Estimation of Value-at-Risk and Expected Shortfall using Copulas

Helina Sumbhoolaul

Submitted to the Department of Statistical Sciences
in partial fulfillment of the requirements for the degree of

Master of Science in Statistical Sciences
at the
UNIVERSITY OF CAPE TOWN

April 21, 2008

Supervisor: Professor Linda Haines

Abstract

Recent empirical evidence has shown that the joint distributions of most assets returns are not elliptical, an example being the joint distribution of equity returns. As a consequence, the correlation coefficient between assets does not provide an adequate measure of the dependence structure. Copulas are distribution functions whose parameter provides a dependence measure and they can be used even when the risky assets in a portfolio are not jointly elliptically distributed. A copula can assume different functional forms and connects together multiple univariate distribution functions to form a multivariate distribution function that captures the dependence of the underlying assets (Dowd, 2005a).

This mini-dissertation explores a copula-based approach for computing two popular measures of risk, namely Value-at-Risk (VaR) and expected shortfall (ES). Three estimation methods of the copula parameter have been used, namely maximum likelihood estimation (MLE), inference function for margins (IFM) and canonical maximum likelihood (CML).

A simulation study has been performed to compare the properties and performance of the VaR and ES estimates obtained using the three estimation methods. In the simulation study, two copulas have been considered, namely the Frank and the Gaussian copulas with normal and Student's t distributed marginals. The results show that the VaR and ES estimates obtained from the three estimation methods are very similar, in particular the estimates obtained from the MLE and IFM estimation methods, but in terms of performance the MLE is the preferred method. The CML method also yielded satisfactory results and the estimates were easily computed and were similar to the MLE estimates. However the simulation results also showed that using the CML method may come at the cost of the performance of the estimates.

A real data set has been used to illustrate the copula-based approach of estimating the VaR and the ES of a portfolio of two equally-weighted risky assets, namely Anglo and the Top 40 Index, using five years of daily closing prices from the 11^{th} March 2002 to the 7^{th} March 2007. The properties of the VaR and ES estimates have been investigated and the VaR bounds of the portfolio have been computed. The backtesting technique has been used to evaluate the performance of the VaR estimates generated from the fitted models. However, the results obtained were somewhat disappointing in that the percentage of exceedance levels of the VaR estimates were not close to the 5% nominal level.

Plagiarism Declaration

1. This dissertation is my own work. It has not been submitted before for any degree or examination to any other University.

2. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

3. Each significant contribution to, and quotation in, this dissertation from the work of other people has been cited and referenced.

Signed by candidate

Signature

24/04/2008

Date

Acknowledgements

The financial assistance of the Department of Labour (DST) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DST.

I would like to thank my supervisor, Professor Linda Haines, for her help, guidance and commitment.

I would also like to thank my parents, Kishore and Geeta Sumbhoolaul as well as my fiancé Avinash Andhee, my father-in-law Krish Andhee and my sister Varsha Sumbhoolaul for their unconditional love and moral support.

Lastly I would like to thank Dominique Katshunga for helping me get started and Craig Anderson for helping me with LaTeX.

Contents

List of Figures vii

List of Tables ix

1 Introduction 1

1.1 Background 1

1.2 Rationale for using copula to model dependencies 2

1.3 Aims of the mini-dissertation 3

1.4 Outline of the mini-dissertation 4

2 Copulas 5

2.1 Introduction to copulas 5

2.1.1 Mathematical Definition 5

2.1.2 Sklar's Theorem 6

2.1.3 Statistical Interpretation 6

2.2 Properties 7

2.2.1 Invariance Property 7

2.2.2 The Fréchet-Hoeffding Bounds 7

2.3 Dependence Concepts 8

2.3.1 Linear Correlation 8

2.3.2 Perfect dependence and perfect independence 9

2.3.3 Concordance 10

2.3.4 Tail Dependence 14

2.3.5 Positive Quadrant Dependence 15

2.4 Multivariate Copulas . . . 15
2.5 Common parametric families of bivariate copulas . . . 15
2.5.1 Elliptical Copulas . . . 16
2.5.2 Archimedean Copulas . . . 20
2.6 Estimation procedures for copulas . . . 25
2.6.1 Parametric estimation procedures . . . 25
2.6.2 Semiparametric estimation procedures . . . 28
2.6.3 Nonparametric estimation . . . 29
2.6.4 Choosing a copula from the Archimedean family . . . 29

3 Value-at-Risk and Expected Shortfall 32
3.1 Description of a two-asset portfolio . . . 32
3.2 Value-at-Risk . . . 33
3.2.1 Definition of VaR . . . 33
3.2.2 Estimation of VaR . . . 34
3.2.3 Limitations of VaR . . . 35
3.3 Expected shortfall . . . 35
3.3.1 Definition of expected shortfall . . . 36
3.3.2 Estimation of expected shortfall . . . 36
3.3.3 Limitations of expected shortfall . . . 36
3.4 Computation of VaR and expected shortfall using copulas . . . 36
3.4.1 Computation of VaR and ES by integration . . . 37
3.4.2 Computation of VaR and expected shortfall by Monte Carlo simulation . . . 38
3.5 Computation of VaR bounds . . . 39

4 The simulation study 42
4.1 The objectives of the simulation process . . . 42
4.2 Scenario . . . 43
4.3 Procedures . . . 44
4.4 Properties of the estimators . . . 46

4.4.1 Stage 1: Generate a "real" data set 47

4.4.2 Stage 2: Obtaining the VaR and ES estimates 47

4.4.3 Stage 3: The bootstrapping process 48

4.4.4 Results 51

4.5 The performance of the the VaR and ES estimates 56

4.5.1 The procedure used 56

4.5.2 Results 58

5 Risk measures for a real data set 60

5.1 Description of the portfolio and the data 60

5.2 Selection of the marginals and the copula 63

5.3 Estimating the risk measures of the portfolio 65

5.3.1 Estimation of the marginals and the copula 65

5.3.2 Estimation of VaR and expected shortfall 68

5.3.3 Properties of the VaR and the ES estimates 69

5.4 Computing the VaR bounds 69

5.5 Backtesting the VaR estimates 72

6 Conclusions 74

Bibliography 76

List of Figures

2.1 The graphs of the minimum, maximum and the product copulas. . . . 11

2.2 The contour plots of the minimum, maximum and the product copulas. 12

2.3 The Gaussian copula, its density function and its contour plot for ρ=0.9. 18

2.4 The Student's t copula, its density function and its contour plot for ρ=0.9, ν=3. 19

2.5 The distribution functions of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas. 22

2.6 The density functions of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas. 23

2.7 The contour plots of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas. 24

2.8 A sample of random variables drawn from two distributions having the same standard normal marginal distributions and the same τ=0.7 but different dependence structures. 26

4.1 Flowchart of the computation of the VaR and ES of the two-asset portfolio. 46

4.2 Flowchart of the algorithm used to perform Stage 1 and Stage 2. . . . 48

4.3 Flowchart of the algorithm used to generate the bootstrap samples for the VaR and ES estimates. 50

4.4 Histograms of the bootstrapped VaR and ES estimates obtained for Scenario 1. 52

4.5 Flowchart of the algorithm used for running the simulation process S times. 58

5.1 The histograms of the log returns of Anglo and the TOP 40. 62

5.2 The scatter plot of the log returns of the TOP 40 against the log returns of Anglo. 62

5.3 Plot of the parametric and nonparametric estimates of $K(z)$ against the corresponding z values for the Clayton, Frank and Gumbel copulas. 64

5.4 The fit of the t-location scale distribution to the log returns of Anglo and the TOP 40. 66

5.5 Histograms of the bootstrapped samples of the VaR and ES estimates obtained from the MLE, IFM and CML methods. 71

List of Tables

2.1 The generator function of the Frank, Clayton and Gumbel copulas. . 20

2.2 Kendall's tau and Spearman's rho of the Frank, Clayton and Gumbel copulas. 21

4.1 The twelve scenarios that were used in the simulation study. 44

4.2 The three VaR estimates and the three ES estimates for the twelve scenarios. 51

4.3 The mean, standard error, skewness, kurtosis and bias of the three VaR estimates for the twelve scenarios obtained from the bootstrap procedure. 53

4.4 The mean, standard error, skewness, kurtosis and bias of the three ES estimates for the twelve scenarios obtained from the bootstrap procedure. 54

4.5 The bias and MSE of the VaR and ES estimates obtained by repeating the simulation process $S = 500$ times. 59

5.1 Some descriptive statistics of the log returns of Anglo and the TOP 40. 61

5.2 The estimates of the location, scale and degrees of freedom of the distribution of the log returns of Anglo and the TOP 40. 66

5.3 The MLE, IFM and CML parameter estimates. 67

5.4 The descriptive statistics of the bootstrapped sample of the VaR and ES estimates. 70

5.5 The MLE, IFM and CML bounds which were estimated at the 95% confidence level. 72

5.6 The exceedance level of the VaR estimates obtained from the MLE , IFM and CML methods. 73

Chapter 1

Introduction

1.1 Background

As financial markets across the world are becoming increasingly inter-connected, financial institutions are faced with more competition and are therefore pushed to take more risks in an attempt to increase returns. Hence the ability of financial institutions to effectively identify and manage risks so that they can maintain a competitive edge and thus ensure their survival in the business world is crucial. Risk management, more specifically market risk, which is the risk of incurring a loss due to adverse market movements, has received ample attention in the literature. This is partly due to regulatory capital requirements imposed by the Basel II Accord established in 1998 as well as an attempt to cushion from unacceptable losses during a financial market crisis (Luciano and Marena, 2002).

Two popular numerical measures of market risk are Value-at-Risk (VaR) and expected shortfall (ES). Broadly speaking, VaR measures by how much the market value of a portfolio of risky assets can drop over a given period of time and with a specified confidence level, due to adverse market movements and expected shortfall is given by the average of losses that fall below the VaR estimate (Dowd and Blake, 2006). The estimates of VaR and ES are determined by the joint distribution of the returns of the risky assets in a portfolio such that their success as an effective risk metric depends on the suitability of the choice of the multivariate distribution of the underlying assets (Roch and Alegre, 2006).

There is abundant empirical evidence that the distributions of the individual asset returns are non-Gaussian. For example, they are known to exhibit excess kurtosis, also known as fat tails, and skewness (Patton, 2006b). At the univariate level, this implies that the probability of the occurrence of extreme events compared with the normal distribution is amplified (Patton, 2006a). In a portfolio setting where more than one risky asset is involved, there is usually dependencies between the underlying assets and not considering such dependencies can lead to a significant underestimation of the risk of the portfolio (Roch and Alegre, 2006). Therefore properly identifying the dependencies between the underlying instruments is crucial for portfolio managers to take decisions concerning which assets to include in the portfolio for diversification purposes and the optimal portfolio weights (Patton, 2006a). In most cases, the simplifying assumption made is that the joint distribution of the underlying assets returns is normally or more generally speaking, elliptically distributed, in which case modelling the dependence between the returns using Pearson's correlation coefficient is justified (Dowd, 2005a). However, recent empirical evidence shows that the joint distributions of most assets returns are not elliptical, an example being the joint distribution of equity returns. As a consequence, the correlation coefficient between assets does not provide an adequate measure of the dependence structure (Dowd, 2005a; Patton, 2006b).

