2010

Copula and default correlation

Dongxiang Yan
Louisiana State University and Agricultural and Mechanical College, dyan2@math.lsu.edu

COPULA AND DEFAULT CORRELATION

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Mathematics

by
Dongxiang Yan
B.S., Nanjing University of Science and Technology, 2002
M.S., Shanghai Jiaotong University, 2006
August 2010

Acknowledgments

This thesis would not be possible without several contributions. It is a great pleasure to thank my advising professor, Dr. Ambar Sengupta, for his invaluable guidance, patience and kindness through my study here.

I would like to thank to my committee members, Dr. Guillermo Ferreyra and Dr. Eric T. Hillebrand for their invaluable advice and support.

It is a pleasure also to thank the Department of Mathematics, Louisiana State University, for providing me with a pleasant working environment, and Dr. Neal W. Stoltzfus, for his invaluable advice on correcting my thesis format.

This thesis is dedicated to my wife, Hanjie Zhang, my daughters Catherine Yan and Angelina Yan, and all my friends for their support and encouragement.

Table of Contents

List of Tables

List of Figures

Abstract

This work presents a study of copulas, with special focus on the Gaussian copula model and its behavior under a certain conditioning process. Simulations are carried out to examine the behavior of the moments on conditional copula model, as measured by the behavior of Wick identities which hold for multivariate Gaussians.

Chapter 1

Introduction

In this work we will give an introductory overview of the Gaussian copula model and then present simulations we have carried out to examine the stability of the model under conditioning.

A *copula* is, briefly, the joint distribution function of variables which are individually uniformly distributed on $[0, 1]$. The notion has been widely used in modeling phenomena, especially default events in finance. Copulas, arising from works of Maurice Fréchet, were first studied in the 1950s by Sklar [16]. His fundamental result, stated briefly, says that if F is an n-dimensional distribution function with marginal distribution functions $F_1(x_1)$, $F_2(x_2)$,... , $F_n(x_n)$, then there is a copula function C which satisfies

$$F(x_1, x_2, \cdots, x_n) = C\big(F_1(x_1), F_2(x_2), F_n(x_n)\big).$$

We will discuss how to prove this important theorem in Chapter 2.

In the last twenty years, with the rapid development of mathematical finance, more and more attention has been paid to create practical models to improve competitive performance in finance and insurance world. Copulas form one of the most important classes of these models in mathematical finance.There are so many works on this field, like E.W. Frees's paper[5] on understanding relationships using copulas, P. Embrechts's paper [4] on correlation and dependence in risk management. They gained widespread popularity in the credit derivatives world following the work of Li [10].

The model that gained most popularity after the work of Li is the Gaussian copula model especially the case using one underlying factor. There are several nice papers which are related to this topic, like E. Hillebrand, A. N. Sengupta, and J. Y. Xu's research[8] on temporal correlation of defaults in subprime securitization and M. Chao, A. N. Sengupta's work[12] on collateralized debt obligations tranche sensitivities in the Gaussian copula model. Despite the popularity of this model among practicioners, there are theoretical problems with the model. One problem is that it is not clear that the joint behavior of some of the variables conditional on certain given behavior of other variables preserves the Gaussian copula structure. In this work we have examined this stability question under conditioning. Our findings are not conclusive one way or another but suggest that the conditioned Gaussian copula has moment behavior close to that of a Gaussian copula again.

We begin the work in Chapter 2 with an introduction to the theory of copulas. In the next chapter we describe Gaussian copula models, and specifically factor models. We then turn to describing how default phenomena are modeled using the Gaussian copula. We conclude in Chapter 5 discussing our simulation method and observations.

One principle objective is to study the conditional moments and their deviation from Gaussian moments. For a more careful study of the behavior of the joint distribution of default time proxies, one should use standard tests for normality, such as kolmogorov normality test, but this is outside the range of our work.

Chapter 2

Copula Theory

In this chapter, we present the basic mathematical theory of copulas, including the fundamental Sklar Theorem. The proof we present is not widely known and is due to Rüschendorf [15].

We set up here some notation. As usual, $\mathbb{R}$ is the ordinary real line $(-\infty, \infty)$. Let $\mathbb{R}^*$ be the extended real line $[-\infty, \infty]$. Let $(\mathbb{R}^*)^2$ be the extended real plane $\mathbb{R}^* \times \mathbb{R}^*$. The unit square I^2 is the product $I \times I$ where $I = [0, 1]$. A 2-place real function H is a function whose domain, $D(H)$, is a subset of $(\mathbb{R}^*)^2$ and whose range, $R(H)$, is a subset of $\mathbb{R}$.

2.1 Subcopulas and Copulas

We also need some preliminary terminology (from [2]):

Definition 2.1. *Let S_1 and S_2 are two non-empty subsets of $\mathbb{R}^*$. Let s_1 be the least element of S_1, and s_2 be the least element of S_2. A real function f is called grounded if, for every $(u \times v)$ in $S_1 \times S_2$,*

$$f(s_1, u) = 0 = f(v, s_2) \tag{2.1}$$

For any rectangle $[u_1, u_2] \times [v_1, v_2]$ whose vertices lie in $S_1 \times S_2$ with $u_1 \leqslant u_2, v_1 \leqslant v_2$, a real function $f : S_1 \times S_2 \to \mathbb{R}$ is called 2-increasing if

$$f(u_2, v_2) + f(u_1, v_1) - f(u_1, v_2) - f(u_2, v_1) \geqslant 0 \tag{2.2}$$

Using this language we have the notion of a subcopula (Nelsen [14]):

Definition 2.2. *Let S_1 and S_2 be non-empty subsets of $I = [0, 1]$ containing 0 and 1. A two-dimentional subcopula (or 2-subcopula) is a real function $C' : S_1 \times S_2 \to \mathbb{R}$ with the following properties:*

1. C' is grounded and 2-increasing;

2. For every $(u \times v)$ in $S_1 \times S_2$,

$$C'(u,1) = u \text{ and } C'(1,v) = v. \tag{2.3}$$

Groundedness for C' implies that

$$C'(u,0) = 0 = C'(0,v) \tag{2.4}$$

A copula is a multivariate distribution function with uniform marginals on $[0,1]$. Formally:

Definition 2.3. *A two-dimensional subcopula C' with domain I^2 is called a two-dimensional copula (or a copula). That is to say that a copula is a real function $C' : I^2 \to I$ with the following properties:*

1. For every (u,v) in $I \times I$,

$$C(u,1) = u \text{ and } C(1,v) = v \tag{2.5}$$

$$C(u,0) = 0 = C(0,v) \tag{2.6}$$

2. For every u_1, u_2, v_1, v_2 in I with $u_1 \leqslant u_2, v_1 \leqslant v_2$,

$$C(u_2,v_2) + C(u_1,v_1) - C(u_1,v_2) - C(u_2,v_1) \geqslant 0 \tag{2.7}$$

In order to apply copulas to financial markets, the probabilistic interpretation is needed. The copula is used to model the joint behavior of variables with known marginal distribution functions. Sklar's theorem describes this relationship clearly.

2.2 Transforming to Uniform Variables

An extremely useful technique in the application of copulas is the transformation of random variables to uniformly distributed variables.

Proposition 2.4. *If the distribution function F of a random variable X is continuous and strictly increasing then the variable*

$$U = F(X)$$

is uniformly distributed on $[0, 1]$*, and*

$$X = F^{-1}(U).$$

Proof. For any $t \in (0, 1)$ we have

$$\begin{aligned} P[U < t] &= P[X < F^{-1}(t)] \\ &= F\big(F^{-1}(t)\big) \\ &= t. \end{aligned} \tag{2.8}$$

Thus, U is uniformly distributed on $[0, 1]$. □

In the argument above it was essential to have the inverse function F^{-1}, and it is for this reason that we assumed F to be strictly increasing and continuous. The following result of Rüschendorf [15] covers the case where F does not have an inverse in the usual sense.

Proposition 2.5 (Distributional transform)**.** *Let F be the distribution function of a real random variable X, and Y an independent variable with uniform distribution on* $(0, 1)$*. Let $F(x, \alpha)$ be the modified distribution function*

$$F(x, \alpha) = P(X < x) + \alpha P(X = x). \tag{2.9}$$

Then the variable

$$U \stackrel{\text{def}}{=} F(X, Y) \tag{2.10}$$

has uniform distribution on $(0, 1)$*, and*

$$X = F^{-1}(U) \quad \textit{almost surely,} \tag{2.11}$$

where the inverse F^{-1} is defined by

$$F^{-1}(s) = \inf\{x \in \mathbb{R} : F(x) \geqslant s\}, s \in (0, 1). \tag{2.12}$$

Proof. Since $F(x, \alpha) = P(X < x) + \alpha P(X = x)$, then $F(x, \alpha)$ is the same with $F(x)$ on the condition that F is continuous distribution function.

Since $U = F(X, Y)$, then we have $U = P(X < x) + Y P(X = x)$.

Let $Q_X(\beta)$ be the lower β-quantile which is defined by

$$Q_X(\beta) = \sup\{x : P(X \leqslant x) < \beta\},\ 0 < \beta < 1 \tag{2.13}$$

It is equivalent to define that $Q_X(\beta) = \sup\{x : F(X) < \beta\}$ with $0 < \beta < 1$, since $F(X) = P(X \leqslant x)$.

If $F(X, Y) \leqslant \beta$, then we have $(X, Y) \in \{(x, \alpha) : P(X < x) + \alpha P(X = x) \leq \beta\}$. In converse, if $(X, Y) \in \{(x, \alpha) : P(X < x) + \alpha P(X = x) \leq \beta\}$, then we will have $F(X, Y) \leqslant \beta$. These implies the above two equations have equivalent relationship, i.e. $\{(x, y) : F(X, Y) \leqslant \beta\} = \{(x, y) : P(X < x) + yP(X = x) \leq \beta\}$. The left part is sufficient and necessary for the right part,

Define

$$\gamma = P[X = Q_X(\beta)]. \tag{2.14}$$

Define

$$\lambda = P[X < Q_X(\beta)]. \tag{2.15}$$

Then we can see that $\{(x, \alpha) : P(X < x) + \alpha P(X = x) \leq \beta\}$ is equivalent to

$$\{X < Q_X(\beta)\} \cup \{X = Q_X(\beta), \lambda + Y\gamma \leqslant \beta\}. \tag{2.16}$$

Then we have

$$P[F(X, Y) \leqslant \beta] = P[X < Q_X(\beta)] + P[X = Q_X(\beta)]P(\lambda + Y\gamma \leqslant \beta). \tag{2.17}$$

Suppose $\gamma > 0$, then

$$P[F(X, Y) \leqslant \beta] = P[X < Q_X(\beta)] + P[X = Q_X(\beta)]P\{Y \leqslant \frac{\beta - P[X < Q_X(\beta)]}{P[X = Q_X(\beta)]}\}. \tag{2.18}$$

Since $P(U \leqslant \beta) = P[F(X,Y) \leqslant \beta]$, and

$$\frac{\beta - P[X < Q_X(\beta)]}{P[X = Q_X(\beta)]} = \frac{\beta - \lambda}{\gamma}, \tag{2.19}$$

then we have

$$P(U \leqslant \beta) = \lambda + \gamma\frac{\beta - \lambda}{\gamma} = \beta. \tag{2.20}$$

Suppose $\gamma = 0$, then we have

$$P[F(X,Y) \leqslant \beta] = P[X < Q_X(\beta)] = P[X \leqslant Q_X(\beta)] = \beta. \tag{2.21}$$

Thus we have $U = U(0,1)$.

Furthermore, we have $P(X < x) \leqslant U \leqslant P(X \leqslant x)$. This implies that for any u in $(P(X < x), P(X \leqslant x)]$, we have $x = F^{-1}(u)$. Thus, we have $X = F^{-1}(U)$ a.s. □

2.3 Sklar's Theorem

Here is the fundamental result of copula theory:

Theorem 2.6 (Sklar's Theorem)**.** *Let F be an n-dimensional joint distribution function which has marginal distribution functions F_i, for $i = 1, 2, \cdots, n$. Then there exists a copula C on $[0,1]^n$ such that*

$$F(x_1, x_2, \cdots, x_n) = C\big(F_1(x_1), F_2(x_2), \cdots, F_n(x_n)\big) \tag{2.22}$$

for all $(x_1, ..., x_n) \in \mathbb{R}^n$.

The theorem tells that the n-dimensional distribution function F can be understood as being composed of two parts: C, the copula function, and F_i, the marginal distribution functions. This implies that F can be transformed to a unique subcopula of marginals. Then a copula of uniform marginals can be obtained by extending this subcopula.

Now we can use Proposition 2.5 to prove Sklar's theorem in a short way, following Rüschendorf[15].

Proof. Let $X = (X_1, X_2, \cdots, X_n)$ be a random vector with distribution function F. Let Y be uniformly distributed on $(0, 1)$, and independent of X. By Propositon 2.5, we know that the distributional transform of $U_i = F_i(X_i, Y)$ is uniformly distributed on $(0, 1)$. Furthermore, we know that $X_i = F_i^{-1}(U_i)$ a.s. If we let the copula C be the distribution function of $U = (U_1, U_2, \cdots, U_n)$, then we have

$$\begin{aligned} F(X) &= P(X \leq x) \\ &= P(F_i^{-1}(U_i) \leq x_i, 1 \leq i \leq n) \\ &= P(U_i \leq F_i(x_i), 1 \leq i \leq n) \\ &= C(F_1(x_1), \cdots, F_n(x_n)). \end{aligned}$$

This implies that C is the copula of F. □

The standard proof of Sklar's theorem can be found in Nelsen's book [14].

