | 1.18E-05 | 1.59 |

Table 6.5: Evaluating DL with Different Degrees of Freedom

| parameters | $n$ | $k$ | $T$ | $\rho$ | $\lambda^{-1}$ | $r$ | $R$ | $m$ | $m_1$ | $m_2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 3 | 2 | 0.5 | 400 | 0.05 | 0.4 | 100000 | 10 | 10000 |

| | Basic MC | | Conditional Importance Sampling | | |
|---|---|---|---|---|---|
| *df* | $DL$ | $SE$ | $DL$ | $SE$ | Ratio |
| 1 | 2.02E-03 | 1.07E-04 | 2.33E-03 | 2.26E-04 | 0.47 |
| 2 | 1.34E-03 | 8.76E-05 | 1.45E-03 | 1.26E-04 | 0.69 |
| 3 | 9.53E-04 | 7.37E-05 | 9.81E-04 | 8.84E-05 | 0.83 |
| 4 | 7.79E-04 | 6.65E-05 | 6.87E-04 | 6.38E-05 | 1.04 |
| 5 | 5.98E-04 | 5.84E-05 | 5.19E-04 | 4.18E-05 | 1.40 |
| 6 | 4.62E-04 | 5.14E-05 | 4.54E-04 | 3.67E-05 | 1.40 |
| 7 | 3.42E-04 | 4.41E-05 | 3.76E-04 | 2.88E-05 | 1.53 |
| 8 | 3.34E-04 | 4.35E-05 | 3.70E-04 | 2.90E-05 | 1.50 |
| 9 | 3.06E-04 | 4.16E-05 | 2.90E-04 | 2.22E-05 | 1.87 |
| 10 | 2.26E-04 | 3.57E-05 | 2.23E-04 | 1.44E-05 | 2.48 |
| 11 | 2.28E-04 | 3.61E-05 | 2.01E-04 | 1.14E-05 | 3.18 |
| 12 | 1.86E-04 | 3.25E-05 | 2.33E-04 | 1.69E-05 | 1.92 |
| 13 | 2.28E-04 | 3.60E-05 | 2.06E-04 | 1.33E-05 | 2.70 |
| 14 | 1.83E-04 | 3.23E-05 | 1.86E-04 | 1.25E-05 | 2.58 |
| 15 | 1.19E-04 | 2.60E-05 | 1.74E-04 | 1.35E-05 | 1.93 |
| 16 | 1.98E-04 | 3.34E-05 | 1.68E-04 | 9.67E-06 | 3.45 |
| 17 | 2.25E-04 | 3.55E-05 | 1.58E-04 | 8.57E-06 | 4.14 |
| 18 | 9.57E-05 | 2.32E-05 | 1.58E-04 | 1.03E-05 | 2.25 |
| 19 | 1.64E-04 | 3.05E-05 | 1.38E-04 | 7.61E-06 | 4.01 |
| 20 | 1.63E-04 | 3.03E-05 | 1.36E-04 | 7.79E-06 | 3.88 |

# Chapter 7

# Conclusion

In our study, we have confirmed the effectiveness of direct importance sampling under Gaussian copula and conditional importance sampling under Student-T copula for evaluating DL in BDS. We use the idea of efficient importance sampling under Large Deviation Theory in estimating joint default probability and pricing basket default swaps. We first extend CYH's conditional on marginal factors approach to the conditional on common factor approach. Then we formulate our own direct importance sampling scheme. Under Gaussian copula, we test these different algorithms with different correlation strengths and default intensity. We discover that direct change of measure is a stable approach. It is more accurate except in extreme cases when correlation is very high or low. Also, we discover that consistent with large deviation theory, as reciprocal of default intensity $\lambda^{-1}$ increases, $SE$ reduction ratio increases as well. This allows us to effectively use the direct importance sampling algorithm when default is extremely rare, causing default threshold to be very small.