When the joint distribution of the underlying assets in the portfolio is not elliptically distributed, it implies that there are asymmetries in the dependence structure (Embrechts et al., 2001). One example of asymmetric dependence is that asset returns tend to exhibit stronger correlation during market crashes than when the market is performing well (Embrechts, et al. 2001; Patton, 2006b). One possible reason for this phenomenon is the negative market sentiment emanating from investors who fear that the economy is going down during a financial market crisis and this causes the correlation between stock markets to increase (Patton, 2006a). Furthermore, financial instruments are becoming more sophisticated thus rendering their dependency structures more complex (Cherubini et al., 2004). This is why there is a need for a measure of dependence between asset returns which provides more flexibility other than that describing the dependence structure of multivariate elliptical distributions.

1.2 Rationale for using copula to model dependencies

Copulas are distribution functions whose parameter provides a dependence measure and they can be used even when the risky assets in a portfolio are not jointly elliptically distributed. Copulas were introduced in a statistical sense by Abe Sklar in 1959 but their use in finance and economics has only recently attracted significant attention, as evidenced by the number of papers that have been published in that area (Hürlimann, 2002; Patton, 2006a). More specifically, there has been a growing interest in using copulas for modelling multivariate dependence in a number of fields including finance. Frees and Valdez (1997) and Embrechts et al. (2002) provide a good introduction to the dependence properties of copulas and also highlight the weaknesses of the correlation coefficient. A mathematical and statistical introduction to copulas can be found in Joe (1997) and in Nelsen (2006) as well as in an article by Embrechts et al. (2001). On the other hand, Luciano and Marena (2003) and Cherubini et al. (2004) demonstrate the financial applications of copulas, for example in risk management, in the pricing of financial instruments and in credit risk. Furthermore, Patton (2006a) provides a survey on the applications of copulas in the fields of finance and economics whilst an article by Dowd (2005a) focuses on the application of copulas to risk management.

A copula can assume different functional forms and connects together multiple univariate distribution functions to form a multivariate distribution function that captures the dependence of the underlying assets in the portfolio (Dowd, 2005a). The parameter of the copula measures the degree of dependence between the underlying assets where the larger the absolute value of the copula parameter, the greater the dependence (Hürlimann, 2004). Copulas provide a flexible way of modelling the joint distribution of risky assets by separately modelling the univariate distributions and the dependence structure between them (Patton, 2006a). This two-step approach enables the construction of a multivariate distribution function having different univariate distribution functions. Another advantage of using copulas is that there is a large selection of well-documented families of copula in the literature which cater for a broad range of different dependence structures (Dowd, 2005a). Copulas also provide an excellent tool to stress test portfolios, which are procedures used in an attempt to estimate the vulnerability of portfolios, in the event of extreme adverse moves in the market (Cherubini et al., 2004; Dowd, 2005b).

1.3 Aims of the mini-dissertation

The degree to which portfolio and risk managers rely upon the VaR metric computed by assuming normality, is quite alarming in the case of non-normal dependence between the risky assets in the portfolio (Patton, 2006a). To address this problem, a number of authors including Luciano and Marena (2003), Cherubini et al. (2004) and Dowd (2005a) have taken a copula-based approach in order to compute VaR.

The aims of this mini-dissertation are as follows:

1. To provide an introduction to copulas and demonstrate their application in the computation of VaR and ES of a two-asset portfolio.

2. To investigate the properties and performance of the VaR and ES estimates obtained from various estimation techniques of the copula parameter and to compare the VaR and ES estimates obtained from different copulas and marginals, through a simulation study.

3. To illustrate and critically appraise the estimation of the copula-based VaR and ES of a portfolio using a set of real data, within a South African context.

1.4 Outline of the mini-dissertation

A broad introduction to the theory of copulas and their properties as well as some common families of copula are given in Chapter 2. Furthermore, the chapter describes the different methodologies of estimating the parameter of a copula together with the method of selecting the copula which best fits the data. Chapter 3 discusses VaR and ES and the methodology of estimating these two risk metrics using copulas as well as the computation of the VaR bounds of the portfolio. Chapter 4 is devoted to a simulation study. Chapter 5 investigates the estimation of the copula-based Value-at-Risk and expected shortfall of a real data set, more specifically for a portfolio consisting of a South African stock index and a particular stock. In the concluding chapter, Chapter 6, the aims which have been achieved in this mini-dissertation are highlighted, recommendations are made and future work is noted.

Chapter 2

Copulas

In this chapter, the concept and general properties of copulas are introduced. Most of the definitions and theorems have been extracted from Joe (1997), Cherubini et al. (2004) and Nelsen (2006) and are presented fairly generally in an attempt to broadly introduce the ideas. The structure of the chapter is as follows. Firstly, the copula function is defined and its properties are discussed. Then dependence concepts are introduced and some common parametric families of copulas are presented. Lastly some estimation procedures for copulas and the procedure for selecting the copula that provides the best fit for a particular data set is discussed.

2.1 Introduction to copulas

In this section, the mathematical definition of a copula for the bivariate case is presented and its statistical interpretation is then discussed.

2.1.1 Mathematical Definition

A two-dimensional copula (2-copula) is a function $C : I \times I \to I$ where $I = [0, 1]$ with the following properties (Nelsen, 2006, p. 10):

1. It is grounded, that is $C(u, 0) = 0 = C(0, v)$ for every (u, v) in I.
2. It satisfies $C(u, 1) = u$ and $C(1, v) = v$.
3. It is 2-increasing, that is

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

   for every u_1, u_2, v_1, v_2 in I where $u_1 \leq u_2$ and $v_1 \leq v_2$.

2.1.2 Sklar's Theorem

Consider two random variables X and Y having marginal distribution functions $F(x)$ and $G(y)$ respectively. From the definition of the distribution function, it is known that $F(x)$ and $G(y)$ are uniformly distributed, that is $F(x) \sim U(0,1)$, $G(y) \sim U(0,1)$ and thus $F(x)$ and $G(y)$ lie in I. Sklar's Theorem is presented as follows (Nelsen, 2006, p. 18):

1. If $H(x, y)$ is the joint distribution function of X and Y, then there exists a copula C such that
$$H(x, y) = C(F(x), G(y)). \tag{2.1}$$
The copula C is unique if and only if the two marginal distribution functions are continuous.

2. Conversely, if C is a copula and $F(x)$ and $G(y)$ are univariate distribution functions, then the function $C(F(x), G(y))$ is a bivariate distribution function with marginal distribution functions $F(x)$ and $G(y)$.

Thus Sklar's Theorem provides a powerful and direct link between the 2-copula and the bivariate distribution function.

2.1.3 Statistical Interpretation

From Sklar's Theorem, the 2-copula can be defined statistically as a bivariate distribution function. Furthermore, Sklar's Theorem demonstrates that a 2-copula function separates a bivariate joint distribution function into its dependence structure and its univariate marginal distribution functions since it is a function of the latter. That is, the copula function expresses the bivariate distribution function in terms of variables that have been transformed into uniform variates which lie in I, through the marginal distribution functions. So the copula approach provides a flexible way of modelling a bivariate distribution in two stages:

1. Modelling the marginal distributions.
2. Modelling the copula function separately.

With the copula construction in Equation (2.1), different marginal distributions can be selected for each random variable X and Y. This differs from the "traditional" way of constructing a bivariate distribution which generally limits the marginal distributions to be of the same type. In a financial context, as for example when making portfolio decisions, this key characteristic of copulas of being able to select different marginals enables firstly the modelling of the marginal distribution functions of each equity return series in the portfolio and then secondly the modelling of the dependence structure between the different equity returns. These steps can be done independently of each other.

2.2 Properties

In this section two important features of a copula function, namely its invariance property and its bounds, are explored.

2.2.1 Invariance Property

One important characteristic of a copula is its invariance property under strictly monotonic increasing transformations of the underlying random variables (Cherubini et al., 2004, p. 72; Nelsen, 2006, p. 25). Consider the continuous random variables X and Y which are combined by a copula function C. If a_1 and a_2 are two monotonic increasing functions, then the transformed variables $a_1(X)$ and $a_2(Y)$ are linked by C as well since

$$Pr\left(X \leq x, Y \leq y\right) = Pr\left(a_1(X) \leq a_1(x), a_2(Y) \leq a_2(y)\right).$$

Therefore, irrespective of any monotonic transformations of the marginal variables, the copula function captures the dependence structure between the underlying variables.

2.2.2 The Fréchet-Hoeffding Bounds

Any copula C satisfies the inequality

$$\max(u+v-1,0) \leq C(u,v) \leq \min(u,v) \tag{2.2}$$

for all $u, v \in I$, which is known as the Fréchet-Hoeffding inequality (Cherubini et al., 2004, p. 52). This inequality can be derived immediately from the fact that for events A and B

$$\max\{Pr(A) + Pr(B) - 1, 0\} \;\leq\; Pr(A \cap B) \;\leq\; \min\{Pr(A), Pr(B)\}.$$

Thus by taking A and B to be the events that $X \leq x$ and $Y \leq y$ respectively, it follows that:

$$\max\{F(x) + G(y) - 1, 0\} \;\leq\; H(x,y) \;\leq\; \min\{F(x), G(y)\}.$$

In the inequality (2.2), both the Fréchet-Hoeffding lower bound and the upper bound are copulas known as the minimum copula and the maximum copula, denoted by C^- and C^+ respectively. In other words, any copula function C is constrained to lie between the minimum and maximum copulas as follows:

$$C^-(u,v) \leq C(u,v) \leq C^+(u,v). \tag{2.3}$$

2.3 Dependence Concepts

The joint distribution of two random variables fully captures the dependence structure between the variables (Embrechts et al., 2002). It follows from Sklar's Theorem that as copulas separate the marginal behaviour from the dependence structure, they provide an intuitive way of understanding and measuring the association amongst random variables. In addition, due to their invariance property, copulas capture the co-movements of the underlying variables, irrespective of the scale in which each variable is measured and can thus be used as a basis for measures of dependence (Frees and Valdez, 1997).

In this section, the notion of linear correlation between two random variables is explored and its weaknesses are highlighted. The concepts of dependence and independence stated in terms of the copula function are appraised as well. Furthermore, some alternative measures of dependence formulated on the basis of copulas are presented.