2.4 Monotonicity

Copulas have a fundamental feature of invariance under monotone transformations ([14]):

Theorem 2.7. *Let C_{XY} be the copula of continuous random variables X and Y. If μ and ν are both strictly increasing functions on the range of X and on the range of Y, respectively, then*

$$C_{\mu(X)\nu(Y)} = C_{XY}. \tag{2.23}$$

Proof. Suppose Q_1, W_1, Q_2 and W_2 are corresponding distribution functions of X, Y, $\mu(X)$ and $\nu(Y)$.Since the strictly increasing property of μ and ν, then we have

$$Q_2(x) = P[\mu(X) \leqslant x] = P[X \leqslant \mu^{-1}(x)] = Q_1(\mu^{-1}(x)) \tag{2.24}$$

and

$$W_2(y) = P[\nu(Y) \leqslant y] = P[Y \leqslant \nu^{-1}(y)] = W_1(\nu^{-1}(y)) \tag{2.25}$$

Furthermore, we have

$$
\begin{aligned}
C_{\mu(X)\nu(Y)}(Q_2(x), W_2(y)) &= P[\mu(X) \leqslant x, \nu(Y) \leqslant y] \\
&= P[X \leqslant \mu^{-1}(x), Y \leqslant \nu^{-1}(y)] \\
&= C_{XY}(Q_1(\mu^{-1}(x)), W_1(\nu^{-1}(y))) \\
&= C_{XY}(Q_2(x), W_2(y)).
\end{aligned} \tag{2.26}
$$

Since X, Y are both continuous, then we have

$$R(Q_2) = R(W_2) = I. \tag{2.27}$$

This implies that on I^2

$$C_{\mu(X)\nu(Y)} = C_{XY}. \tag{2.28}$$

□

Here is another result on the effect of transformation of variables and the copula:

Theorem 2.8. *Let C_{XY} be the copula of continuous random variables X and Y. Suppose μ and ν are both strictly decreasing on $R(X)$ and $R(Y)$, then we have*

$$C_{\mu(X)\nu(Y)}(Q_2(x), W_2(y)) = Q_2(x) + W_2(y) - 1 + C_{XY}(1 - Q_2(x), 1 - W_2(y)). \tag{2.29}$$

Proof. Suppose Q_1, W_1, Q_2 and W_2 are corresponding distribution functions of X, Y, $\mu(X)$ and $\nu(Y)$.Since the strictly decreasing property of μ and ν, then we have

$$
\begin{aligned}
Q_2(x) = P[\mu(X) \leqslant x] &= P[X \geqslant \mu^{-1}(x)] \\
&= 1 - P[X < \mu^{-1}(x)] \\
&= 1 - Q_1(\mu^{-1}(x)).
\end{aligned} \tag{2.30}
$$

and

$$
\begin{aligned}
W_2(y) = P[\nu(Y) \leqslant y] &= P[Y \geqslant \nu^{-1}(y)] \\
&= 1 - P[Y < \nu^{-1}(y)] \\
&= 1 - W_1(\nu^{-1}(y)).
\end{aligned} \tag{2.31}
$$

Furthermore, we have

$$
\begin{aligned}
&C_{\mu(X)\nu(Y)}(Q_2(x), W_2(y)) \\
&\qquad = P[\mu(X) \leqslant x, \nu(Y) \leqslant y] \\
&\qquad = P[X \geqslant \mu^{-1}(x), Y \geqslant \nu^{-1}(y)] \\
&\qquad = 1 - Q_1(\mu^{-1}(x)) + 1 - W_1(\nu^{-1}(y)) - \\
&\qquad\qquad (1 - P[X \leqslant 1 - \mu^{-1}(x), Y \leqslant 1 - \nu^{-1}(y)] \\
&\qquad = Q_2(x) + W_2(y) - 1 + C_{XY}(1 - Q_2(x), 1 - W_2(y)).
\end{aligned}
\tag{2.32}
$$

Since X, Y are both continuous, then on I^2 we have

$$C_{\mu(X)\nu(Y)}(Q_2(x), W_2(y)) = Q_2(x) + W_2(y) - 1 + C_{XY}(1 - Q_2(x), 1 - W_2(y)). \tag{2.33}$$

□

We close with another similar property:

Theorem 2.9. *Let C_{XY} be the copula of continuous random variables X and Y. Suppose μ is strictly increasing on $R(X)$ and ν is strictly decreasing on $R(Y)$, then we have*

$$C_{\mu(X)\nu(Y)}(Q_2(x), W_2(y)) = Q_2(x) - C_{XY}(Q_2(x), 1 - W_2(y)). \tag{2.34}$$

Proof. Suppose Q_1, W_1, Q_2 and W_2 are corresponding distribution functions of X, Y, $\mu(X)$ and $\nu(Y)$.Since the strictly increasing property of μ, then we have

$$Q_2(x) = P[\mu(X) \leqslant x] = P[X \leqslant \mu^{-1}(x)] = Q_1(\mu^{-1}(x)) \tag{2.35}$$

Since the strictly decreasing property of ν, then we have

$$
\begin{aligned}
W_2(y) = P[\nu(Y) \leqslant y] &= P[Y \geqslant \nu^{-1}(y)] \\
&= 1 - P[Y \leqslant \nu^{-1}(y)] \\
&= 1 - W_1(\nu^{-1}(y)).
\end{aligned}
\tag{2.36}
$$

Furthermore, we have

$$
\begin{aligned}
C_{\mu(X)\nu(Y)}&(Q_2(x), W_2(y)) \\
&= P[\mu(X) \leqslant x, \nu(Y) \leqslant y] \\
&= P[X \leqslant \mu^{-1}(x), Y \geqslant \nu^{-1}(y)] \\
&= P[X \leqslant \mu^{-1}(x)] - P[X \leqslant \mu^{-1}(x), Y < \nu^{-1}(y)] \\
&= Q_1(\mu^{-1}(x)) - C_{XY}(Q_1(\mu^{-1}(x)), W_1(\nu^{-1}(y))) \\
&= Q_2(x) - C_{XY}(Q_2(x), 1 - W_2(y)).
\end{aligned}
\tag{2.37}
$$

Since X, Y are both continuous, then on I^2 we have

$$Q_2(x) = P[\mu(X) \leqslant x] = P[X \leqslant \mu^{-1}(x)] = Q_1(\mu^{-1}(x)) \tag{2.38}$$

Chapter 3

Gaussian Copula and Correlation

In the financial world, there are several famous families of copulas, such as the Gaussian copula, Student's t copula, and the Archimedean copula.

Among these, the Gaussian copula is most widely known and used. Briefly, a Gaussian copula is simply a multivariate Gaussian distribution whose marginals are standard Gaussian. In order to be consistent with the terminology of copulas used before, one should require a transformation so as to make the marginals uniform on $(0, 1)$. It is only in the specifics of application that the significance becomes clear.

In this chapter we will look at the Gaussian copula, focusing on the two-dimensional case to be specific [2].

3.1 The Two-dimensional Gaussian Copula

Definition 3.1. *A two-dimensional Gaussian copula is a copula function of the form*

$$C(\omega, \nu) = \Phi_{\rho_{XY}}(\Phi^{-1}(\omega), \Phi^{-1}(\nu)) \tag{3.1}$$

where $\Phi_{\rho_{XY}}$ denotes the joint distribution of a 2-dimensional standard normal marginal distributions, with correlation coefficient ρ_{XY},and Φ is the standard normal distribution function.

Thus, by substituting Gaussian probability density function into above equation, we have

$$\Phi_{\rho_{XY}}(\Phi^{-1}(\omega), \Phi^{-1}(\nu)) = \int_{-\infty}^{\Phi^{-1}(\omega)} \int_{-\infty}^{\Phi^{-1}(\nu)} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} \exp\left(\frac{2r_{XY}st - s^2 - t^2}{2(1-\rho_{XY}^2)}\right) dsdt \tag{3.2}$$

We can also write above equation in an equivalent form:

$$C(\omega,\nu) = \int_0^{\omega} \Phi\left(\frac{\Phi_{-1}(\nu) - \rho_{XY}\Phi_{-1}(t)}{\sqrt{1-\rho_{XY}^2}}\right) dt \tag{3.3}$$

Definition 3.2. *The density $c(\omega,\nu)$ associated to a copula $C(\omega,\nu)$ is*

$$c(\omega,\nu) = \frac{\partial^2 C(\omega,\nu)}{\partial\omega\partial\nu} \tag{3.4}$$

So, the density of the Gaussian copula is

$$c(\omega,\nu) = \frac{1}{\sqrt{1-\rho_{XY}^2}} \exp\left(\frac{x^2+y^2}{2} + \frac{2r_{XY}xy - x^2 - y^2}{2(1-\rho_{XY}^2)}\right) \tag{3.5}$$

where $x = \Phi^{-1}(\omega)$, $y = \Phi^{-1}(\nu)$.

By integrating the density, since the copula is absolutely continuous, the following equivalent expression for the copula can be obtained:

$$c(\omega,\nu) = \int_0^{\omega}\int_0^{\nu} \exp\left(\frac{x^2+y^2}{2} + \frac{2r_{XY}xy - x^2 - y^2}{2(1-\rho_{XY}^2)}\right) dsdt \tag{3.6}$$

where $x = \Phi^{-1}(s)$, $y = \Phi^{-1}(t)$.

According to Sklar's theorem, the Gaussian copula generates the joint normal standard distribution function if and only if the margins are standard normal. That is to say

$$C(F_1(x), F_2(y)) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} \exp\left(\frac{2r_{XY}st - s^2 - t^2}{2(1-\rho_{XY}^2)}\right) dsdt \tag{3.7}$$

if and only if $\Phi^{-1}(F_1(x)) = x$ and $\Phi^{-1}(F_2(y)) = y$, i.e. if and only if $F_1 = F_2 = \Phi$. For any other marginal choice, the Gaussian copula does not give a standard jointly normal vector.

3.2 Concordance and Discordance

Dependence and association relationship between random variables are widely researched in probability and statistics field. Copulas are usually used to study this issue. There are so many ways to find and measure dependence. In order to make this studies be meaningful, some assumptions are needed to build some statistical models.

The first concept we need to know about dependence and association is called concordance.Institutively, if we say a pair of random variables are concordant, this means that large values of each random variables prefer to be tied together, while small values of each random variables prefer to tied together. In other words, large is with large, and small is with small.

Definition 3.3. *Let (x_i, y_i) and (x_j, y_j) be two observations of (X, Y). Then (x_i, y_i) and (x_j, y_j) are concordant if the following relations hold*

$$x_i < x_j \Longleftrightarrow y_i < y_j \tag{3.8}$$

or

$$x_i > x_j \Longleftrightarrow y_i > y_j \tag{3.9}$$

On the other hand, (x_i, y_i) and (x_j, y_j) are disconcordant if the following relations hold

$$x_i < x_j \Longleftrightarrow y_i > y_j \tag{3.10}$$

or

$$x_i > x_j \Longleftrightarrow y_i < y_j \tag{3.11}$$

It's clearly that, if (x_i, y_i) and (x_j, y_j) are concordant, then $(x_i - x_j)(y_i - y_j) > 0$; if (x_i, y_i) and (x_j, y_j) are discordant, then $(x_i - x_j)(y_i - y_j) < 0$.

One of the most important measure of association is known as Kendall's tau which is first mentioned by WH Kruskal in 1958 [9].

Definition 3.4. *Suppose X and Y are continuous random variables. Let (x_i, y_i) $(i = 1, \ldots, n)$ be n random observations of (X, Y). Totally, we know that there are $\frac{n(n-1)}{2}$ distinct pairs. Let a be the number of pairs which are concordant. Let b be the number of pairs which are discordant. Kendall's tau for this sample is defined to be*

$$t = \frac{a - b}{a + b} = \frac{2(a - b)}{n(n - 1)} \tag{3.12}$$

Now let X, Y be random variables. Create independent copies (X_1, Y_1) and (X_2, Y_2) of (X, Y) ('copies' means that (X_j, Y_j) has the same distribution as (X, Y)). Then Kendall's tau for the pair X, Y is defined by defined as

$$\tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{3.13}$$

We can see that the Kendall's tau represents the difference between probability of concordant pairs and probability of discordant pairs.

The best known correlation measure is Pearson's correlation coefficient which is defined as the quotient between covariance and the product of two standard variances [1].

Definition 3.5. *The Pearson correlation coefficient between random variables X and Y is*

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}. \tag{3.14}$$

where μ and σ correspond to mean and standard deviation, respectively.

For Gaussian variables, the relation between Kendall's tau and Pearson's correlation coefficient is known to be

$$\tau = \frac{2arcsin(\rho)}{\pi}. \tag{3.15}$$

3.3 Gaussian Factor Models

The Gaussian copula when used in practice involves not the abstract formulation of general copula theory but concrete Gaussian *factors*. Specifically, consider the problem of using a Gaussian copula to model variables $Y_1, ..., Y_N$ whose marginal distributions

$$F_i(t) = P[Y_i \leq t]$$

are known, and are assumed to be continuous and strictly increasing. Form the standard Gaussians

$$Z_i = \Phi^{-1}\big(F_i(Y_i)\big).$$

The Gaussian copula model states that the variables $Z_1, ..., Z_N$ are *jointly Gaussian*. Hence the joint distribution of these variables is completely specified by the *correlations*

$$\rho_{jk} = E[Z_j Z_k]. \tag{3.16}$$

These correlations form an $N \times N$ symmetric matrix

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1N} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{N1} & \rho_{N2} & \rho_{N3} & \cdots & 1 \end{bmatrix}$$

Instead of specifying this matrix explicitly, often a *correlation structure* is imposed by expressing $Z_1, ..., Z_N$ in terms of some underlying independent standard Gaussian variables called *factors*.

For example, for Z_1, Z_2, Z_3, there may be one common factor Z and three variables $\epsilon_1, \epsilon_2, \epsilon_3$, and then

$$Z_j = \rho_j Z + a_j \epsilon_j \tag{3.17}$$

where ρ, a_j are constants, and $Z, \epsilon_1, \epsilon_2, \epsilon_3$ are independent standard Gaussians. The correlations between the Z_j is then given by

$$\rho_{jk} = E[Z_j Z_k] = \rho_j \rho_k + a_j a_k$$

The requirement that Z_j is standard Gaussian imposes the restriction that the variance be 1:

$$\rho_j^2 + a_j^2 = 1.$$

We will use this type of model again later.