Also, we prefer importance sampling because of its versatility. The importance sampling algorithm we used for estimating joint default probability is easily extendable to more complicated problems such as working with order statistics and evaluating BDS. Though Quasi MC is more effective in evaluating multivariate normal and Student-T CDF, it is not easily extendable to evaluating BDS. We compared importance sampling with Quasi MC method and discover that though importance sampling is less accurate, it nonetheless performs reasonably well. Therefore, we adopt importance sampling as the main method in this study.

Another advantage of the direct importance sampling scheme is that it only involves one expectation instead of double expectation. Monte Carlo methods become less stable and more inaccurate when working with multiple layers of expectations. Direct importance sampling removes one layer of expectation. We take advantage of this feature and apply it to Student-T copula. Under Student-T copula, our importance sampling scheme under Gaussian Copula cannot be applied directly, so we first condition on the $\chi^2$ variable. Then we are back to working with multivariate normal variable. Now we can apply our importance sampling scheme. We do not apply the condition on common or marginal factors scheme because that would require three layers of expectations. We discovered that our conditional importance sampling scheme consistently reduced $SE$ under different correlation strengths and default intensity. We also discovered that $SE$ reduction ratio increases as degrees of freedom increased. This implies that with very low degrees of freedom, it is better to use Basic MC while with higher degrees of freedom, it is better to use conditional importance sampling. However, in most cases, it is still better to use conditional importance sampling.

Here, we conclude this study by recommending direct importance sampling under Gaussian copula and conditional importance sampling under Student-T copula for pricing BDS.

# Reference


[1] Black, F. and J. Cox. Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance*, 31:351-367, 1976.

[2] Bucklew, J. *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley, 1990.

[3] Bucklew, J. *Introduction to Rare Event Simulation*, Springer-Verlag New York, Inc, 2004.

[4] Cherubini, U., E. Luciano, and W. Vecchiato. *Copula Methods in Finance*, Wiley, 2004.

[5] Chiang, M. H., M. L. Yueh, and M. H. Hsieh. An Efficient Algorithm for Basket Default Swap Valuation. *Journal of Derivatives*, Winter: 1-22, 2007.

[6] Delianedis, G. and R. Geske. Credit Risk and Risk Neutral Default Probabilities: Information about Rating Migrations and Defaults. UCLA, The Anderson School, working paper, 1998

[7] Duffie, D. and K. J. Singleton. *Credit Risk Modeling: Pricing, Measurement and Management. Princeton University Press,* Princeton, New Jersey, 2003.

[8] Genz, A. and F. Bretz. Numerical Computation of Multivariate t Probabilities with Application to Power Calculation of Multiple Contrasts, *Journal of Statistical Computation and Simulation*, 63:361-378. 1999.

[9] Geske, R. The Valuation of Corporate Liabilities as Compound Options. *Journal of Financial and Quantitative Analysis,* 5:541-552, 1977.

[10] Gregory, J. and J. P. Laurent. In the Core of Correlation. *Risk*, October: 87-91, 2004.

[11] Hull, J. and A. White. Valuing Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives,* 8(3): 12-22, 2001.

[12] Lando, D. *Credit Risk Modeling: Theory and Applications.* Princeton University Press, Princeton, New Jersey, 2004.

[13] Li, D. X. On Default Correlation: A Copula Function Approach. *Journal of Fixed Income, 9:43-54,* 2000.

[14] Li, D. X. Constructing a Credit Curve. *Risk* Special report on Credit Risk, Nov. 1988, 40-44.

[15] Longstaff, F. and E. S. Schwartz. A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. *Journal of Finance,* 50:789-819, 1995.

[16] Leland, H.E. Risky Debt, Bond Covenants and Optimal Capital Structure. *Journal of Finance,* 49:1213-1252, 1994.

[17] Merton, R. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29:449-470, 1974.

[18] Zhou, C. An Analysis of Default Correlations and Multiple Defaults. *Review of Financial Studies*, 14(2):555-576, 2001.

[19] Schönbucher, P. *Credit Derivatives Pricing Models: Model, Pricing and Implementation.* Wiley, Chichester, England, 2003.