2.3.1 Linear Correlation

In practice, Pearson's (linear) correlation coefficient is the most frequently used measure of dependence. The linear correlation coefficient of random variables X and Y having finite variances is defined as:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \tag{2.4}$$

where σ_{XY} is the covariance of X and Y and σ_X^2 and σ_Y^2 are their corresponding variances. Within the context of copulas, Pearson's correlation coefficient can be defined as follows (Cherubini et al., 2004, p. 41):

$$\rho_{XY} = \frac{1}{\sqrt{\sigma_X^2 \sigma_Y^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [C(F(x), G(y)) - F(x)G(y)] \, dx \, dy. \tag{2.5}$$

Pearson's correlation coefficient is a measure of linear dependence and it has the following properties:

1. It is invariant under strictly increasing *linear* transformations. However, ρ_{XY} is influenced by nonlinear transformations as it depends on the marginal distribution functions. For example, if variable X has a standard normal distribution and if $Y = X^2$, using Pearson's correlation coefficient as a measure of dependence yields a very small value even though a perfect relationship exists between the two variables.

2. The measure ρ_{XY} lies between -1 and +1 where $\rho_{XY} = 1$ if and only if there exists a perfect positive linear relationship between X and Y and $\rho_{XY} = -1$

2.3 Dependence Concepts

The joint distribution of two random variables fully captures the dependence structure between the variables (Embrechts et al., 2002). It follows from Sklar's Theorem that as copulas separate the marginal behaviour from the dependence structure, they provide an intuitive way of understanding and measuring the association amongst random variables. In addition, due to their invariance property, copulas capture the co-movements of the underlying variables, irrespective of the scale in which each variable is measured and can thus be used as a basis for measures of dependence (Frees and Valdez, 1997).

In this section, the notion of linear correlation between two random variables is explored and its weaknesses are highlighted. The concepts of dependence and independence stated in terms of the copula function are appraised as well. Furthermore, some alternative measures of dependence formulated on the basis of copulas are presented.

2.3.1 Linear Correlation

In practice, Pearson's (linear) correlation coefficient is the most frequently used measure of dependence. The linear correlation coefficient of random variables X and Y having finite variances is defined as:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \tag{2.4}$$

where σ_{XY} is the covariance of X and Y and σ_X^2 and σ_Y^2 are their corresponding variances. Within the context of copulas, Pearson's correlation coefficient can be defined as follows (Cherubini et al., 2004, p. 41):

$$\rho_{XY} = \frac{1}{\sqrt{\sigma_X^2 \sigma_Y^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [C(F(x), G(y)) - F(x)G(y)]\, dx\, dy. \tag{2.5}$$

Pearson's correlation coefficient is a measure of linear dependence and it has the following properties:

1. It is invariant under strictly increasing *linear* transformations. However, ρ_{XY} is influenced by nonlinear transformations as it depends on the marginal distribution functions. For example, if variable X has a standard normal distribution and if $Y = X^2$, using Pearson's correlation coefficient as a measure of dependence yields a very small value even though a perfect relationship exists between the two variables.

2. The measure ρ_{XY} lies between -1 and +1 where $\rho_{XY} = 1$ if and only if there exists a perfect positive linear relationship between X and Y and $\rho_{XY} = -1$

if and only if there exists a perfect negative linear relationship between X and Y. However, the minimum and maximum values ρ_{XY} can attain are not always -1 and +1 respectively (Cherubini et al., 2004, p. 41). The maximum value that ρ_{XY} can reach, denoted by ρ^+_{XY}, can be calculated by substituting the Fréchet-Hoeffding upper bound into Equation (2.5) to give

$$\rho^+_{XY} = \frac{1}{\sqrt{\sigma^2_X \sigma^2_Y}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [min(F(x), G(y)) - F(x)G(y)]dx\ dy.$$

Similarly, the minimum value ρ_{XY} can attain, denoted by ρ^-_{XY}, can be calculated by substituting the Fréchet-Hoeffding lower bound into Equation (2.5) and is given by

$$\rho^-_{XY} = \frac{1}{\sqrt{\sigma^2_X \sigma^2_Y}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\max(F(x) + G(y) - 1, 0) - F(x)G(y)]dx\ dy.$$

The measures ρ^+_{XY} and ρ^-_{XY} are expected to yield +1 and -1 respectively. However, this is only the case when the underlying variables arise from elliptical distributions such as the normal and the t-distributions (Embrechts et al., 2001) and is not generally so.

3. If X and Y are independent then $\rho_{XY} = 0$ but the converse is not generally true, that is $\rho_{XY} = 0$ does not explicitly imply independence.

4. The correlation coefficient ρ_{XY} is only defined if the random variables have a finite variance (Embrechts et al., 2002). This is problematic if one wants to model variables using a heavy-tailed distribution such as a Cauchy distribution which has an infinite second moment.

2.3.2 Perfect dependence and perfect independence

As an alternative to Pearson's correlation coefficient, measures of dependence based on copulas can be used (Embrechts et al., 2002). In particular, the random variables X and Y are said to be perfect positive dependent, also termed comonotone, if and only if their copula is given by the maximum copula, that is $H(x, y) = C^+(u, v)$ (Cherubini et al., 2004, p. 70).

Similarly, X and Y are said to be perfect negative dependent, also termed countermonotone, if and only if their copula is given by the minimum copula, that is $H(x, y) = C^-(u, v)$.

On the other hand, if X and Y are independent then their copula is given by the *product copula*, $C^\perp$, which is defined as follows:

$$H(x, y) = C^\perp(u, v) = uv. \tag{2.6}$$

Graphs of the maximum, minimum and independent copulas are presented in Figure 2.1. The contour plots of these copulas are also presented in Figure 2.2. A contour plot is a practical way of representing a copula. The contour diagram, defined by Nelsen (2006, p. 12) as "the set in $I \times I$ given by $C(u, v) =$ a constant, for selected constants in I", plots the level curves of the copula $C(u, v)$.

2.3.3 Concordance

Concordance is a broad concept that provides a basis for measures of dependence between random variables. Loosely speaking, X and Y are considered to be concordant if small and large values of X are associated with small and large values of Y respectively and discordant if the converse is true (Nelsen, 2006, p. 157). More particularly, a measure of dependence between X and Y whose copula is given by C is said to be a measure of concordance if it satisfies seven specific properties which are listed in, for example, Cherubini et al. (2004, p. 96). The main characteristics of a concordant measure are as follows:

1. If X and Y are independent, then the concordance measure is 0 but the converse is not generally true.
2. Bounds corresponding to comonotonicity and countermonotonicity exist and equal -1 if X and Y countermonotone and +1 if X and Y are comonotone.
3. The measure is invariant with respect to strictly increasing transformations.

Since Pearson's linear correlation coefficient is not invariant under strictly increasing transformations, it does not qualify as a measure of concordance (Embrechts et al., 2001). Two nonparametric measures, Kendall's tau and Spearman's rho, are commonly used as alternatives to Pearson's linear correlation coefficient and they are both concordant (Embrechts et al., 2001). These measures are nonparametric in the sense that they are not dependent on the marginal density functions of the underlying variables and they can therefore be used as a dependence measure for nonelliptical distributions.

Kendall's tau

Kendall's tau for a population is defined as follows. Consider two independent and identically distributed random vectors (X_1, Y_1) and (X_2, Y_2) having the same copula function given by C. Then Kendall's tau is described by Nelsen (2006, p. 158) as " the probability of concordance minus the probability of discordance" and it is defined by:

$$\tau = Pr\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - Pr\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right]. \tag{2.7}$$

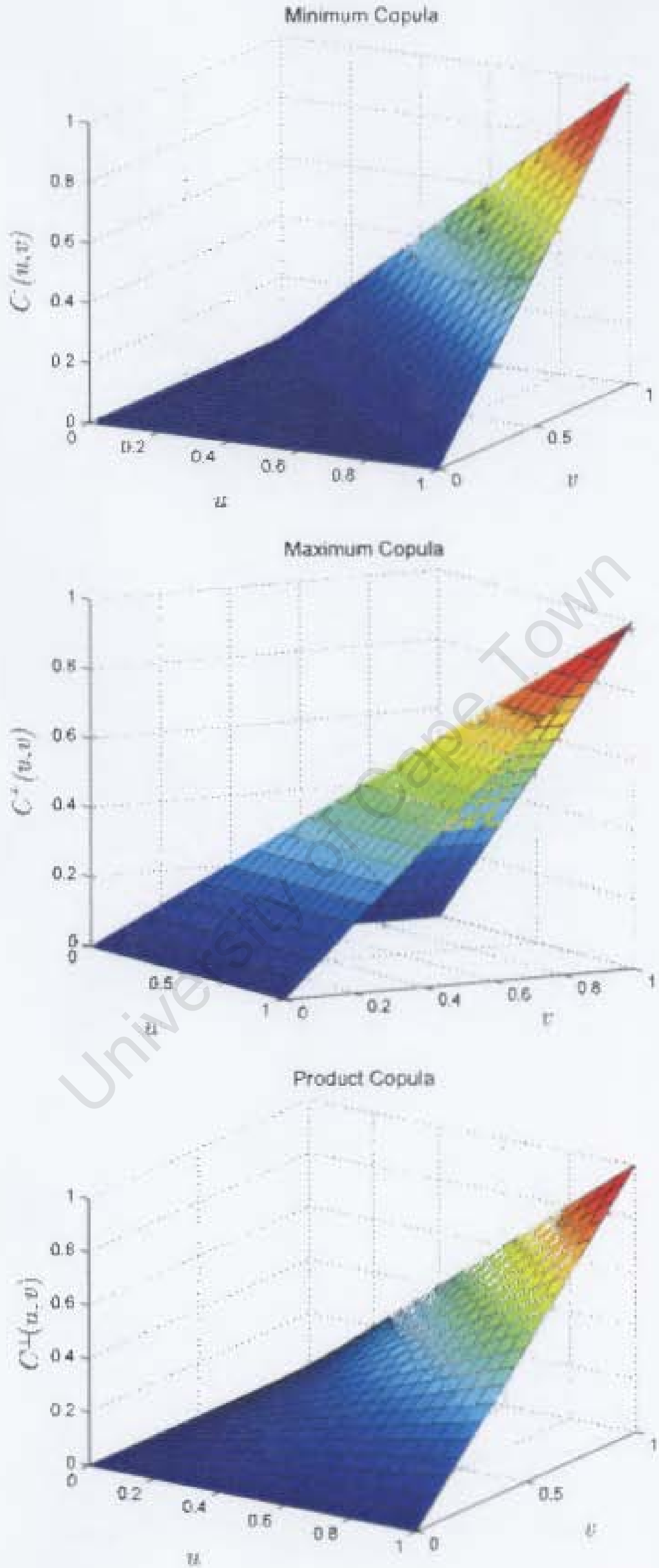

Figure 2.1: The graphs of the minimum, maximum and the product copulas.