3.4 Other Copulas: A Quick Look

We take a quick look at two other copulas.

Student's t copula is essentially a multivariate distribution function having Student's t distribution as marginals (suitably transformed to uniform). The bivariate Student's t copula can be defined as followed [2]. Suppose t_u, a function, satisfies one-dimensional Student's t distribution with degree of freedom u. This means the marginal distribution is Student's t distribution. Then we have

$$t_u(x) = \int_{-inf}^{x} \frac{\Gamma((u+1)/2))}{\sqrt{\Pi u}\Gamma(u/2)}(1+\frac{h^2}{u})^{-\frac{u+1}{2}}dh \tag{3.18}$$

where Γ is called Euler function which is defined as $\Gamma(x) = \prod_{k=1}^{\infty}(1-x^k)$.

Let $v \in I = [0,1]$, then we have two-dimentional Student's t distribution function as following

$$t_(v,u)(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\Pi\sqrt{1-v^2}}(1+\frac{h^2+l^2-2vhl}{u(1-v^2)})^{-\frac{u+2}{2}}dhdl. \tag{3.19}$$

Then the corresponding Student's t copula is

$$\begin{aligned} C_(v,u)(u,z) &= t_{v,u}(t_u^{-1}(u), t_u^{-1}(z)) \\ &= \int_{-\infty}^{t_u^{-1}(u)}\int_{-\infty}^{t_u^{-1}(z)} \frac{1}{2\Pi\sqrt{1-v^2}}(1+\frac{h^2+l^2-2vhl}{u(1-v^2)})^{-\frac{u+2}{2}}dhdl. \end{aligned} \tag{3.20}$$

Archimedean copula is copula with form

$$C(x,y) = \varphi^{-1}(\varphi(x)+\varphi(y)) \tag{3.21}$$

where φ is a generator function with following properties:

$$\varphi(1) = 0 \tag{3.22}$$

$$\varphi'(x) < 0 \tag{3.23}$$

$$\varphi''(x) > 0 \tag{3.24}$$

More detailed discussions can be found in C.Genest's paper [7].

Chapter 4

Default Correlation and Conditioning

Finding and measuring dependencies of related instruments is a necessary skill for risk managers. With the rapid development of derivatives market, more and more financial instruments are created. At the same time, the derivatives market is slowly transformed from over the counter market to standardized, liquid and transparent market. This change provides enough available financial instruments to meet different investor's risk-return expectations in a large range. In order to control risk of derivative, researchers devote a lot energy to find and modify related models. Modeling and measuring dependencies of many financial products are focused on correlation.

Default correlation understanding is essential for managers to manage the risk of a credit portfolio. In order to control the risk, we need to valuate the default correlations between credit derivatives. In 1995, D.Lucus defined a default correlation in discrete model [11].

For example, let D_1 and D_2 be default events of two credit derivatives over 1 year. The probability for each security to default is defined as $p_1 = P[D_1]$ and $p_2 = P[D_2]$ respectively. The probability for both securities to default within one year is defined as $p_{1,2} = P[D_1 \cap D_2]$. Then the default correlation between these two securities may be defined as

$$\rho = \frac{p_{1,2} - p_1 p_2}{\sqrt{p_1 p_2 (1 - p_1)(1 - p_2)}}. \tag{4.1}$$

This definition seems a good valuation of default correlation. When we check it carefully, we may find some limitations. From above, we may see that the probability of one security to default is dependent on the time interval. Then the default correlation is also tied with time period. For fixed one year period, some important information about securities may be missed.

In 2000, David X. Li introduced a new copula model in the credit derivatives market [10]. It is called the default time (or survival time) Gaussian copula. This model is widely used. The Gaussian copula is the joint distribution of variables with marginal normal distributions. The key point in Li's default time Gaussian copula is that it is a model for the joint distribution of default times, rather than the correlations of default events. The concept of time-until-default is created to characterize the default. The correlation coefficient between survival times of two securities were used to replace their default correlation.

4.1 Gaussian Copula for Default Times

Consider a portfolio of N defaultable securities. The distribution of the default time τ_j of the j-th security is

$$F_j(t) = \mathbb{P}[\tau_j < t] \tag{4.2}$$

Assume that this is strictly increasing and continuous, and set

$$X_j = \Phi^{-1}\big(F_j(\tau_j)\big). \tag{4.3}$$

As seen before, this is standard Gaussian.

The event that the j-th security defaults by time T is

$$[\tau_j \leq T]$$

and, in terms of X_j, this is

$$[X_j \leq c_j(T)],$$

where

$$c_j(T) = \Phi^{-1}\big(F_j(T)\big). \tag{4.4}$$

When we do simulations we check if X_j is below the *default threshold* and if it is then we say that the j-th security has defaulted.

The Gaussian copula model for default times posits that $X_1, ..., X_N$ have *joint Gaussian distribution.*

The most basic and widely used Gaussian copula model is the *one factor* model. In this we assume that there are independent standard Gaussians

$$Z, \varepsilon_1, ..., \varepsilon_N,$$

and assume that

$$X_j = \sqrt{\rho}\, Z + \sqrt{1-\rho}\, \varepsilon_j, \tag{4.5}$$

for all $j \in \{1, ..., N\}$, where $\rho > 0$ is a fixed parameter. Thus Z is a *common factor* and ε_j is an *idiosyncratic* variable for the j-th security.

It is possible to view this as a simplified proxy to the default model of Merton [13].

It is easy to see that X_i, as specified in the model, does have mean 0 and variance 1:

$$\begin{aligned} E[X_i] &= E[\sqrt{\rho}Z + \sqrt{1-\rho}\varepsilon_i] \\ &= E[\sqrt{\rho}Z] + E[\sqrt{1-\rho}\varepsilon_i] \\ &= 0 + 0 \\ &= 0 \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} Var(X_i) &= Var(\sqrt{\rho}Z + \sqrt{1-\rho}\varepsilon_i) \\ &= Var(\sqrt{\rho}Z) + Var(\sqrt{1-\rho}\varepsilon_i) \\ &= \rho Var(Z) + (1-\rho)Var(\varepsilon_i) \\ &= \rho + 1 - \rho \\ &= 1 \end{aligned} \tag{4.7}$$

The key correlation measure between default behavior of the securities is

$$E[X_j X_k] = \sqrt{\rho}\sqrt{\rho} = \rho$$

for all $j \neq k$.

4.2 Conditional Defaults

Our objective is to examine the stability of the Gaussian copula model under defaults of some of the securities.

We will assume that each security has the same default time distribution F, that is, all F_j are equal to a common F. The consequence of this is that for each time t_0, the corresponding default threshold

$$c = \Phi^{-1}\big(F(t_0)\big).$$

is the same for all the securities.

For example, what is the distribution of $\tau_3, ..., \tau_N$ if the first two securities have defaulted by time t_0? This is

$$P[\tau_3 \leq t_3, ..., \tau_N \leq t_N \,|\, \tau_1 \leq t_0, \tau_2 \leq t_0],$$

for $t > t_0 > 0$.

In particular, we have the conditional marginals

$$F_{j,c}(t) = P[\tau_j \leq t \,|\, \tau_1 \leq t_0, \tau_1 \leq t_0].$$

We use this to convert τ_j to a standard Gaussian:

$$X_{j,c} = \Phi^{-1}\big(F_{j,c}(\tau_j)\big). \tag{4.8}$$

for $j \in \{3, ..., N\}$.

The big question is if the variables $X_{3,c}, ..., X_{N,c}$ are jointly Gaussian. Theoretically this seems unlikely, and so the failure of the Gaussian nature is a negative feature of the Gaussian copula model.

4.3 Wick's Theorem

In this section we discuss an important property of jointly Gaussian distributed variables. It is what we shall use in simulations to see how close the conditional variables $X_{j,c}$ are to being jointly Gaussian.

Wick's theorem expresses general moments of jointly Gaussian variables in terms of the second-order moments (covariances) [18].

Theorem 4.1. *Suppose* $(W_1, ..., W_m)$ *is Gaussian with mean* $0 \in \mathbb{R}^m$*. Then*

$$E[W_1...W_m] = \sum_{P \in \mathcal{P}_m} \prod_{\{j,k\} \in P} [W_j W_k] \tag{4.9}$$

where $\mathcal{P}_m$ *is the set of all partitionings of* $\{1, 2, ..., m\}$ *into pairs of distinct elements.*

If f m is odd, $\mathcal{P}$ is empty and so the sum in (4.9) gives 0.

For another special case take $m = 2n$ and all the W_i equal to a common variable which is standard Gaussian. Then

$$E(W^{2n}) = |\mathcal{P}_{2n}| = (2n-1)(2n-3)...3.1. \tag{4.10}$$

For $\{1, 2, 3, 4\}$, there are three pairs of partitions $\{\{1, 2\}, \{3, 4\}\}$, $\{\{1, 4\}, \{3, 2\}\}$ and $\{\{1, 3\}, \{2, 4\}\}$. This implies that

$$E[W_1W_2W_3W_4] = E[W_1W_2]E[W_3W_4] + E[W_1W_4]E[W_3W_2] + E[W_1W_3]E[W_2W_4]. \tag{4.11}$$

For $\{1, 2, 3, 4, 5, 6\}$, there will have fifteen pairs of partitions

$$
\begin{aligned}
&\{(1,2),(3,4),(5,6)\},\\
&\{(1,2),(3,5),(4,6)\},\{(1,2),(3,6),(4,5)\},\\
&\{(1,3),(2,4),(5,6)\},\{(1,3),(2,5),(4,6)\},\\
&\{(1,3),(4,5),(2,6)\},\{(1,4),(2,3),(5,6)\},\\
&\{(1,4),(3,6),(2,5)\},\{(1,4),(2,6),(3,5)\},\\
&\{(1,5),(2,3),(4,6)\},\{(1,5),(2,4),(3,6)\},\\
&\{(1,5),(2,6),(3,4)\},\{(1,6),(2,3),(4,5)\},\\
&\{(1,6),(2,4),(3,5)\},\{(1,6),(2,5),(3,4)\}.
\end{aligned}
$$

The corresponding Wick formula is:

$$
\begin{aligned}
E[W_1W_2W_3W_4W_5W_6] = {} & E[W_1W_2]E[W_3W_4]E[W_5W_6] + E[W_1W_2]E[W_3W_5]E[W_4W_6]\\
&+ E[W_1W_2]E[W_3W_6]E[W_5W_4] + E[W_1W_3]E[W_2W_4]E[W_5W_6]\\
&+ E[W_1W_3]E[W_2W_5]E[W_4W_6] + E[W_1W_3]E[W_5W_4]E[W_2W_6]\\
&+ E[W_1W_4]E[W_3W_2]E[W_5W_6] + E[W_1W_4]E[W_3W_6]E[W_5W_2]\\
&+ E[W_1W_4]E[W_3W_5]E[W_2W_6] + E[W_1W_5]E[W_3W_2]E[W_4W_6]\\
&+ E[W_1W_5]E[W_2W_4]E[W_3W_6] + E[W_1W_5]E[W_3W_4]E[W_2W_6]\\
&+ E[W_1W_6]E[W_3W_2]E[W_5W_4] + E[W_1W_6]E[W_2W_4]E[W_5W_3]\\
&+ E[W_1W_6]E[W_3W_4]E[W_5W_2].
\end{aligned}
\tag{4.12}
$$

Chapter 5

Simulation

In this chapter we describe our simulation method for studying the one-factor Gaussian copula and its behavior under conditioning. In our simulations we will never involve the default times τ_i, and instead will work exclusively with the proxy variables X_i. The number of securities will be denoted n. The number of simulations will be denoted by N.

5.1 Simulation of Correlated Gaussians

MatLab or other software can be used to simulate standard Gaussian variables. In this way we generate (simulate) iid standard Gaussians

$$Z, \epsilon_1, ..., \epsilon_n.$$

Then we form the variables

$$X_j = \sqrt{\rho}\, Z + \sqrt{1-\rho}\, \epsilon_j$$

for $j \in \{1, ..., n\}$, with a fixed choice of the correlation parameter $\rho > 0$.

This produces Gaussian variables $X_1, ..., X_n$ which have the correlation structure

$$E[X_j X_k] = \rho, \quad \text{for } j \neq k.$$

We also fix a threshold c. If a simulated value of X_j is $\leq c$ then we interpret that to mean that in that simulation the j-th security defaults within a given time horizon t_0.

5.2 Conditioning

For the sake of describing the idea we consider the task of conditioning to default of the last two securities within a given time horizon t_0. Thus, we want to understand the joint behavior of $X_1, ..., X_{n-2}$ conditional on *only* $X_{n-1}, ..., X_n$ being $\leq c$.

We generate the Z and ϵ_j as above, and compute the X_j. Now drop from the simulations those for which *only* $X_{n-1}, ..., X_n$ are below c. Retaining the others we determine the empirical distribution functions

$$H_{j,c}(x) = P[X_j \le t \,|\, X_1 > c, ..., X_{n-2} > c, X_{n-1} \le c, X_n \le c]$$

for suitable values of t, in fact just those values of X_j which have been generated for $j \in \{1, ..., n-2\}$. Thus, for each simulated value of $X_j = x$ we compute the proportion y of values, among the retained simulations, which are $\le x$. We replace the entry x by the value y. This process generates/simulates values of the variable

$$W_j = \Phi^{-1}\big(H_{j,c}(X_j)\big).$$

Note that this is standard Gaussian. Our task is to study the departure from Gaussian of the joint distribution of the variables $W_1, ..., W_{n-2}$. We do this be comparing the predicted value of, say, $E[W_1, ..., W_{n-2}]$ with that predicted by Wick's formula.

5.3 Simulation Code

The main program codes can be found in the Appendices.

5.4 Simulation Discussion

The general algorithm steps are as follows. We denote by N the number of simulations.