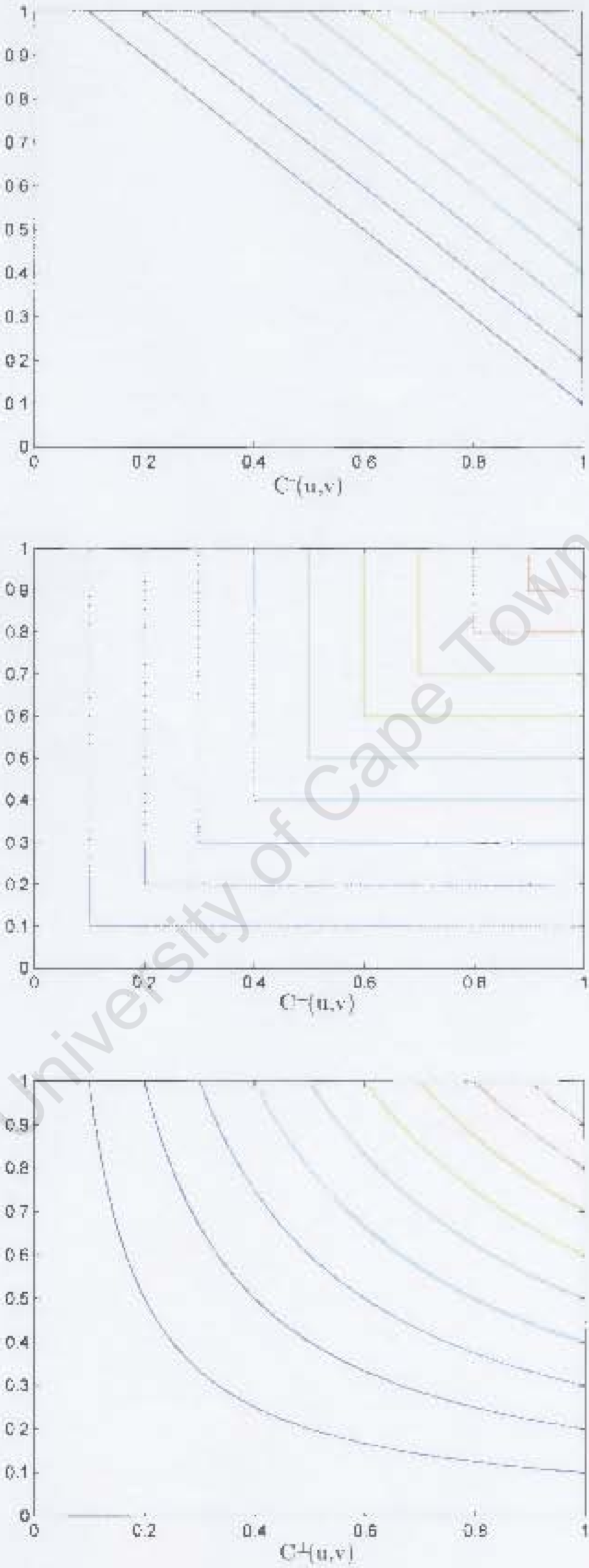

Figure 2.2: The contour plots of the minimum, maximum and the product copulas.

Kendall's tau can also be expressed in terms of the copula function C as (Cherubini et al., 2004, pp. 97-98):

$$\tau = 4\int_I\int_I C(u,v)dC(u,v) - 1. \tag{2.8}$$

The measure τ lies between -1 and +1. In particular, $\tau = -1$ if and only if $C(u,v) = C^-(u,v)$ and $\tau = +1$ if and only if $C = C^+(u,v)$. Furthermore, $\tau = 0$ if and only if $C = C^\perp(u,v)$.

For a random sample $(x_1,y_1),\ldots,(x_n,y_n)$, taken from a bivariate distribution whose copula function is given by C, the sample version of Kendall's tau can be computed following, for example, Frees and Valdez (1997) as

$$\hat{\tau} = \frac{2}{n(n-2)}\sum_{i<j}^{n} \text{sign}\Big[(x_i - x_j)(y_i - y_j)\Big]. \tag{2.9}$$

Spearman's rho

The population version of Spearman's rho is defined as follows (Nelsen, 2006, p. 167). Consider three independent and identically distributed random vectors (X_1,Y_1), (X_2,Y_2) and (X_3,Y_3) taken from the bivariate distribution function $H(x,y)$ whose copula function is given by C. Spearman's rho is described by Nelsen (2006, p. 167) to be "proportional to the probability of concordance minus the probability of discordance" for any pair of vectors "with the same margins, but one vector has distribution function H, while the components of the other are independent". It is defined as

$$\rho_s = 3\Big(Pr\left[(X_1 - X_2)(Y_1 - Y_3) > 0\right] - Pr\left[(X_1 - X_2)(Y_1 - Y_3) < 0\right]\Big). \tag{2.10}$$

Furthermore, Spearman's rho can be expressed in terms of the copula function C as follows (Nelsen, 2006, p. 167):

$$\rho_s = 12\int_I\int_I C(u,v)dv\, du - 3. \tag{2.11}$$

As for Kendall's tau, Spearman's rho covers the range from -1 to +1 where $\rho_s = -1$ if and only if X and Y are countermonotonic, $\rho_s = +1$ if and only if X and Y are comonotonic and lastly $\rho_s = 0$ if and only if X and Y are independent.

Spearman's sample rho can be obtained as follows (Cherubini et al., 2004, p. 101). Consider a random sample $(x_1,y_1),\ldots,(x_n,y_n)$, taken from a bivariate distribution whose copula function is given by C and where the ranks of x_i and y_i in the sample for $i = 1,\ldots,n$ are denoted by r_i and s_i respectively. Spearman's sample rho is then given by the sample correlation coefficient of the ranks, that is

$$\hat{\rho}_s = \frac{\sum_{i=1}^n (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^n (r_i - \bar{r})^2 \sum_{i=1}^n (s_i - \bar{s})^2}} \tag{2.12}$$

where $\bar{r} = \frac{1}{n}\sum_{i=1}^n r_i$ and $\bar{s} = \frac{1}{n}\sum_{i=1}^n s_i$ are the averages of r_i and s_i respectively.

2.3.4 Tail Dependence

Tail dependence relates to concordance between extreme values of the underlying variables, that is values in the upper-right or lower-left quadrant tails of the bivariate distribution. The amount of dependence in the upper-right and lower-left quadrant tails of the joint distribution function is described by the upper tail dependence and lower tail dependence, denoted by λ_U and λ_L respectively. These are defined as follows (Joe, 1997, p. 33; Cherubini et al., 2004, p. 43). Consider the random variables X and Y whose copula is given by C. Then the probability that X takes a large value with a probability greater than ν, given that Y also assumes a large value with a probability greater than ν, can be introduced as a measure of upper tail dependence, specifically;

$$\begin{aligned} \lambda_U(\nu) &= Pr\left(F(x) > \nu \mid G(y) > \nu\right) \\ &= \frac{Pr\left(F(x) > \nu, G(y) > \nu\right)}{Pr(G(y) > \nu)} \\ &= \left[\frac{1 - 2\nu + C(\nu,\nu)}{1-\nu}\right]. \end{aligned}$$

If $\lambda_U(\nu)$ is computed for very large values of ν, then $\lambda_U(\nu)$ corresponds to the limit of the conditional probability that the distribution function of X is greater than ν as ν tends to one, given that the corresponding distribution function for Y exceeds ν, that is

$$\lambda_U(\nu) = \lim_{\nu \to 1^-} \left[\frac{1 - 2\nu + C(\nu,\nu)}{1-\nu}\right]. \tag{2.13}$$

If $\lambda_U(\nu)$ exists, then C has an upper tail dependence if and only if $0 < \lambda_U(\nu) \leq 1$, as there is a positive conditional probability that X assumes a large value given that Y also does so. Also $\lambda_U(\nu) = 0$ implies independence and is obtained by substituting the product copula in Equation (2.13).

Similarly, the probability that X takes a small value with a probability less than ν, given that Y also assumes a small value with a probability less than ν, can be introduced as a measure of lower tail dependence, specifically;

$$\begin{aligned} \lambda_L(\nu) &= Pr\left(F(X) \leq \nu \mid G(Y) \leq \nu\right) \\ &= \frac{Pr\left(F(X) \leq \nu, G(Y) \leq \nu\right)}{Pr(G(Y) \leq \nu)} \\ &= \left[\frac{C(\nu,\nu)}{\nu}\right]. \end{aligned}$$

If $\lambda_L(\nu)$ is computed for very small values of ν, then $\lambda_L(\nu)$ corresponds to the limit of the conditional probability that the distribution function of X is less than ν as ν tends to zero, given that the corresponding distribution function for Y is less than ν, that is

$$\lambda_L(\nu) = \lim_{\nu \to 0^+} \left[\frac{C(\nu,\nu)}{\nu}\right]. \tag{2.14}$$

If $\lambda_L(\nu)$ exists, then C has a lower tail dependence if and only if $0 < \lambda_L(\nu) \leq 1$ while $\lambda_L(\nu) = 0$ implies independence.

2.3.5 Positive Quadrant Dependence

The variables X and Y are said to be positive quadrant dependent (PQD) if and only if the following condition holds (Cherubini et al., 2004, p. 110):

$$C(u, v) \geq C^{\perp}(u, v) = uv. \tag{2.15}$$

In other words, X and Y are said to be PQD if the probability that they assume small or large values together is at least as great as the probability were they independent. Analogously, X and Y are negative quadrant dependent (NQD) if the probability that they assume small or large values together is at least as small as the probability were they independent (Joe, 1997, p. 20).
Furthermore, if X and Y are PQD, then it can be shown that Kendall's tau, Spearman's rho and Pearson's correlation coefficient will all be positive as these three coefficients are zero in the case of independence of the underlying variables (Cherubini et al., 2004, p. 111).

2.4 Multivariate Copulas

The definition of an n-copula as well as Sklar's Theorem and Fréchet-Hoeffding bounds for the n-copula follow immediately from the bivariate setting and details about the n-copula can be found in, for example, Nelsen (2006). Since this mini-dissertation is concerned with the bivariate case, multivariate copulas are not further discussed.

2.5 Common parametric families of bivariate copulas

In this section, some well-known parametric families of bivariate copulas which are popular in financial applications and which are used in this mini-dissertation are now introduced. These are the elliptical copulas, in particular the Gaussian and the Student's t copulas, and the Archimedean copulas, in particular the Frank, Gumbel and Clayton copulas. For each copula, the distribution and density function will be presented as well as the properties.

2.5.1 Elliptical Copulas

Elliptical copulas are the copulas derived from elliptical distributions (Embrechts et al., 2001). The main features of elliptical copulas are that they "share many of the tractable properties of the multivariate normal distribution" (Embrechts et al., 2001, p. 22) and that draws from them can be easily simulated. Furthermore, the computation of Kendall's tau, Spearman's rho and the measures of tail dependence is straightforward. Nevertheless, Embrechts et al. (2001) also point out that elliptical copulas have certain drawbacks, in particular the following:

1. Elliptical copulas do not have closed form expressions.
2. Elliptical copulas cannot be used to model asymmetries in the data as, for example, the well-known asymmetries between market losses and market gains.

The two common elliptical copulas that will be discussed are the Gaussian copula and the Student's t copula.

Gaussian Copula

The bivariate Gaussian copula or the normal copula whose parameter is given by ρ, is defined as follows (Cherubini et al., 2004, p. 112):

$$C_\rho^{Ga}(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$$

where Φ denote the distribution function of a standard normal N(0,1) with its inverse function given by Φ^{-1} and Φ_ρ represent the distribution function of the standard bivariate normal. The distribution function of the Gaussian copula is thus given by:

$$C_\rho^{Ga}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right\} ds\, dt \quad (2.16)$$

where $-1 \leq \rho \leq 1$, $s = \Phi^{-1}(u)$ and $t = \Phi^{-1}(v)$.

The Gaussian copula yields the joint bivariate normal distribution if and only if the univariate marginal distribution functions of X and Y are normally distributed. However, the marginal distributions can take any other and possibly different functional forms. Furthermore, even if the marginal distribution functions are normal, selecting a non-Gaussian copula will yield a nonnormal joint distribution function (Luciano and Marena, 2003).

The density of the Gaussian copula is given by (Joe, 1997, p. 141):

$$\begin{aligned} c_\rho^{Ga}(u,v) &= \left(1-\rho^2\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(1-\rho^2\right)^{-1}\left(s^2+t^2-2\rho st\right)\right\} \\ &\quad \exp\left\{\frac{1}{2}\left(s^2+t^2\right)\right\} \end{aligned} \quad (2.17)$$

The Gaussian copula exhibits the following dependence properties (Cherubini et al., 2004, p. 116). Firstly, it has neither lower nor upper tail dependence except when $\rho = 1$, in which case $\lambda_U = \lambda_L = 1$ and when $\rho = 0$ it is simply the product copula. Furthermore, if $\rho \geq 0$, it exhibits PQD. Lastly, ρ is linked to both Kendall's tau and Spearman's rho as follows:

$$\tau = \frac{2}{\pi} \arcsin \rho \quad \text{and} \quad \rho_s = \frac{6}{\pi} \arcsin \frac{\rho}{2}.$$

Student's t copula

The Student's t copula is similar to the Gaussian copula except that it has a second parameter ν, which accommodates tails fatter than those of the normal distribution and hence its use in finance. If T denote the distribution function of the univariate Student's t distribution with ν degrees of freedom, then the Student's t copula is given by (Cherubini et al., 2004, p. 116):

$$C^T_{\rho,\nu}(u,v) = \int_{-\infty}^{T^{-1}(u)} \int_{-\infty}^{T^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1 + \frac{s^2 - 2\,\rho st + t^2}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2} ds\; dt \quad (2.18)$$

where $-1 \leq \rho \leq 1, \nu > 2, s = T^{-1}(u)$ and $t = T^{-1}(v)$.

The corresponding density of the Student's t copula is given by

$$c^T_{\rho,\nu}(u,v) = \rho^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)^2} \frac{\left(1 + \frac{s^2+t^2-2\rho st}{\nu(1-\rho^2)}\right)^{-(\nu+2)/2}}{\left(1+\frac{s^2}{\nu}\right)^{-(\nu+2)/2}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+2)/2}} \quad (2.19)$$

where Γ denotes the gamma function.

For large values of ν, the Student's t and the Gaussian copulas behave similarly and in fact as $\nu \to \infty$, the Student's t copula converges to the Gaussian copula (Cherubini et al., 2004, p. 116). However, for a finite ν these two copulas behave quite differently. The Student's t copula exhibits equal upper and lower tail dependency (Embrechts et al., 2001) and for finite ν, $\lambda_U = \lambda_L$ increases as ρ increases and $\lambda_U = \lambda_L = 1$ if and only if ρ =1.

The distribution function, the density function and the contour plot of the Gaussian and the Student's t copulas are shown in Figure 2.3 and in Figure 2.4.

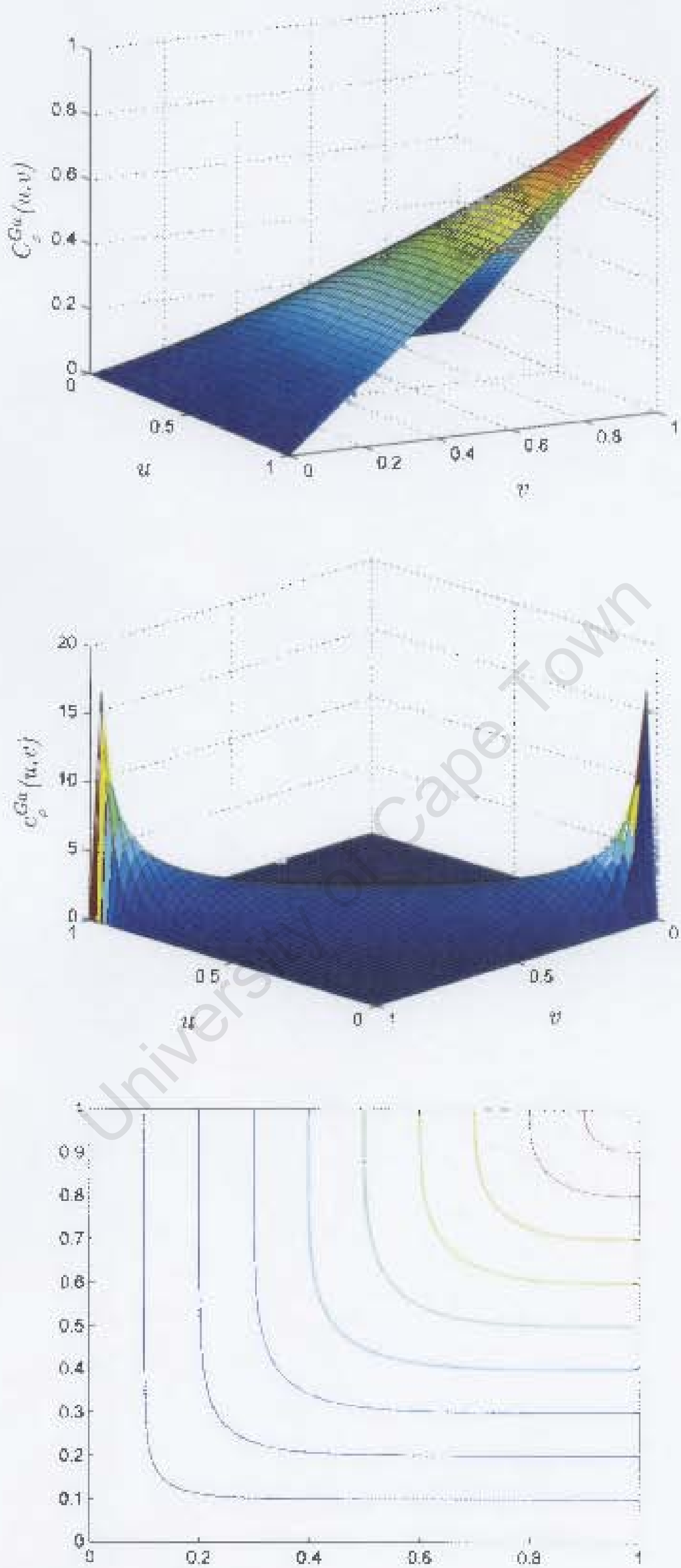

Figure 2.3: The Gaussian copula, its density function and its contour plot for ρ=0.9.

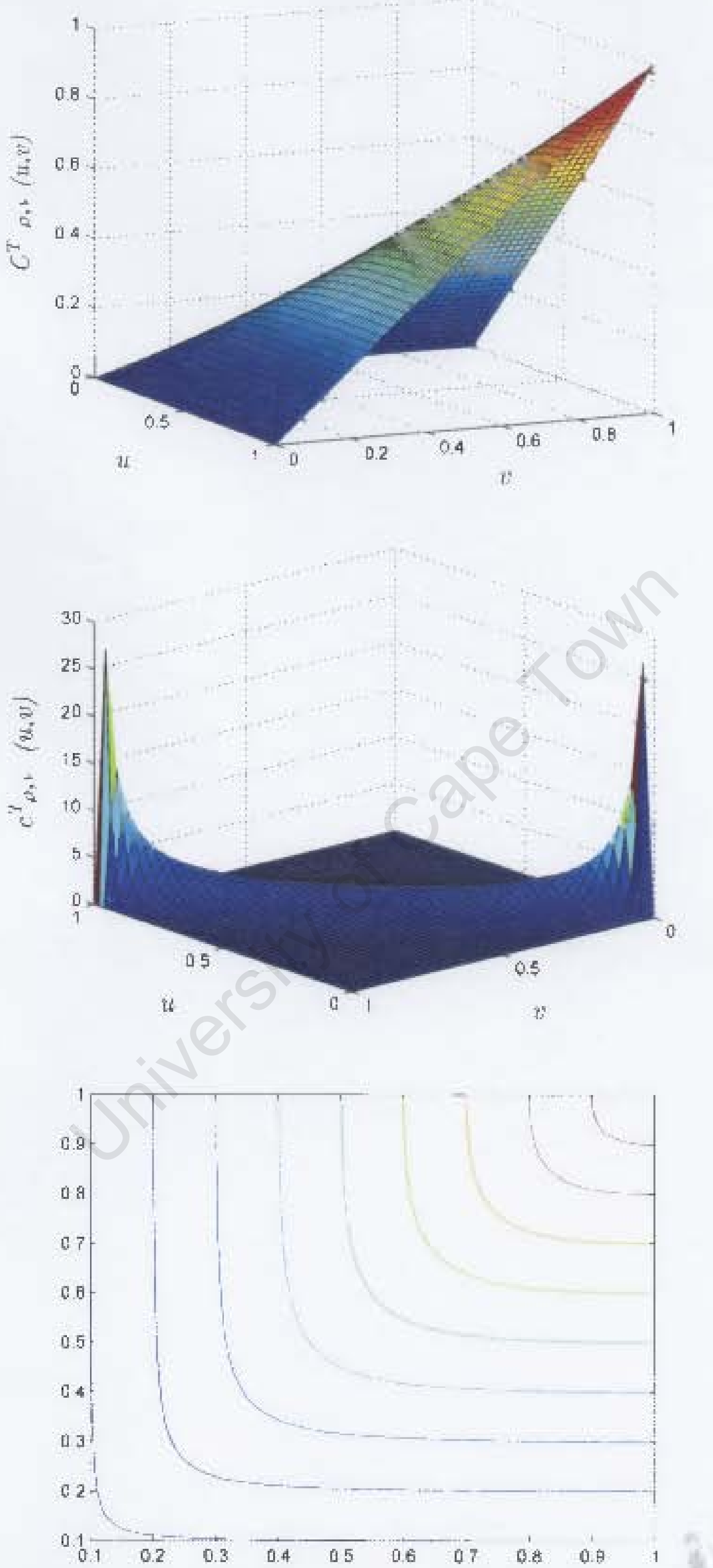

Figure 2.4: The Student's t copula, its density function and its contour plot for $\rho=0.9$, $\nu=3$.