1. By formula $X_i = \sqrt{\rho}Z + \sqrt{1-\rho}\varepsilon_i$, let $i = 1, \cdots, n$, n names X_i are generated, where Z is generated as one standard normal random variable, and ε_i are generated as independent n size standard normal random variables. We may repeat to generate X_i as many groups as we want.

2. We set up the condition to pick up the data groups which satisfy the limitation that only the first 3 names don't default, which means only X_1, X_2, X_3 are greater than the threshold. Then we may get N groups of data. We just keep the first three

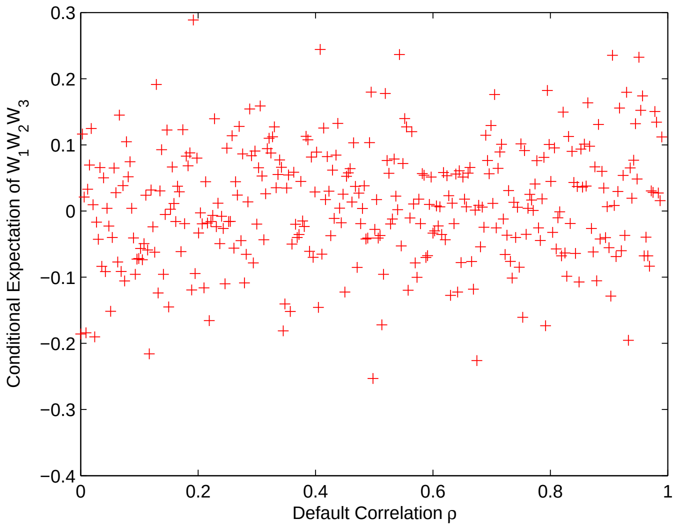

FIGURE 5.1. Conditional Expectations of Product $W_1W_2W_3$ with $n = 4$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

values for each group data and form a matrix Y with size $(N, 3)$. The first column data are all about X_1. The second column data are all about X_1, and so is the third column.

3. Then for each column, we calculate cumulative probability values for each data. By using inverse standard normal distribution function, we transform these cumulative probability data to corresponding standard normal values. Then we get a data matrix ZZ with size $(N, 3)$.

4. We notice that, there are several inf (infinity) terms in ZZ. In order to calculate expectation, we delete the whole rows where inf terms are. Then we look each column of ZZ as data with individual distribution. Finally, we calculate the expectation for these 3 joint distributions. W is used to represent ZZ in simulation figures and following discussion.

First, let's see the simulation results of conditional test for three names.

For Figure5.1, conditioning on only $X_4 < -1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.01085.

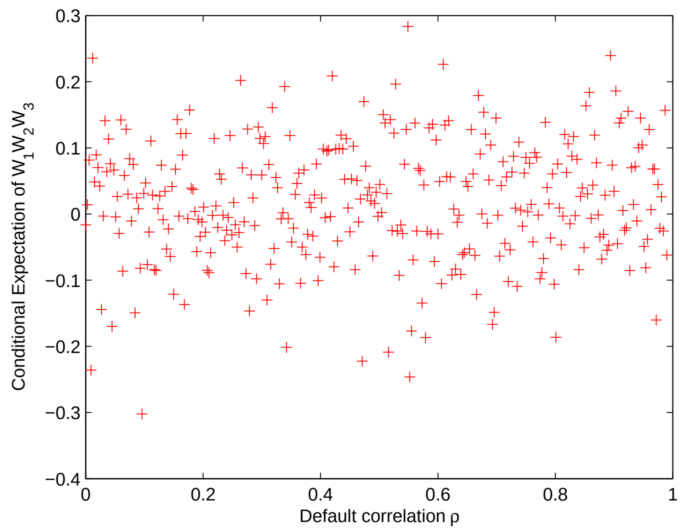

FIGURE 5.2. Conditional Expectations of Product $W_1W_2W_3$ with $n = 4$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

For Figure5.2, conditioning on only $X_4 < 0$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.01848.

For Figure5.3, conditioning on only $X_4 < 1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.009889.

For Figure5.4, conditioning on only $X_4 < -1$ and $X_5 < -1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.00591.

For Figure5.5, conditioning on only $X_4 < 0$ and $X_5 < 0$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.01123.

For Figure5.6, conditioning on only $X_4 < 1$ and $X_5 < 1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The

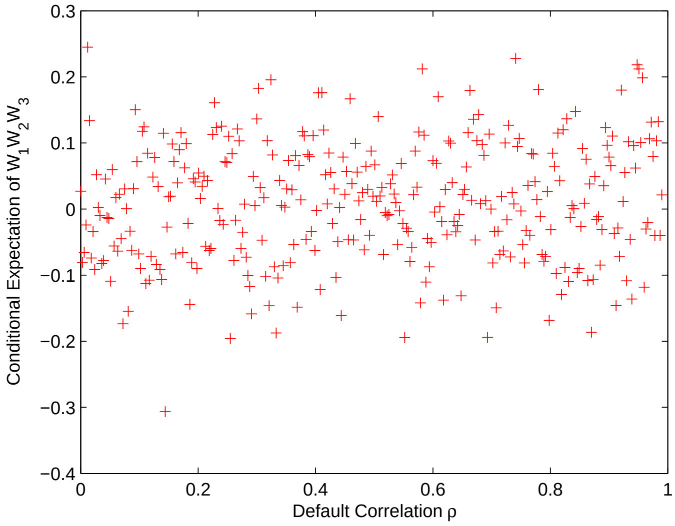

FIGURE 5.3. Conditional Expectations of Product $W_1W_2W_3$ with $n = 4$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

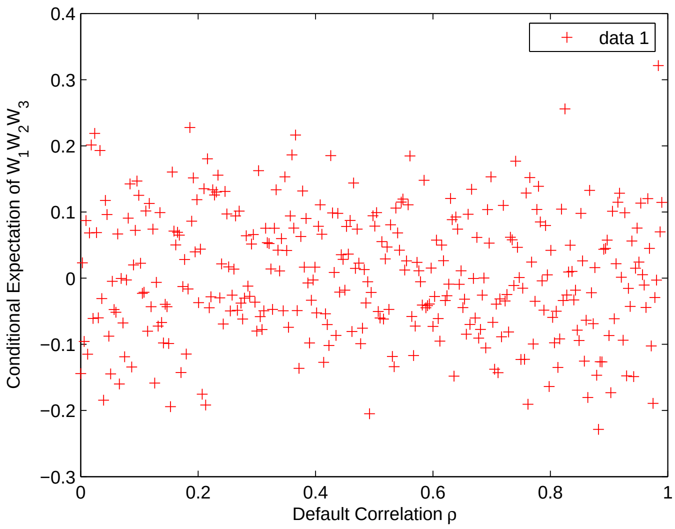

FIGURE 5.4. Conditional Expectations of Product $W_1W_2W_3$ with $n = 5$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

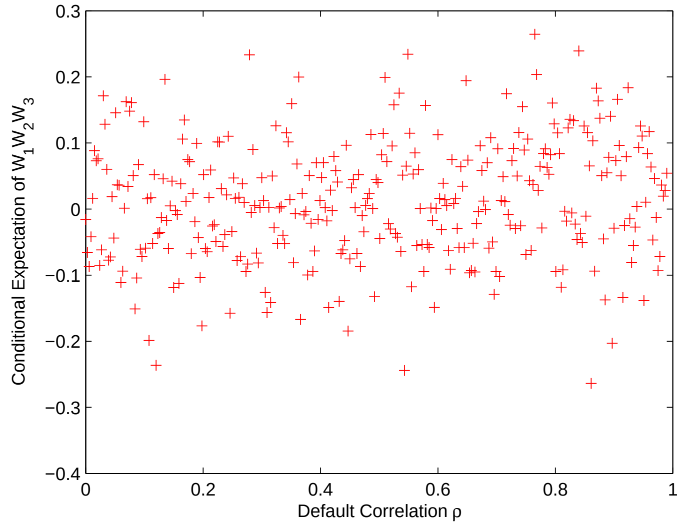

FIGURE 5.5. Conditional Expectations of Product $W_1W_2W_3$ with $n = 5$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

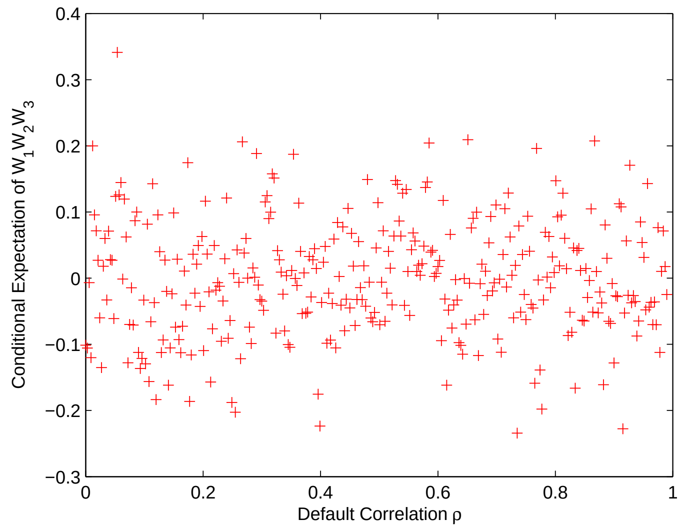

FIGURE 5.6. Conditional Expectations of Product $W_1W_2W_3$ with $n = 5$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

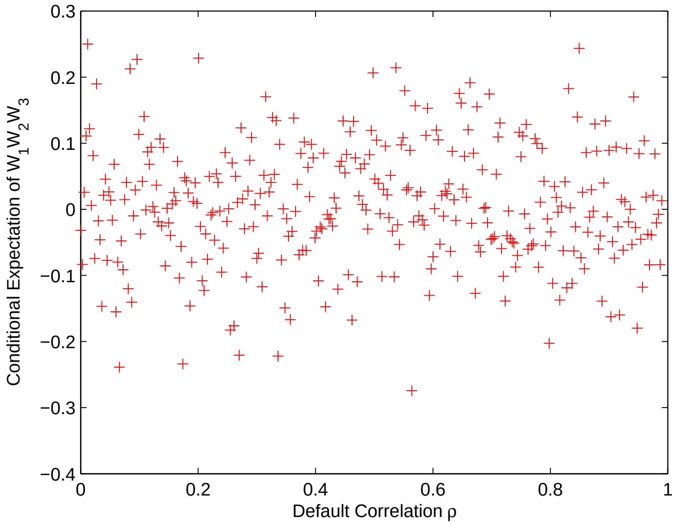

FIGURE 5.7. Conditional Expectations of Product $W_1W_2W_3$ with $n = 6$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

data size for each X_i is 100. The total number of points is 331. The mean of expectation is -0.0003618.

For Figure5.7, conditioning on only $X_4 < -1$, $X_5 < -1$ and $X_6 < -1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.005266.

For Figure5.8, conditioning on only $X_4 < 0$, $X_5 < 0$ and $X_6 < 0$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is -0.009402.

For Figure5.9, conditioning on only $X_4 < 1$, $X_5 < 1$ and $X_6 < 1$, we calculate expectations of $W_1W_2W_3$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is -0.0003505.

In summary of above figures of conditional expectation of $W_1W_2W_3$, we have Table5.1 which displays the relationship among mean of conditional expectations of $W_1W_2W_3$,

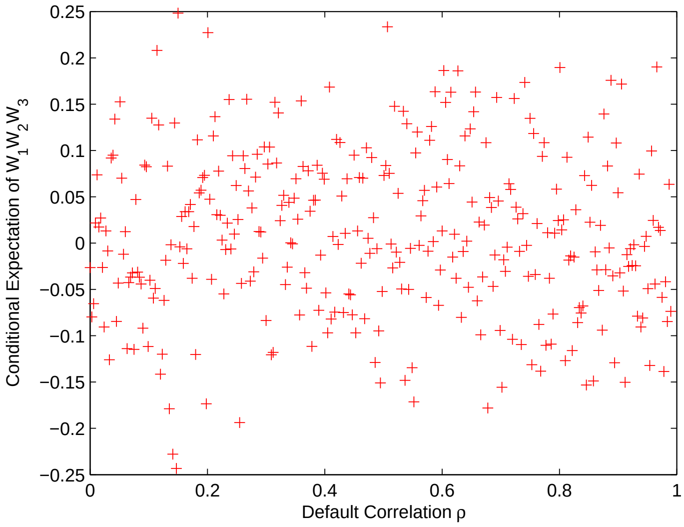

FIGURE 5.8. Conditional Expectations of Product $W_1W_2W_3$ with $n = 6$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

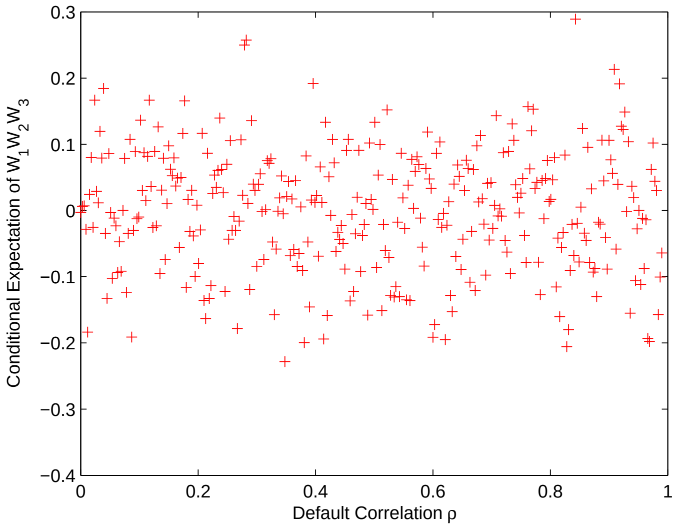

FIGURE 5.9. Conditional Expectations of Product $W_1W_2W_3$ with $n = 6$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

TABLE 5.1. Mean of Conditional Expectations of $W_1W_2W_3$. For more explanation see text

Total Names	Mean of Conditional Expectations of $W_1W_2W_3$ with Different c		
	c = -1	c = 0	c = 1
4	0.01085	0.01848	0.009889
5	0.00591	0.01123	-0.0003618
6	0.005266	-0.009402	-0.0003505

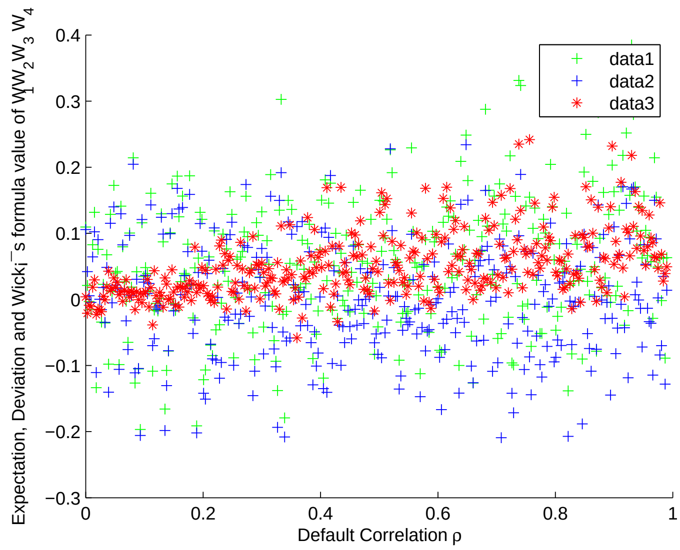

FIGURE 5.10. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 5$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

total names of portfolio, and threshold c. According to this table, we can see that, fixing threshold as -1, 0, and 1, the mean of conditional expectation of $W_1W_2W_3$ is closing to 0 as increasing of total names. Fixing total names, the means for threshold $c = 1$ are more less than the other two.