2.5.2 Archimedean Copulas

In a financial context, it is a common occurrence that the level of dependence between big losses for example during a period of crisis, is greater than that between big gains. As stated above, elliptical copulas cannot model such asymmetries. Another class of copulas which does accommodate an asymmetric type of dependence structure is that of the Archimedean copulas (Genest and Rivest, 1993). Furthermore, these copulas have the benefit of having a closed form expression given by:

$$C_\phi(u,v) = \phi^{-1}(\phi(u)+\phi(v)) \tag{2.20}$$

where u and v lie in $(0,1)$, ϕ is a convex, decreasing function with domain (0,1] such that $\phi(1) = 0$, and ϕ^{-1} is the inverse function of ϕ. The function ϕ is known as the generator of the copula C_ϕ. Different generators will yield different families of copula with the three most common members being the Frank, Clayton and Gumbel copulas. The generator functions for these three copulas as well as the range of their parameter, θ are shown in Table 2.1 (Nelsen, 2006, p. 116). The distribution and

Family	Generator $\phi(t)$	Range of θ
Frank	$-\ln\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$	$(-\infty,\infty)\setminus\{0\}$
Clayton	$(t^{-\theta}-1)/\theta$	$[-1,\infty)\setminus\{0\}$
Gumbel	$[-\ln(t)]^\theta$	$[1,\infty)$

Table 2.1: The generator function of the Frank, Clayton and Gumbel copulas.

density functions of the Frank, Clayton and Gumbel copulas are presented below.

Frank Copula

The distribution function of the bivariate Frank copula is given by (Nelsen, 2006, p. 116):

$$C_\theta^F(u,v) = -\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right) \tag{2.21}$$

and the density function is given by (Joe, 1997, p. 141):

$$c_\theta^F(u,v) = \frac{\theta e^{-\theta(u+v)}(1-e^{-\theta})}{[1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})]^2}. \tag{2.22}$$

The Frank copula exhibits the following properties (Embrechts et al., 2001). As θ tends to zero, it converges to the product copula and as θ tends to ∞, it is given by the maximum copula. On the other hand, as θ tends to $-\infty$, it is given by the minimum copula.

Clayton Copula

The distribution function of the bivariate Clayton copula is given by (Nelsen, 2006, p. 116):

$$C_\theta^C(u,v) = \max\left[\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}},0\right] \tag{2.23}$$

and the density function of the Clayton copula is given by:

$$c_\theta^C(u,v) = (1+\theta)(uv)^{-\theta-1}C(u,v;\theta)[u^{-\theta}+v^{-\theta}-1]^{-2}. \tag{2.24}$$

The Clayton copula exhibits the following properties (Embrechts et al., 2001). When $\theta > 0$, it exhibits a low tail dependency and as θ tends to zero, it converges to the product copula. Furthermore, when $\theta = -1$, it is the minimum copula and as θ tends to ∞, it converges to the maximum copula.

Gumbel Copula

The distribution function of the bivariate Gumbel copula is given by (Nelsen, 2006, p. 116):

$$C_\theta^G(u,v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}. \tag{2.25}$$

The density function of the Gumbel copula is quite complicated and can be found in, for example, Joe (1997, p. 142). The Gumbel copula exhibits the following properties (Embrechts et al., 2001). When $\theta = 1$, it is the product copula and as θ tends to ∞, it is given by the maximum copula. Furthermore, it exhibits upper tail dependence.

Kendall's tau and Spearman's rho can be written in terms of the copula parameter θ for some of the three copulas as shown in Table 2.2 where D_i is the Debye function of order i for $i = 1, 2$ (Cherubini et al., 2004, p. 126).

Family	Kendall's τ	Spearman's ρ
Frank	$1+4[D_1(\theta)-1]/\theta$	$1-12[D_2(-\theta)-D_1(-\theta)]/\theta$
Clayton	$\theta/(\theta+2)$	complicated expression
Gumbel	$1-\theta^{-1}$	no closed form

Table 2.2: Kendall's tau and Spearman's rho of the Frank, Clayton and Gumbel copulas.

The distribution function, the density function and the contour plot of the Frank, Clayton and Gumbel copulas with parameter $\theta = 12.03$, $\theta = 4.97$, and $\theta = 3.48$ are shown in Figures 2.5, 2.6 and 2.7 respectively.

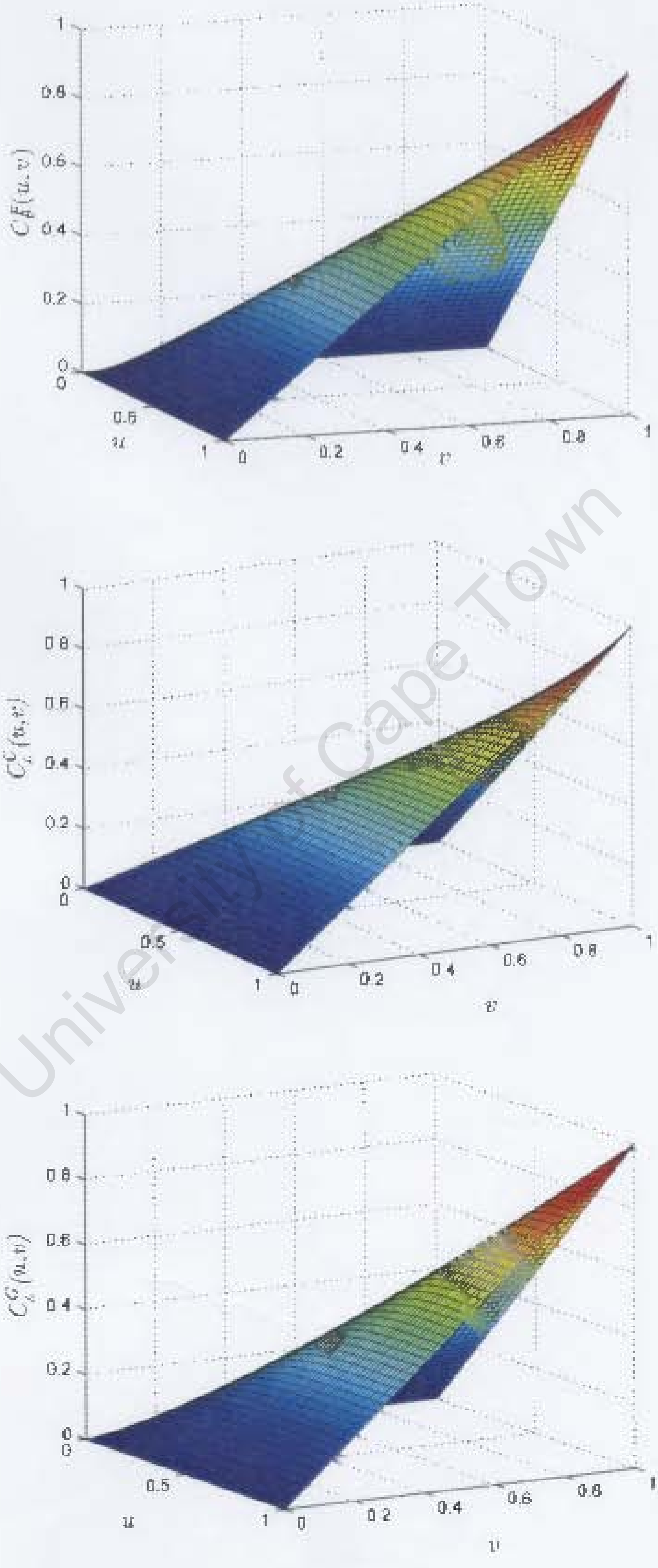

Figure 2.5: The distribution functions of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas.

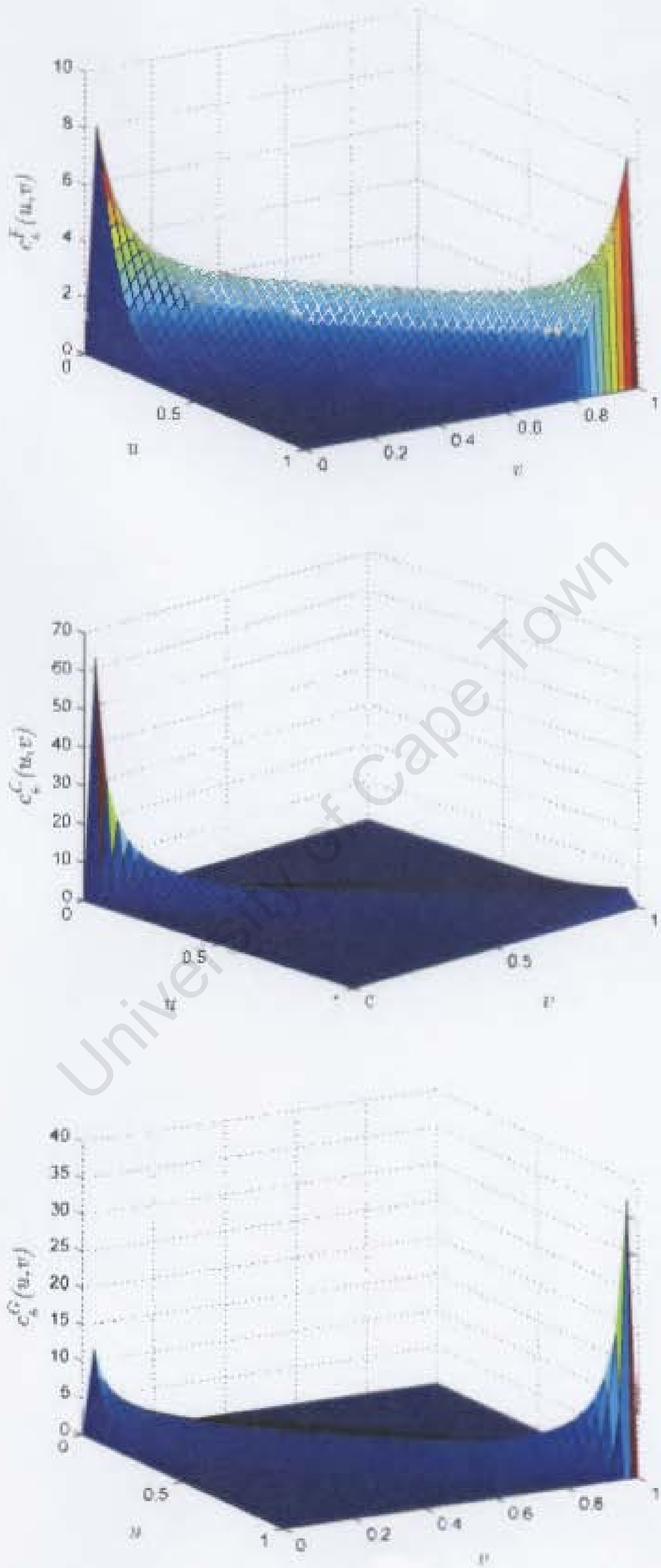

Figure 2.6: The density functions of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas.