Basing on these data, the means are close to 0 which implies that the conditional variables $W_1W_2W_3$ are 'approximately' Gaussian as far as moments go, but it does not appear that they are exactly Gaussian.

Second, let's see the simulation results of conditional test for four names.

For Figure5.10, conditioning on only $X_5 < -1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.05213. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.05183. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.0002998.

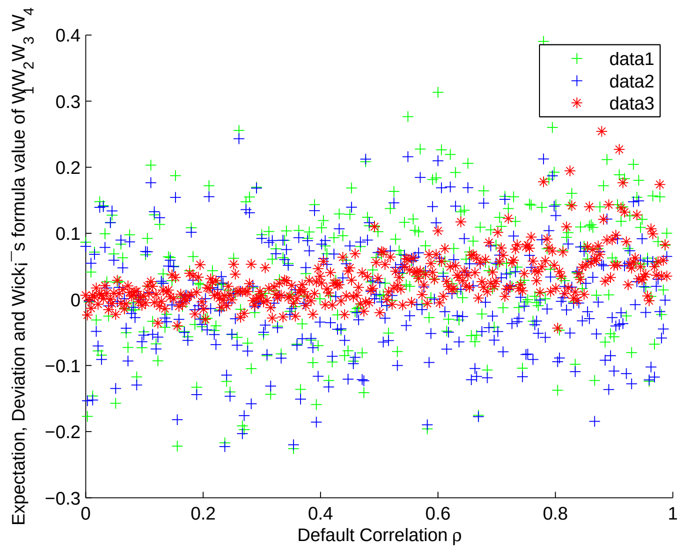

FIGURE 5.11. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 5$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

For Figure5.11, conditioning on only $X_5 < 0$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.03611. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.03168. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.004434.

For Figure5.12, conditioning on only $X_5 < 1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.03022. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.008838. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.02138.

For Figure5.13, conditioning on only $X_5 < -1$ and $X_6 < -1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total

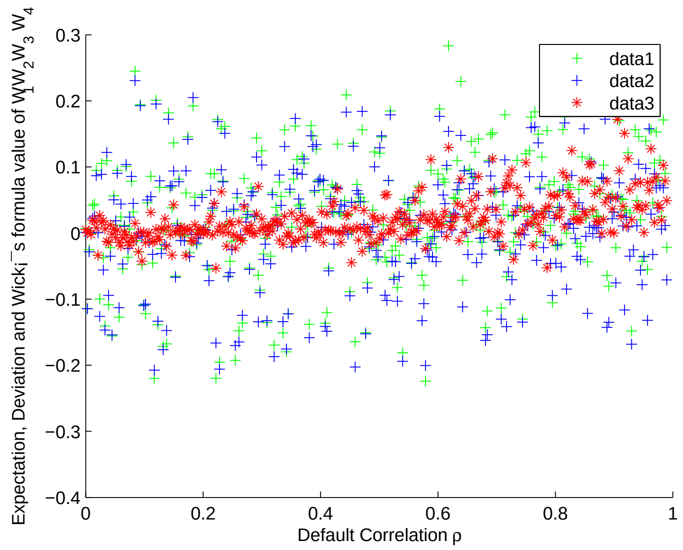

FIGURE 5.12. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 5$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

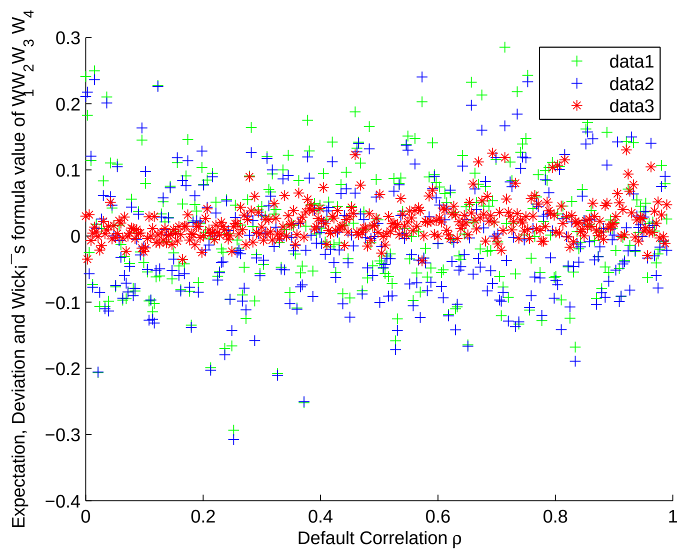

FIGURE 5.13. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 6$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

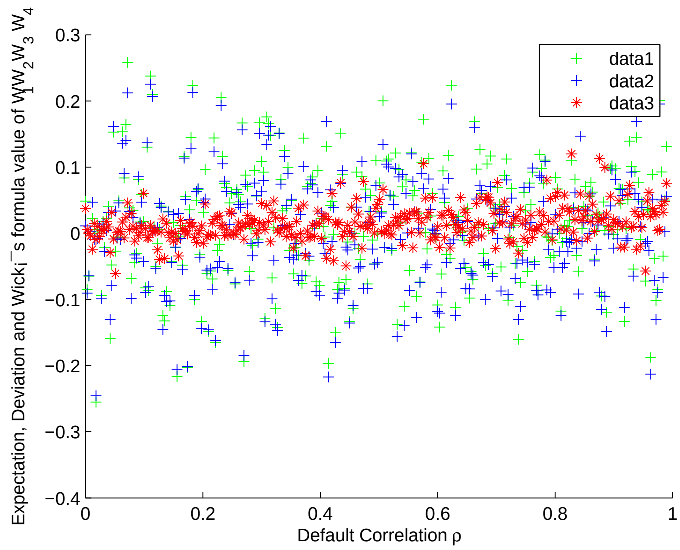

FIGURE 5.14. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 6$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.01433. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.01948. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.005145.

For Figure5.14, conditioning on only $X_5 < 0$ and $X_6 < 0$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.01662. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.01491. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.001704.

For Figure5.15, conditioning on only $X_5 < 1$ and $X_6 < 1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

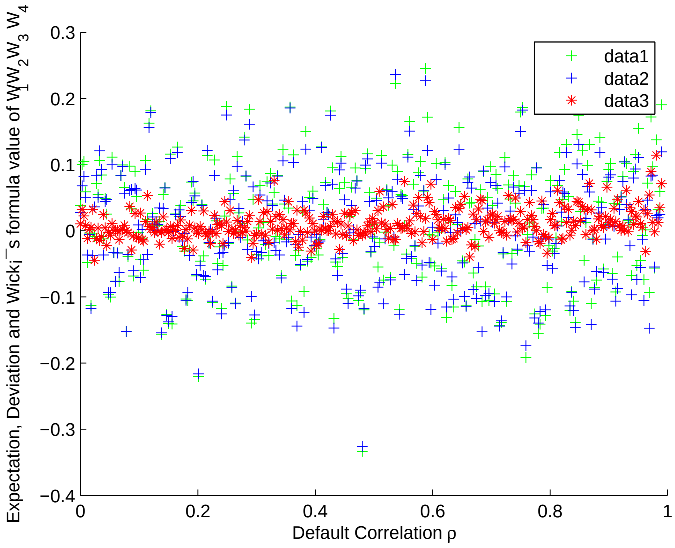

FIGURE 5.15. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 6$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

Data 1 (green +) display the conditional expectations. The mean is 0.01189. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.01035. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.001541.

For Figure5.16, conditioning on only $X_5 < -1$, $X_6 < -1$ and $X_7 < -1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.02245. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.01141. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.01104.

For Figure5.17, conditioning on only $X_5 < 0$, $X_6 < 0$ and $X_7 < 0$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

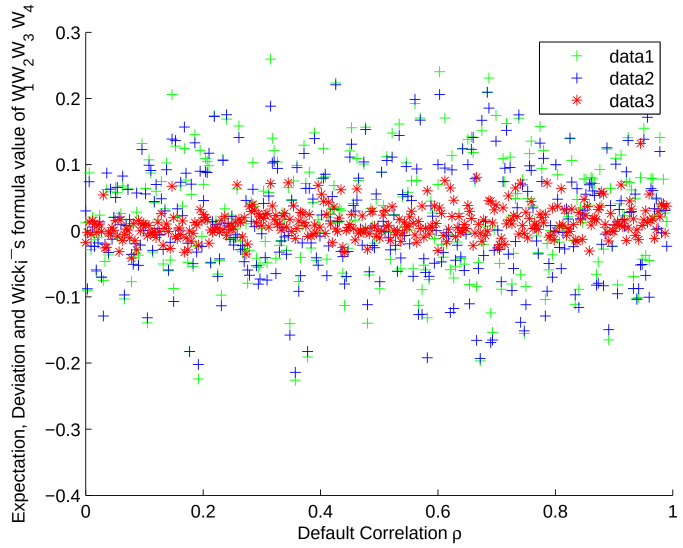

FIGURE 5.16. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 7$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

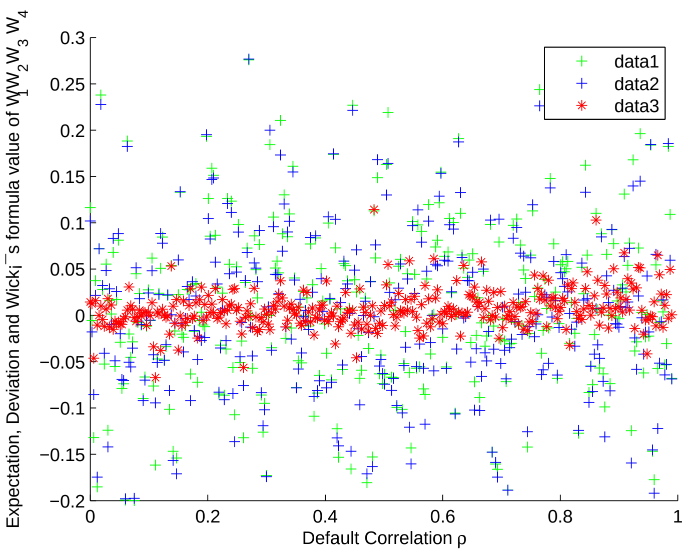

FIGURE 5.17. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 7$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

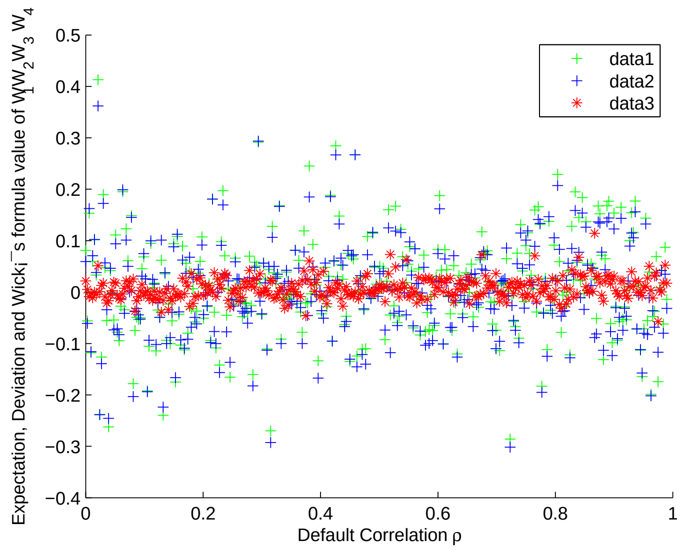

FIGURE 5.18. Conditional Expectations of Product $W_1W_2W_3W_4$ with $n = 7$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

TABLE 5.2. Deviation of Conditional Expectations of $W_1W_2W_3W_4$. For more explanation see text.

Total Names	Deviation of Conditional Expectations of $W_1W_2W_3W_4$ with Different c		
	c = -1	c = 0	c = 1
5	0.0002998	0.004434	0.02138
6	-0.005145	0.001704	0.001541
7	0.01104	0.003794	0.003218

Data 1 (green +) display the conditional expectations. The mean is 0.01028. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.006484. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.003794.

For Figure5.18, conditioning on only $X_5 < 1$, $X_6 < 1$ and $X_7 < 1$, we calculate expectations, deviations and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.01053. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.007309. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.003218.

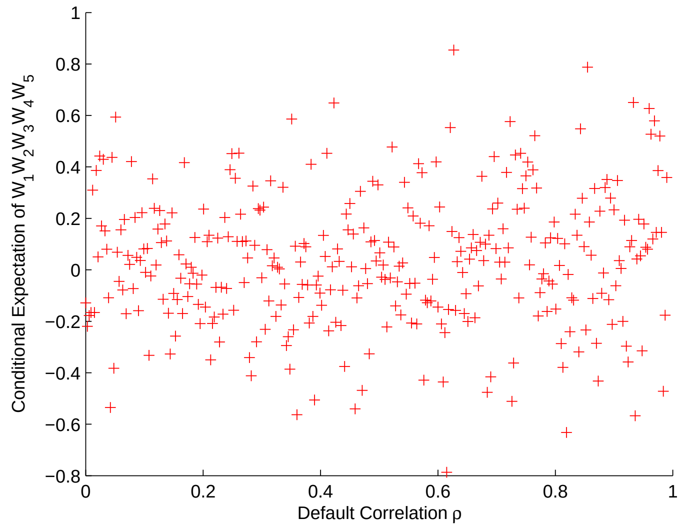

FIGURE 5.19. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 6$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

In summary of above figures of conditional expectation of $W_1W_2W_3W_4$, we have Table5.2 which displays the relationship among deviation of conditional expectations of $W_1W_2W_3W_4$, total names of portfolio, and threshold c. According to this table, we can see that, fixing threshold as -1, the deviation of conditional expectation of $W_1W_2W_3W_4$ is increasing as increasing of total names. Fixing total names, like 6 and 7, the absolute deviation seems decreasing with increasing of threshold. But for total names 5, the absolute deviation seems increasing with increasing of threshold.

Basing on these data, the deviations are close to 0 which implies that the conditional variables $W_1W_2W_3W_4$ are 'approximately' Gaussian as far as moments go, but it does not appear that they are exactly Gaussian.

Thirdly, let's see the simulation results of conditional test for five names.

For Figure5.19, conditioning on only $X_6 < -1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.03555.

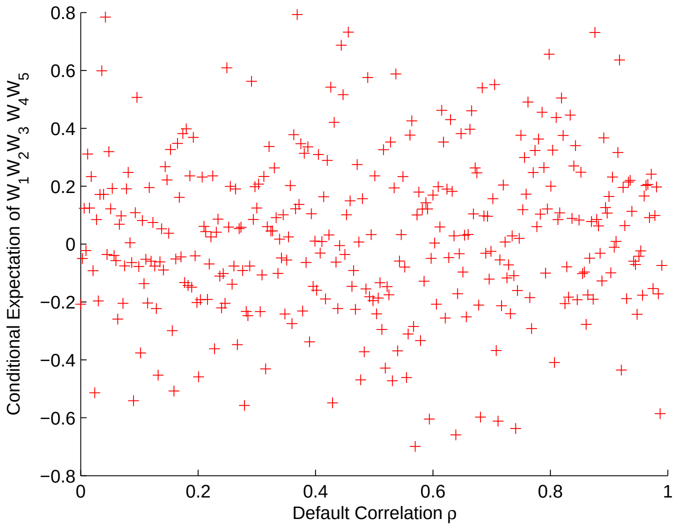

FIGURE 5.20. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 6$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

For Figure5.20, conditioning on only $X_6 < 0$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.03239.

For Figure5.21, conditioning on only $X_6 < 1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.0545.

For Figure5.22, conditioning on only $X_6 < -1$ and $X_7 < -1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.005199.

For Figure5.23, conditioning on only $X_6 < 0$ and $X_7 < 0$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.01062.

For Figure5.24, conditioning on only $X_6 < 1$ and $X_7 < 1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment

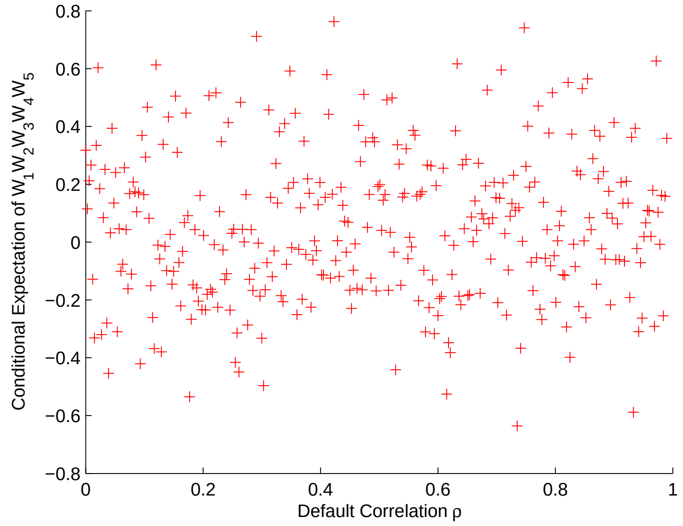

FIGURE 5.21. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 6$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

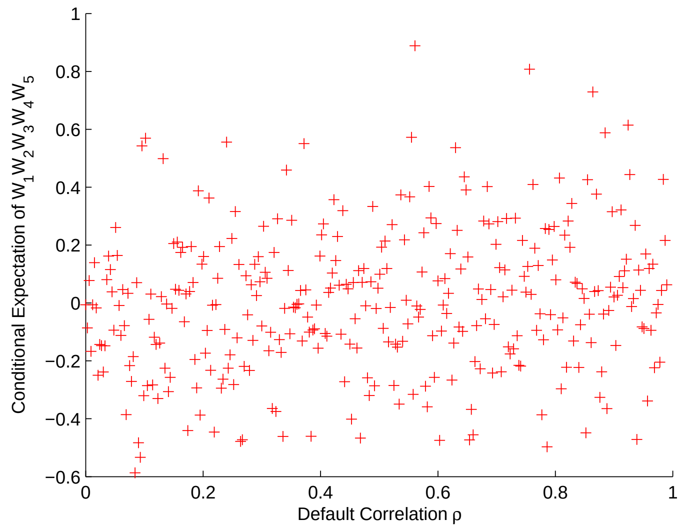

FIGURE 5.22. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 7$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

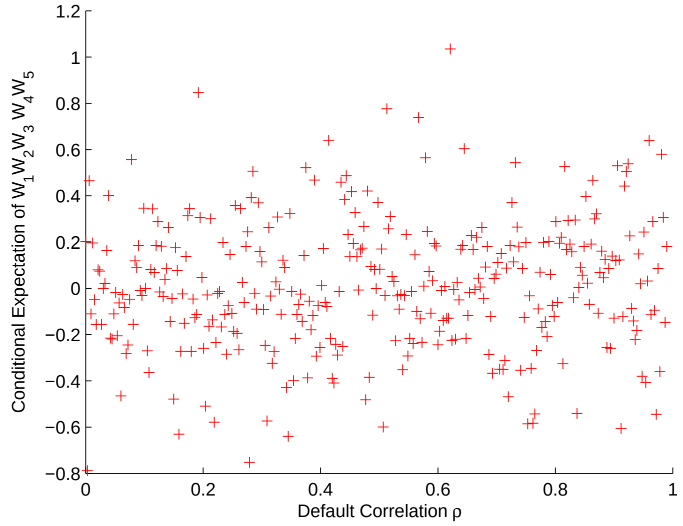

FIGURE 5.23. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 7$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

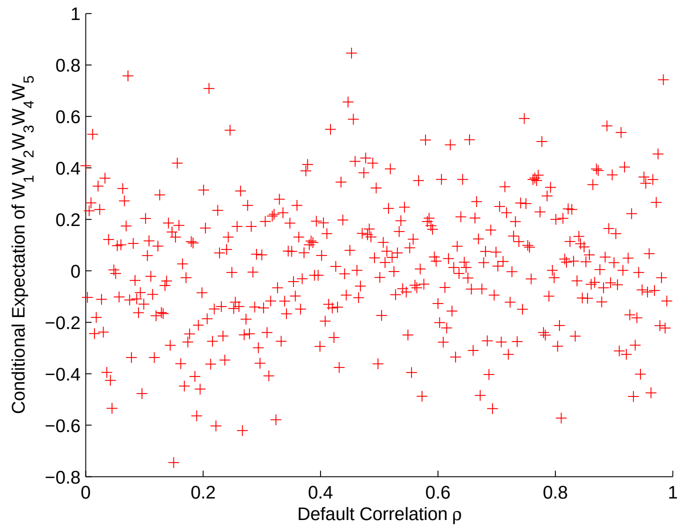

FIGURE 5.24. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 7$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

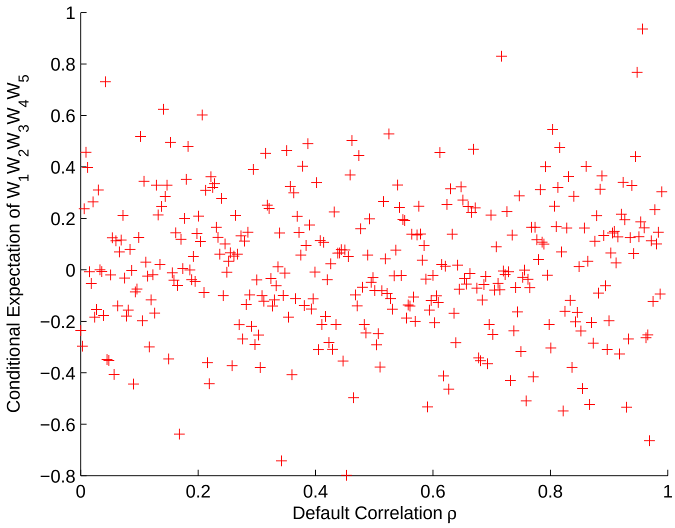

FIGURE 5.25. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 8$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$

0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.02374.

For Figure5.25, conditioning on only $X_6 < -1$, $X_7 < -1$ and $X_8 < -1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.01782.

For Figure5.26, conditioning on only $X_6 < 0$, $X_7 < 0$ and $X_8 < 0$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.003146.

For Figure5.27, conditioning on only $X_6 < 1$, $X_7 < 1$ and $X_8 < 1$, we calculate expectations of $W_1W_2W_3W_4W_5$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points is 331. The mean of expectation is 0.009398.

In summary of above figures of conditional expectation of $W_1W_2W_3W_4W_5$, we have Table5.3 which displays the relationship among mean of conditional expectations of

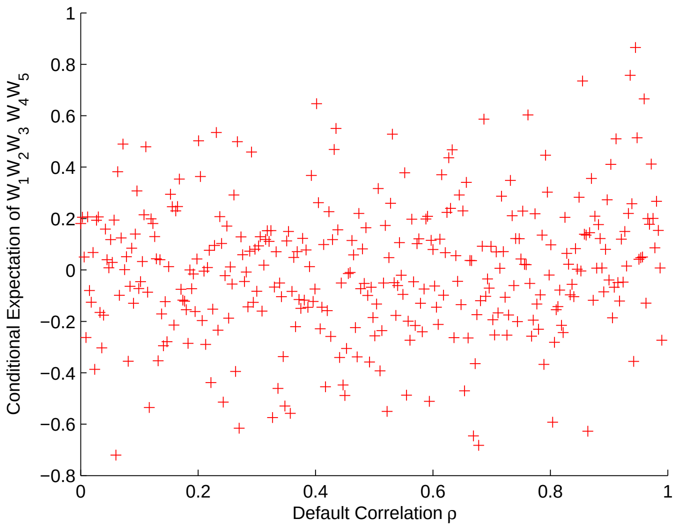

FIGURE 5.26. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 8$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$

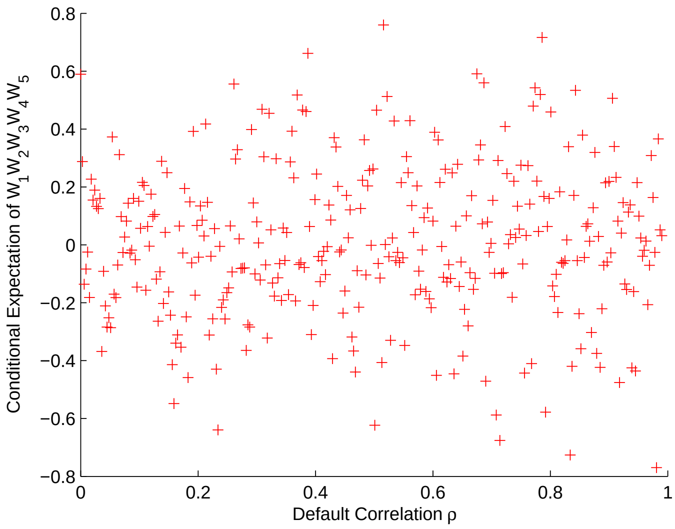

FIGURE 5.27. Conditional Expectations of Product $W_1W_2W_3W_4W_5$ with $n = 8$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$

TABLE 5.3. Mean of Conditional Expectations of $W_1W_2W_3W_4W_5$. For more explanation see text.

Total Names	Mean of Conditional Expectations of $W_1W_2W_3W_4W_5$ with Different c		
	c = -1	c = 0	c = 1
6	0.03555	0.03239	0.0545
7	0.005199	0.01062	0.02374
8	0.01782	0.003146	0.009398

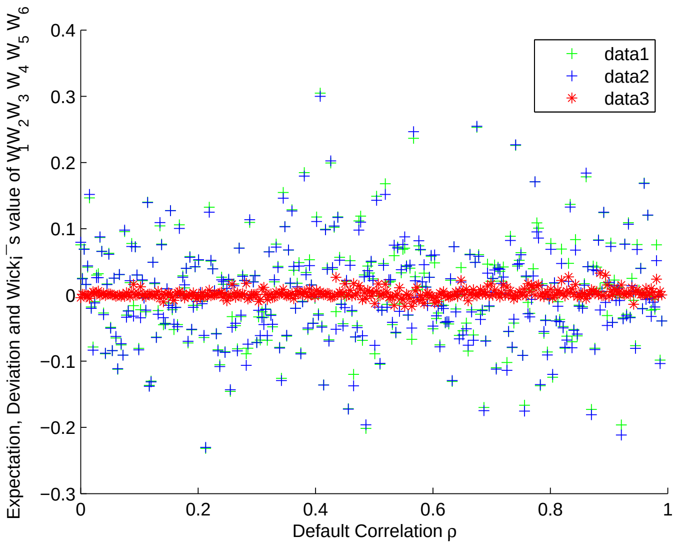

FIGURE 5.28. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 7$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

$W_1W_2W_3W_4W_5$, total names of portfolio, and threshold c. According to this table, we can see that, fixing threshold as -1, 0, and 1, the mean of conditional expectation of $W_1W_2W_3W_4W_5$ is closing to 0 as increasing of total names.