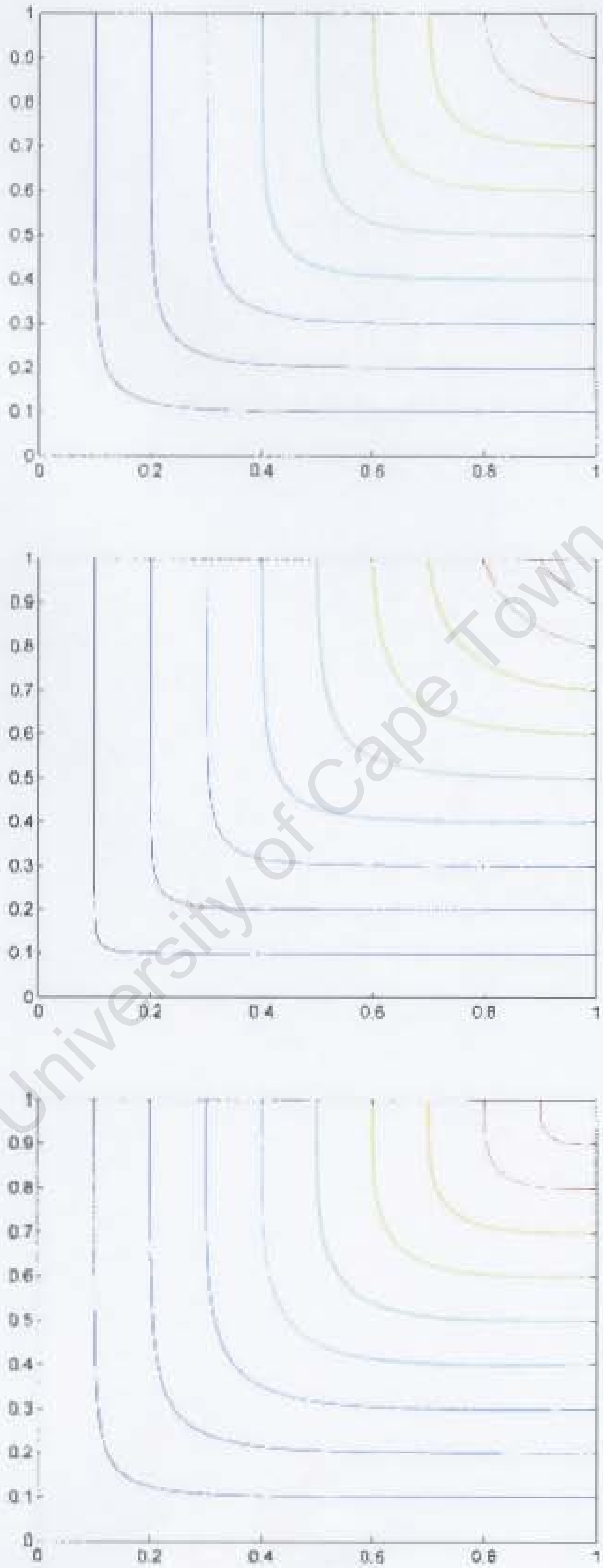

Figure 2.7: The contour plots of the Frank ($\theta = 12.03$), the Clayton ($\theta = 4.97$) and the Gumbel ($\theta = 3.48$) copulas.

The scatter plots of a random sample drawn from two distributions having the same standard normal marginal distribution functions and the same Kendall's tau but different copulas, namely the Gaussian and the Frank copulas are illustrated in Figure 2.8. It is clear that the sample drawn from the Frank copula exhibits more extreme values compared to that from the Gaussian copula.

2.6 Estimation procedures for copulas

In this section, the different techniques of estimating the parameter of the 2-copula function will be presented, namely the parametric, semiparametric and nonparametric techniques. These methods can be extended to the estimation of the n-copula as demonstrated by, for example, Cherubini et al. (2004, pp. 154-161). Details of these estimation methods have been obtained from Joe (1997), Cherubini et al. (2004) and Kim et al. (2007).

Consider the two random variables X and Y. For the parametric and semiparametric estimation methods, the parameters of the marginal distribution function of X denoted by α_x, the parameters of the marginal distribution function of Y denoted by α_y and the parameter of the copula function θ need to be estimated. Let $\Psi = (\alpha_x, \alpha_y, \theta)'$ represent the vector of unknown parameters to be estimated. Furthermore, let the univariate density functions and the joint density function of X and Y be denoted by $f(x; \alpha_x)$, $g(y; \alpha_y)$ and $h(x, y; \Psi)$ respectively. Similarly, let the marginal distribution functions and the joint distribution function of X and Y be denoted by $F(x; \alpha_x)$, $G(y; \alpha_y)$ and $H(x, y; \Psi)$ respectively.

Firstly, two parametric estimation methodologies are discussed namely maximum likelihood estimation (MLE) and inference function for margins (IFM). The semiparametric estimation technique which is presented is called canonical maximum likelihood (CML), also known as semiparametric method, and lastly, nonparametric estimation is presented.

2.6.1 Parametric estimation procedures

These methods are fully parametric in the sense in that it is necessary to specify the model "up to a finite number of unknown parameters" according to Kim et al. (2006, p. 2839). The MLE method is first discussed followed by the IFM method.

Maximum Likelihood Estimation

In this estimation procedure, the parameters are estimated by maximising the log-likelihood with respect to all the parameters simultaneously. It can be recalled that

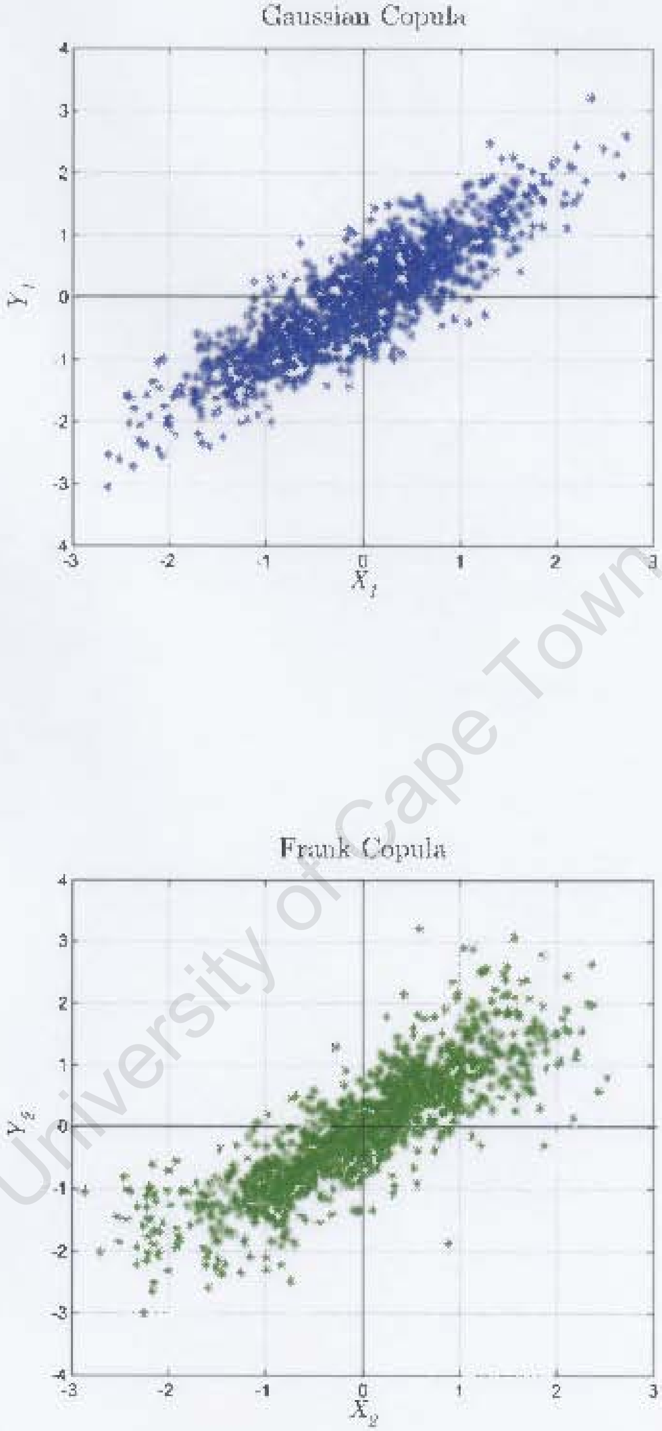

Figure 2.8: A sample of random variables drawn from two distributions having the same standard normal marginal distributions and the same $\tau = 0.7$ but different dependence structures.

$H(x, y; \Psi)$ can be expressed in terms of the copula function as follows:

$$H(x, y; \xi) = C(F(x; \alpha_x), G(y; \alpha_y); \theta)$$

Therefore, $h(x, y; \xi)$ can be expressed as (Cherubini et al., 2004, p. 154):

$$h(x, y; \Psi) = c(F(x; \alpha_x), G(y; \alpha_y); \theta) f(x; \alpha_x) g(y; \alpha_y) \tag{2.26}$$

where

$$c(F(x; \alpha_x), G(y; \alpha_y); \theta) = \frac{\delta^2(C(F(x; \alpha_x), G(y; \alpha_y); \theta)}{\delta F(x; \alpha_x) \delta G(y; \alpha_y)} \tag{2.27}$$

is the density function of the copula.

Consider a random sample of size T from the copula C, $(x_1, y_1), \ldots, (x_T, y_T)$. Hence, the log-likelihood function takes the following form:

$$l(\Psi) = \sum_{t=1}^{T} \log[c(F(x; \alpha_x), G(y; \alpha_y); \theta)) f(x; \alpha_x) g(y; \alpha_y)]. \tag{2.28}$$

Therefore by performing the global maximisation of $l(\Psi)$ with respect to all the parameters Ψ, the maximum likelihood estimates $\hat{\Psi}_{MLE}$ are generated where

$$\hat{\Psi}_{MLE} = \left(\hat{\alpha}_{x(MLE)}, \hat{\alpha}_{y(MLE)}, \hat{\theta}_{MLE} \right)'.$$

Note that $\hat{\Psi}_{MLE}$ follows a normal distribution $\mathrm{N}(\Psi, \mathfrak{I}^{-1})$ asymptotically (Cherubini et al., 2004, p. 154) where $\mathfrak{I}^{-1}$ represents Fisher's information matrix.

Therefore if the model specification is correct, it implies that some asymptotic properties hold for the MLE estimators, which makes MLE the most favored estimation method. However, there are two problems with this method. Firstly, it depends on the correct specification of the distribution functions because if the model has been misspecified, then the MLE estimators no longer enjoy those asymptotic properties. Secondly, and most importantly, the global maximisation of the parameters of the marginal distributions as well as the parameter of the copula function is very computationally intensive and the optimisation routine can frequently run into convergence problems.

Inference Function for Margin Method

The MLE methodology can be modified in such a way that the resulting method is less computationally intensive and more robust (Joe, 1997, p. 297). This method, known as inference function for margins, is discussed in this section.