Basing on these data, the means are close to 0 which implies that the conditional variables $W_1W_2W_3W_4W_5$ are 'approximately' Gaussian as far as moments go, but it does not appear that they are exactly Gaussian.

Finally, let's see the simulation results of conditional test for six names.

For Figure5.28, conditioning on only $X_7 < -1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.003814. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.001961. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.001853.

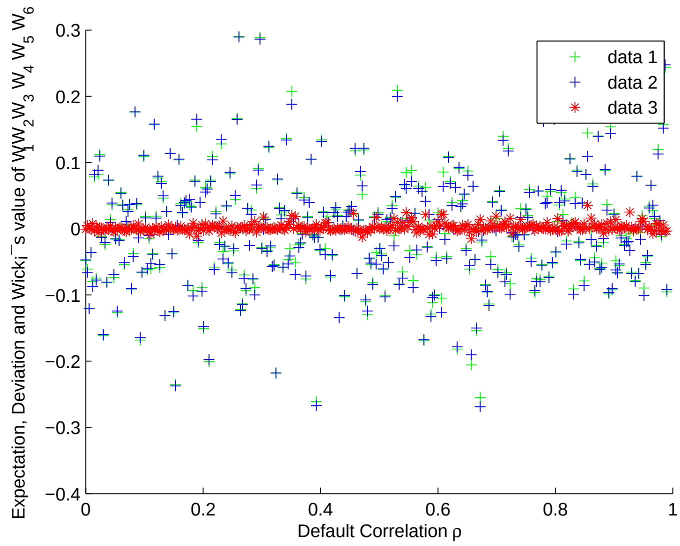

FIGURE 5.29. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 7$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

For Figure5.29, conditioning on only $X_7 < 0$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is -0.001071. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.002165. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.003236.

For Figure5.30, conditioning on only $X_7 < 1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.008645. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.001645. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.007.

For Figure5.31, conditioning on only $X_7 < -1$ and $X_8 < -1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total

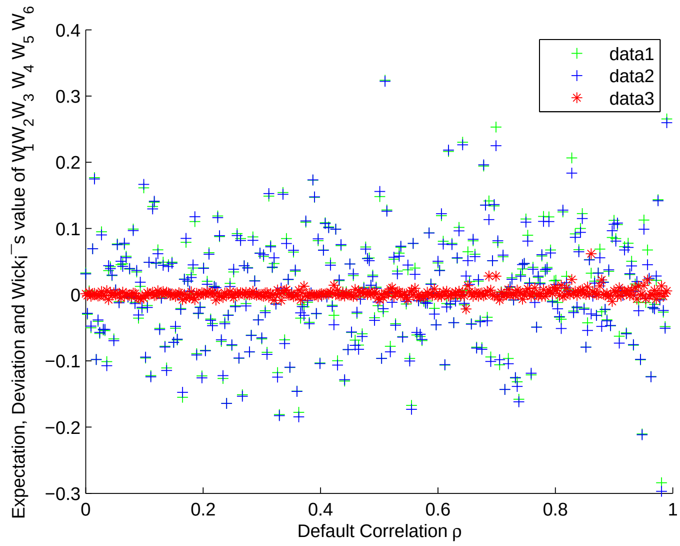

FIGURE 5.30. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 7$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

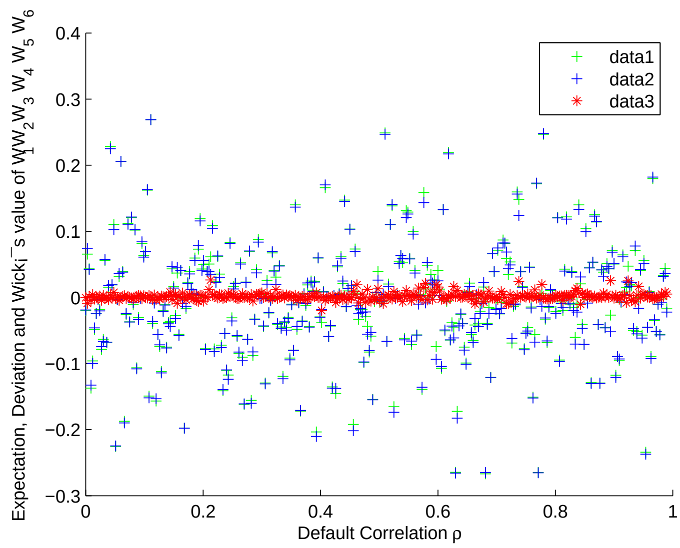

FIGURE 5.31. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 8$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

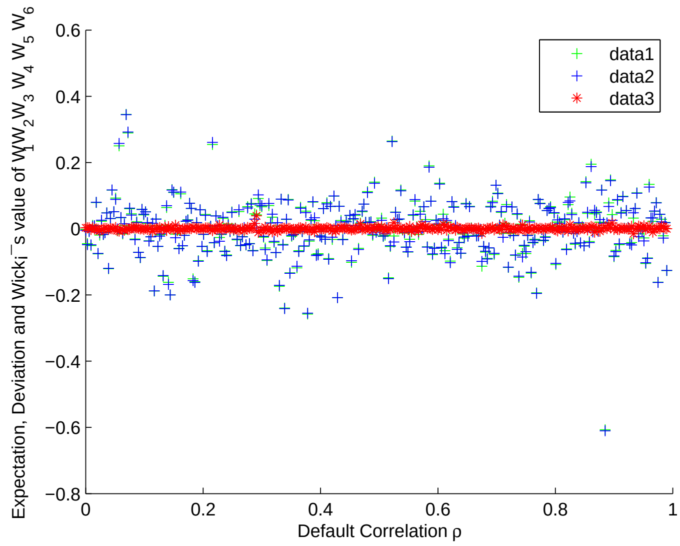

FIGURE 5.32. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 8$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is -0.00409. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.001526. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.005616.

For Figure5.32, conditioning on only $X_7 < 0$ and $X_8 < 0$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.001137. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.0005197. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.000617.

For Figure5.33, conditioning on only $X_7 < 1$ and $X_8 < 1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

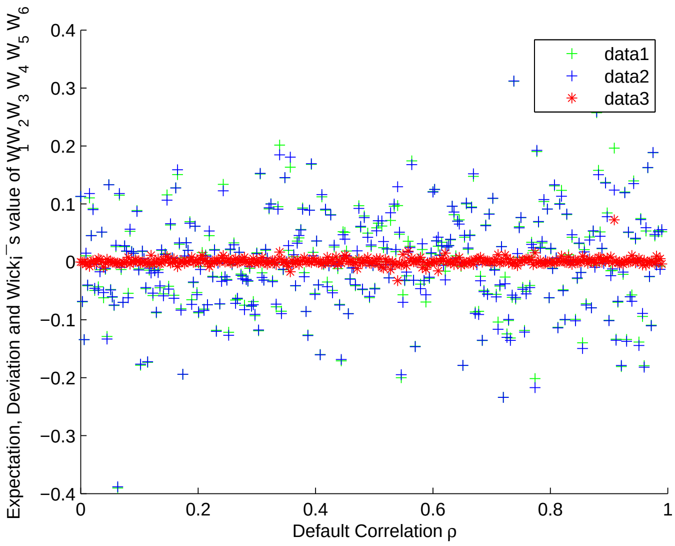

FIGURE 5.33. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 8$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

Data 1 (green +) display the conditional expectations. The mean is -0.00006741. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.000838. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.0009054.

For Figure5.34, conditioning on only $X_7 < -1$, $X_8 < -1$ and $X_9 < -1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.006231. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.0008065. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is 0.005424.

For Figure5.35, conditioning on only $X_7 < 0$, $X_8 < 0$ and $X_9 < 0$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

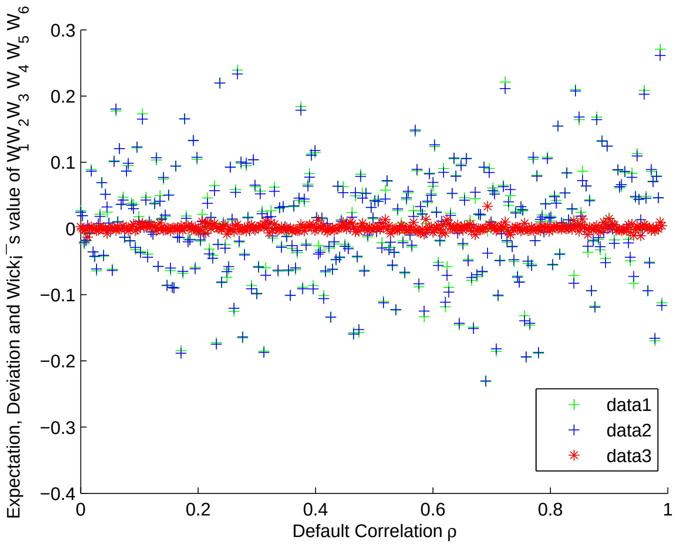

FIGURE 5.34. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 9$, $N = 100$, $c = -1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

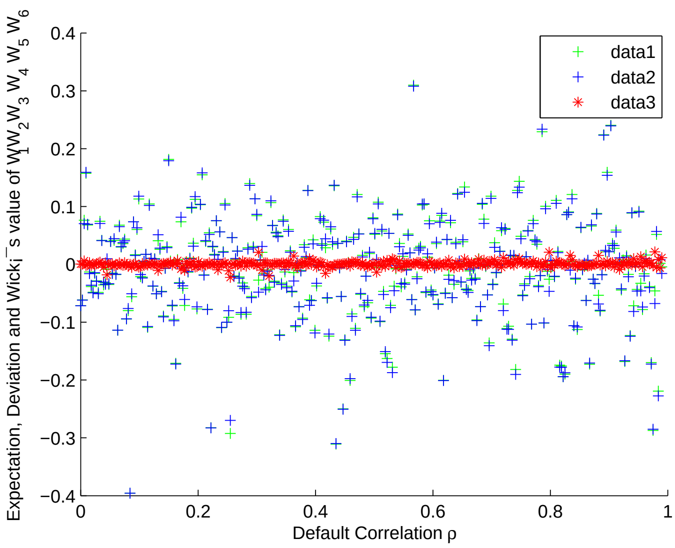

FIGURE 5.35. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 9$, $N = 100$, $c = 0$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

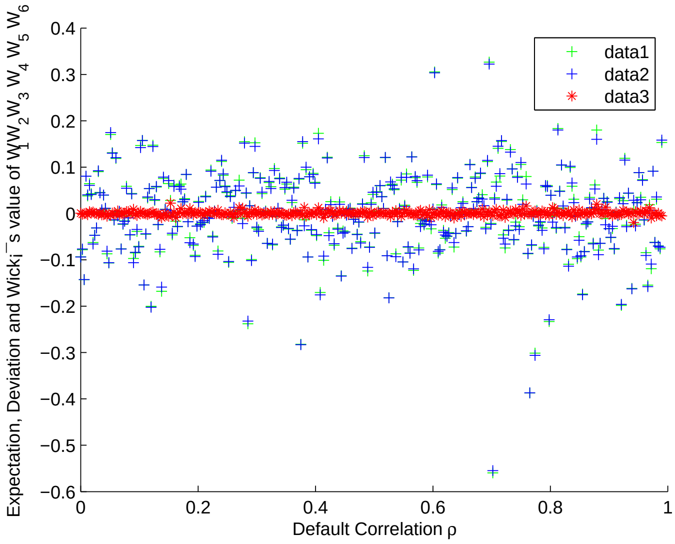

FIGURE 5.36. Conditional Expectations of Product $W_1W_2W_3W_4W_5W_6$ with $n = 9$, $N = 100$, $c = 1$, $\rho = [0 : 0.003 : 0.99]$. For more explanation see text.

TABLE 5.4. Deviation of Conditional Expectations of $W_1W_2W_3W_4W_5W_6$. For more explanation see text.

Total Names	Deviation of Conditional Expectations of $W_1W_2W_3W_4W_5W_6$ with Different c		
	c = -1	c = 0	c = 1
7	0.0002998	0.004434	0.02138
8	-0.005145	0.001704	0.001541
9	0.01104	0.003794	0.003218

Data 1 (green +) display the conditional expectations. The mean is -0.007677. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.0005521. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.008229.

For Figure5.36, conditioning on only $X_7 < 1$, $X_8 < 1$ and $X_9 < 1$, we calculate expectations, errors and Wick's formula values of $W_1W_2W_3W_4$ for different ρ on $[0, 0.99]$ whose total number is 331 with increment 0.003. The data size for each X_i is 100. The total number of points for each term is 331.

Data 1 (green +) display the conditional expectations. The mean is 0.00008357. Data 3 (red *) display the value predicted by the Wick formula. The mean is 0.0006319. Data 2 (blue +) display the deviation Data 1 minus Data 3. The mean is -0.0005483.

In summary of above figures of conditional expectation of $W_1W_2W_3W_4W_5W_6$, we have Table5.4 which displays the relationship among deviation of conditional expectations of $W_1W_2W_3W_4W_5W_6$, total names of portfolio, and threshold c. According to this table, we can see that, fixing threshold as -1, the deviation of conditional expectation of $W_1W_2W_3W_4W_5W_6$ is closing to 0 as increasing of total names. Fixing total names, like 7, the absolute deviation seems increasing with increasing of threshold.

Basing on these data, the deviations are close to 0 which implies that the conditional variables $W_1W_2W_3W_4W_5W_6$ are 'approximately' Gaussian as far as moments go, but it does not appear that they are exactly Gaussian.

5.5 Observations

Above all, extensive simulations for several values of the threshold c as well as the correlation ρ show that the conditional variables W_j are 'approximately' Gaussian as far as moments go, but it does not appear that they are exactly Gaussian.

For a more careful study of the behavior of the joint distribution of default time proxies, one should use standard tests for normality, such as kolmogorov normality test, but this is outside the range of our work.

References

[1] S.K. Chatterjee, *Statistical Thought: A Perspective and History*, Oxford University Press, Oxford, 2003.

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

[3] M. Dorey, P. Joubert, *Modeling Copulas: An Overview*, The Staple Inn Actuarial Society, 2007.

[4] P. Embrechts, A. Mcneil, D. Straumann, *Correlation and dependence in risk management: properties and pitfalls*, Risk Management: Value at Risk and Beyond, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, 176-223, 2002.

[5] E.W. Frees and E.A. Valdez, *Understanding Relationships Using Copulas*, North American Actuarial Journal, 2, 1-25, 1998.

[6] A. Friend, E. Rogge, *Correlation at First Sight*, Economic Notes: Review of Banking, Finance and Monetary Economics, 2004.

[7] C. Genest, J. MacKay, *Copulas archimédiennes et familles de lois bidimensionnelles dont les marges sont données*, Canad. J. Statist. 14, 145-159.

[8] E. Hillebrand, A.N. Sengupta and J.Y. Xu, *Temporal Correlation of Defaults in Subprime Securitization*,16 pages, submitted, available at http://www.math.lsu.edu/ sengupta/papers.

[9] WH. Kruskal, *Ordinal measures of association*, J. Amer. Statist. Assoc., 53,814-861, 1958.

[10] David X. Li, *On Default Correlation: A copula Function Approach*, Journal of Fixed Income, 9, 43-54, 2000.

[11] D. Lucus, *Default Correlation and Credit Analysis*, Journal of Fixed Income, Vol.11, 76-87, 1995.

[12] C. Meng and A.N. Sengupta, *CDO tranche sensitivities in the Gaussian Copula Model*, 16 pages, submitted, available at http://www.math.lsu.edu/ sengupta/papers.

[13] R. Merton, *On the pricing of corporate debt: the risk structure of interest rates*, Journal of Finance, 29, 449-470, 1074.

[14] Roger B. Nelsen, *An Introduction to Copulas*, Second Edition, Springer, New York, 2006.

[15] L. Rüschendorf, *On the distributional transform, Sklar's Theorem, and the empirical copula process*, J. Statist. Plan. Infer., 139 (11), 3921 - 3927, 2009.

[16] A. Sklar, *Fonctions de réepartiton à n dimension et leurs marges*, Publ. Inst. Statist. Univ. Paris. 8, 229-231 (1959).

[17] A. Wald, *Sequential Analysis*, Wiley, New York, 1976.

[18] B. Wlodek and P. Virgil, *Seminar talk: Wick formula for quaternion normal laws*,20 pages, available at http://www.math.uc.edu/ brycw/preprint/quaaternion-wick-08.pdf.

Appendix A: CopulaConditionalTest03.m

The following is a code in Matlab.

```
function CopulaConditionalTest03(n, p, N, c)

global EEEE

for i=1:N
A = randn(1,1);
Z = repmat(A, 1, n);
E = randn(1, n);
X=sqrt(p)*Z + sqrt(1-p)*E;

while min(X(1:3))<=c | norm(X(4:n))>c

      A = randn(1,1);
      Z = repmat(A, 1, n);
      E = randn(1, n);
X = sqrt(p)*Z + sqrt(1-p)*E
end

Y(i,1:3)=X(1:3);
end

Y; for i = 1:N
Y_1(i)=(sum(Y(:,1)<=Y(i,1)))/N;
Y_2(i)=(sum(Y(:,2)<=Y(i,2)))/N;
Y_3(i)=(sum(Y(:,3)<=Y(i,3)))/N;
end

MU = zeros(N,1);
SIGMA = ones(N,1);

Z_1 = norminv(Y_1', MU, SIGMA);

Z_2 = norminv(Y_2', MU, SIGMA);

Z_3 = norminv(Y_3', MU, SIGMA);

ZZ = [Z_1,Z_2,Z_3]

[D,F]=find(ZZ==inf);
ZZ(D,:)=[];
ZZ;
```

```
[B,C] = size(ZZ);

for i = 1:B
E(i) = prod(ZZ(i,:));
end

EEEE = (sum(E))/B;
```

Appendix B: CopulaConditionalTest04.m

The following is a code in Matlab.

```
function CopulaConditionalTest04(n, p, N, c)

global Error
global Wick
global EEEE

for i=1:N
A = randn(1,1);
Z = repmat(A, 1, n);
E = randn(1, n);
X=sqrt(p)*Z+sqrt(1-p)*E;

while min(X(1:4))<=c | max(X(5:n))>c
      A = randn(1,1);
      Z = repmat(A, 1, n);
      E = randn(1, n);
X = sqrt(p)*Z + sqrt(1-p)*E;
end

Y(i,1:4)=X(1:4);
end

Y;
for i = 1:N
Y_1(i) = (sum(Y(:,1)<=Y(i,1)))/N;

Y_2(i) = (sum(Y(:,2)<=Y(i,2)))/N;

Y_3(i) = (sum(Y(:,3)<=Y(i,3)))/N;

Y_4(i) = (sum(Y(:,4)<=Y(i,4)))/N;
end

[m,n] = size(Y_1);

MU = zeros(N,1);

SIGMA = ones(N,1);

Z_1 = norminv(Y_1', MU, SIGMA);
```

```
Z_2 = norminv(Y_2', MU, SIGMA);

Z_3 = norminv(Y_3', MU, SIGMA);

Z_4 = norminv(Y_4', MU, SIGMA);

ZZ = [Z_1,Z_2,Z_3,Z_4];
[D,F]=find(ZZ==inf);
ZZ(D,:)=[];
ZZ;
[B,C]=size(ZZ);

for i = 1:B
E(i) = prod(ZZ(i,:));
end

EEEE = (sum(E))/B
%We want to test if E(Z_1Z_2Z_3Z_4) = E(Z_1Z_2)E(Z_3Z_4)+
%E(Z_1Z_3)E(Z_2Z_4)+ E(Z_1Z_4)E(Z_2Z_3)

EE12 = (ZZ(:,1)'* ZZ(:,2))/B; %E(Z_1Z_2)
EE34 = (ZZ(:,3)'* ZZ(:,4))/B; %E(Z_3Z_4)
EE13 = (ZZ(:,1)'* ZZ(:,3))/B; %E(Z_1Z_3)
EE24 = (ZZ(:,2)'* ZZ(:,4))/B; %E(Z_2Z_4)
EE14 = (ZZ(:,1)'* ZZ(:,4))/B; %E(Z_1Z_4)
EE23 = (ZZ(:,2)'* ZZ(:,3))/B; %E(Z_2Z_3)

Wick = EE12*EE34 + EE13*EE24 + EE14*EE23;

Error = EEEE - (EE12*EE34 + EE13*EE24 + EE14*EE23) %value of difference
```

Appendix C: CopulaConditionalTest05.m

The following is a code in Matlab.

```
function CopulaConditionalTest05(n, p, N, c)

global EEEE

for i=1:N %
A = randn(1,1);
Z = repmat(A, 1, n);
E = randn(1, n);
X = sqrt(p)*Z+sqrt(1-p)*E;

while min(X(1:5))<=c | norm(X(6:n))>c

      A = randn(1,1);
      Z = repmat(A, 1, n);
      E = randn(1, n);
X = sqrt(p)*Z + sqrt(1-p)*E
end

Y(i,1:5)=X(1:5);
end

Y; for i = 1:N
Y_1(i)=(sum(Y(:,1)<=Y(i,1)))/N;
Y_2(i)=(sum(Y(:,2)<=Y(i,2)))/N;
Y_3(i)=(sum(Y(:,3)<=Y(i,3)))/N;
Y_4(i)=(sum(Y(:,3)<=Y(i,3)))/N;
Y_5(i)=(sum(Y(:,3)<=Y(i,3)))/N;
end

MU = zeros(N,1);

SIGMA = ones(N,1);

Z_1 = norminv(Y_1', MU, SIGMA);

Z_2 = norminv(Y_2', MU, SIGMA);

Z_3 = norminv(Y_3', MU, SIGMA);

Z_4 = norminv(Y_4', MU, SIGMA);
```

```
Z_5 = norminv(Y_5', MU, SIGMA);

ZZ = [Z_1,Z_2,Z_3,Z_4,Z_5];

[D,F]=find(ZZ==inf);
ZZ(D,:)=[]; ZZ;
[B,C] = size(ZZ);

for i = 1:B
    E(i) = prod(ZZ(i,:));
end

EEEE = (sum(E))/B;
```

Appendix D: CopulaConditionalTest06.m

The following is a code in Matlab.

```
function CopulaConditionalTest06(n, p, N, c)

 global Error;
 global Wick;
 global EEEE;

for i=1:N %
A = randn(1,1);
Z = repmat(A, 1, n);
E = randn(1, n);
X=sqrt(p)*Z+sqrt(1-p)*E;

while min(X(1:3))<=c | norm(X(4:n))>c
      A = randn(1,1);
      Z = repmat(A, 1, n);
      E = randn(1, n);
X = sqrt(p)*Z + sqrt(1-p)*E
end

Y(i,1:6)=X(1:6);
end

Y; for i = 1:N
Y_1(i)=(sum(Y(:,1)<=Y(i,1)))/N;
Y_2(i)=(sum(Y(:,2)<=Y(i,2)))/N;
Y_3(i)=(sum(Y(:,3)<=Y(i,3)))/N;
Y_4(i)=(sum(Y(:,4)<=Y(i,4)))/N;
Y_5(i)=(sum(Y(:,5)<=Y(i,5)))/N;
Y_6(i)=(sum(Y(:,6)<=Y(i,6)))/N;
end

[m,n] = size(Y_1);

MU = zeros(N,1);

SIGMA = ones(N,1);

Z_1=norminv(Y_1', MU, SIGMA);
Z_2=norminv(Y_2', MU, SIGMA);
Z_3=norminv(Y_3', MU, SIGMA);
Z_4=norminv(Y_4', MU, SIGMA);
```

```
Z_5=norminv(Y_5', MU, SIGMA);
Z_6=norminv(Y_6', MU, SIGMA);

ZZ = [Z_1,Z_2,Z_3,Z_4,Z_5,Z_6];
[D,F]=find(ZZ==inf);
ZZ(D,:)=[];
ZZ;
[B,C]=size(ZZ);

for i = 1:B
E(i) = prod(ZZ(i,:));
end

EEEE=sum(E)/B

EE12 = (ZZ(1:B,1)'* ZZ(1:B,2))/B; %E(X_1X_2)
EE34 = (ZZ(1:B,3)'* ZZ(1:B,4))/B; %E(X_3X_4)
EE56 = (ZZ(1:B,5)'* ZZ(1:B,6))/B; %E(X_5X_6)
EE35 = (ZZ(1:B,3)'* ZZ(1:B,5))/B; %E(X_3X_5)
EE46 = (ZZ(1:B,4)'* ZZ(1:B,6))/B; %E(X_4X_6)
EE36 = (ZZ(1:B,3)'* ZZ(1:B,6))/B; %E(X_3X_6)
EE45 = (ZZ(1:B,4)'* ZZ(1:B,5))/B; %E(X_4X_5)
EE13 = (ZZ(1:B,1)'* ZZ(1:B,3))/B; %E(X_1X_3)
EE24 = (ZZ(1:B,2)'* ZZ(1:B,4))/B; %E(X_2X_4)
EE25 = (ZZ(1:B,2)'* ZZ(1:B,5))/B; %E(X_2X_5)
EE26 = (ZZ(1:B,2)'* ZZ(1:B,6))/B; %E(X_2X_6)
EE14 = (ZZ(1:B,1)'* ZZ(1:B,4))/B; %E(X_1X_4)
EE23 = (ZZ(1:B,2)'* ZZ(1:B,3))/B; %E(X_2X_3)
EE15 = (ZZ(1:B,1)'* ZZ(1:B,5))/B; %E(X_1X_5)
EE16 = (ZZ(1:B,1)'* ZZ(1:B,6))/B; %E(X_1X_6)

Wick = EE12*EE34*EE56 + EE12*EE35*EE46 + EE12*EE36*EE45 +
EE13*EE24*EE56 + EE13*EE25*EE46 + EE13*EE26*EE45 + EE14*EE23*EE56 +
EE14*EE25*EE36 + EE14*EE26*EE35 + EE15*EE23*EE46 + EE15*EE24*EE36 +
EE15*EE26*EE34 + EE16*EE23*EE45 + EE16*EE24*EE35 + EE16*EE25*EE34

Error = EEEE - (EE12*EE34*EE56 + EE12*EE35*EE46 + EE12*EE36*EE45 +
EE13*EE24*EE56 + EE13*EE25*EE46 + EE13*EE26*EE45 + EE14*EE23*EE56 +
EE14*EE25*EE36 + EE14*EE26*EE35 + EE15*EE23*EE46 + EE15*EE24*EE36 +
EE15*EE26*EE34 + EE16*EE23*EE45 + EE16*EE24*EE35 + EE16*EE25*EE34)
```

Vita

Dongxiang Yan was born in June 1981, in Anhui, People's Republic of China. He finished his undergraduate studies at Nanjing University of Science and Technology in June 2002 where he earned a Bachelor of Engineering degree in pharmaceutical engineering. He finished his graduate studies at Shanghai Jiaotong University in June 2006 where he earned a Master of Medicine degree in pharmaceutical chemistry. In January 2008 he came to Louisiana State University to pursue graduate studies in mathematics. He is currently a candidate for the degree of Master of Science in mathematics with concentration in finance, which will be awarded in August 2010.