Similar to the MLE methodology, the IFM is also fully parametric. However the difference lies in the optimisation procedure which is made less intensive by mirroring the notion of separating the marginal distribution functions and the dependence structure. Therefore instead of maximising the log-likelihood of all the parameters simultaneously, the parameters associated with the different marginal distribution functions and the copula function are maximised separately in two stages to yield $\hat{\Psi}_{IFM} = (\hat{\alpha}_{x(IFM)}, \hat{\alpha}_{y(IFM)}, \hat{\theta}_{IFM})'$:

1. In the first stage, α_x and α_y are estimated separately by maximising their log-likelihood functions to generate $\hat{\alpha}_{x(IFM)}$ and $\hat{\alpha}_{y(IFM)}$ as shown below:

$$\hat{\alpha}_{x(IFM)} = \arg\max_{\alpha_x} \sum_{t=1}^{T} \log f(x; \alpha_x) \tag{2.29}$$

$$\hat{\alpha}_{y(IFM)} = \arg\max_{\alpha_y} \sum_{t=1}^{T} \log g(y; \alpha_y). \tag{2.30}$$

2. In the second stage, θ is estimated by substituting $\hat{\alpha}_{x(IFM)}$ and $\hat{\alpha}_{y(IFM)}$ for α_x and α_y respectively in the log-likelihood function (2.28) which is then maximised with respect to θ to obtain $\hat{\theta}_{IFM}$, that is

$$\hat{\theta}_{IFM} = \arg\max_{\theta} \sum_{t=1}^{T} \log c\left(F\left(x; \hat{\alpha}_{x(IFM)}\right), G\left(y; \hat{\alpha}_{y(IFM)}\right); \theta\right). \tag{2.31}$$

The MLE estimators and the IFM estimators are generally not equal (Cherubini et al., 2004, p. 157). Under a set of regularity conditions, similar to those of the MLE estimators, $\sqrt{T}(\hat{\Psi}_{IFM} - \Psi)$ is asymptotically normally distributed as $N(0, \mathcal{G}^{-1})$ (Joe, 1997, p. 301) where $\mathcal{G}$ is the Godambe information matrix. In practice, the computation of the covariance matrix $\mathcal{G}$ is very involved and it is common to use either the jacknife or the bootstrap method to estimate it (Joe, 1997, p. 302).

2.6.2 Semiparametric estimation procedures

The drawback of the MLE and IFM methodologies is that because they are fully parametric, the number of possible combinations of the parametric forms of the marginal distribution functions and the copula virtually has no limits. This leads to two problems; firstly finding the best combination of marginals and copula can be very time consuming and secondly the parametric estimation procedures suffer from model misspecification such that even if only one distribution function has been incorrectly specified, the estimate of θ which is computed from these methods can provide a poor fit (Kim et al., 2007). The semiparametric estimation procedure which is discussed in this section is known as the canonical maximum likelihood method.

Canonical Maximum Likelihood

The canonical maximum likelihood estimation procedure (CML) is semiparametric in the sense that the marginal distribution functions do not have to be specified while the copula function has to be specified. Similar to the IFM methodology, this method also estimates the parameter of the copula in two stages to yield $\hat{\theta}_{CML}$, but differs in that the marginal distribution functions are estimated nonparametrically as follows (Kim et al., 2007):

1. In the first stage, the empirical distribution functions are used to estimate the marginal distributions. Let $\tilde{F}(x_t)$ and $\tilde{G}(y_t)$ denote the empirical distribution functions of X and Y respectively.

2. In the second stage, θ is estimated by plugging $\tilde{F}(x_t)$ and $\tilde{G}(y_t)$ into the log-likelihood function and maximising it to give:

$$\hat{\theta}_{CML} = \arg\max_{\theta} \sum_{t=1}^{T} \log[c\{\tilde{F}(x_t), \tilde{G}(y_t)\}; \theta] \tag{2.32}$$

2.6.3 Nonparametric estimation

So far, it has been assumed that the copula has a specific parametric form and the parameter of the copula has to been estimated. In nonparametric estimation, no such assumption is made and the resulting copula is known as the *empirical copula*. Consider a random sample (x_t, y_t) for $t = 1, \ldots, T$. The x's and the y's are ordered to generate the ordered statistics $x_{(1)}, \ldots, x_{(T)}$ and $y_{(1)}, \ldots, y_{(T)}$ respectively. Then the empirical copula, $\hat{C}$, is given by (Nelsen, 2006, p. 219):

$$\hat{C}\left(\frac{i}{T}, \frac{j}{T}\right) = \frac{\{Number\ of\ pairs\ (x,y)\ in\ the\ sample : x \leq x_{(i)}\ and\ y \leq y_{(j)}\}}{T} \tag{2.33}$$

where $x_{(i)}$ and $y_{(j)}$ represent the i^{th} and j^{th} ordered statistics respectively for $i \geq 1, j \leq T$.

2.6.4 Choosing a copula from the Archimedean family

It is assumed that the marginal distribution functions are correctly fitted either parametrically or nonparametrically such that the focus is on fitting the copula function. Once a copula function has been fitted to a particular data set, the next step is to verify whether the chosen copula appropriately captures the dependency structure of that data set. In practice, a number of copulas should be fitted, their performance evaluated and the copula that provides the best fit selected. Given that the focus of this mini-dissertation will be on fitting the Frank, Gumbel and Clayton copulas because of their desirable tail dependence properties, the method introduced by Genest and Rivest (1993) for selecting the Archimedean copula that provides the best fit for a particular data set is discussed in this section.

It can be recalled that an Archimedean copula is characterised by the generator function ϕ. Thus the problem of model selection reduces to selecting the generator function that best fits the data. Consider the random variable $Z = H(x, y)$ whose distribution function is given by $K(z) = Pr(Z \leq z)$ (Frees and Valdez, 1997). Genest and Rivest (1993) showed that for an Archimedean copula, ϕ and K are linked as follows:

$$K(z) = z - \frac{\phi(z)}{\phi'(z)} \tag{2.34}$$

where ϕ' is the derivative of ϕ.

The identification of ϕ can then be done by following the three steps outlined below (Genest and Rivest, 1993 and Frees and Valdez, 1997):

1. $K(z)$ is estimated nonparametrically to generate $\hat{K}_T(z)$. This is done by firstly determining the values of Z such that

   $$Z_t = \frac{Number\ of\ pairs\ \{(X_j, Y_j) : X_j < X_t, Y_j < Y_t\}}{T-1}, \tag{2.35}$$

   for all $t = 1, \ldots, T$ and $1 \leq j \leq T$.

   Then $\hat{K}_T(z)$ is computed as follows:

   $$\hat{K}_T(z) = \frac{Number\ of\ \{Z_t \leq z\}}{T}, \tag{2.36}$$

   where $z \in (0, 1)$.

2. Estimate $K(z)$ parametrically for a particular Archimedean family using the right hand side of Equation (2.34). The parameter θ associated with the chosen Archimedean copula can be estimated in two ways. It can be generated by computing Kendall's tau defined in Equation (2.9) and then using the appropriate relationship between τ and θ, as given in Table 2.2, to estimate θ. It can also be estimated using the MLE or IFM method described in Section 2.6.1. Once an estimate of θ is obtained, $\hat{K}_\phi(z)$ is evaluated as:

   $$\hat{K}_\phi(z) = z - \frac{\hat{\phi}(z)}{\hat{\phi}'(z)} \tag{2.37}$$

   This procedure can be repeated using various Archimedean copulas, that is with various choices of ϕ. This selection process does not only apply to the

Frank, Clayton and Gumbel copulas but to all Archimedean copulas. Indeed Nelsen (2006, pp. 116-117) provides a wide selection of Archimedean copulas and their generator functions.

3. Comparison of $\hat{K}_T(z)$ and $\hat{K}_\phi(z)$.
To choose the Archimedean copula which best fits the given data set, the performance of each parametric estimate computed in Step 2 is evaluated against the nonparametric estimated computed in Step 1. This can be done in two ways (Frees and Valdez, 1997):

   (i) Graphically by plotting $\hat{K}_T(z)$ and $\hat{K}_\phi(z)$ against z. The copula which yields an estimate $\hat{K}_\phi(z)$ which is closest to $\hat{K}_T(z)$ is selected.

   (ii) By Evaluating the integrated distance, d_ϕ given by:

$$d_\phi = \int_0^1 \left[\hat{K}_\phi(z) - \hat{K}_T(z)\right]^2 dz. \tag{2.38}$$

   The copula having the minimum d_ϕ is selected.

Chapter 3

Value-at-Risk and Expected Shortfall

The aim of this chapter is to introduce two popular measures of risk, namely Value-at-Risk (VaR) and expected shortfall (ES), as well as to address the problem of estimating the VaR and the ES of a portfolio of dependent assets without necessarily making any joint normality assumptions, through the use of copulas. The structure of the chapter is as follows. Firstly, VaR is defined and the most common methods of estimating the VaR of a portfolio are briefly discussed followed by some comments on the limitations of this risk measure. Then the ES is introduced and its computation and limitations are noted. After the two risk measures have been discussed, the methodology of calculating a "copula-based" VaR and ES, when applied to a portfolio of two risky assets, is presented. Lastly, the computation of the VaR bounds of the portfolio is introduced.

3.1 Description of a two-asset portfolio

Consider a portfolio consisting of two risky assets whose log returns are denoted by the random variables X and Y. Suppose historical daily prices of the two assets represented by $p_{1,t}$ and $p_{2,t}$ are available for $t = 1, \ldots, T$ where T is the total number of observations. Then, the daily log returns of the assets denoted by x_t and y_t are computed as follows for $t = 1, \ldots, T$:

$$\begin{aligned} x_t &= ln\left(\frac{p_{1,t}}{p_{1,t-1}}\right) \\ y_t &= ln\left(\frac{p_{2,t}}{p_{2,t-1}}\right). \end{aligned} \tag{3.1}$$

Furthermore, suppose that the portfolio weights associated with X and Y are given by w_1 and w_2 respectively such that $w_1 + w_2 = 1$. Then if Z denotes the random variable representing the portfolio return, Z is given by:

$$Z = w_1 X + w_2 Y. \tag{3.2}$$

Note that Z is a weighted sum of log returns and is referred to as the portfolio return. The observed portfolio return at time t is given by $z_t = w_1 x_t + w_2 y_t$, $t = 1, \ldots, T$.

3.2 Value-at-Risk

VaR has become a standard measure of risk exposure for banks, insurance companies and other institutions since its introduction in the 1990's as an operational measure of risk by the risk management systems developed by J.P.Morgan (1995) (Dowd and Blake, 2006). An introduction to VaR together with a discussion of various estimation techniques associated with it can be found in for example Alexander (2001), Luciano and Marena (2002) and Dowd (2005b) whilst Dowd and Blake (2006) provide a broader survey on the theory and estimation of various risk measures.

3.2.1 Definition of VaR

VaR is a statistical risk metric that measures the maximum loss that an investment or a portfolio of risky assets can incur due to adverse market movements with a confidence level of $100(1-\alpha)\%$, where $0 < \alpha < 1$, over a given period of time. There are two parameters associated with this risk metric:

1. The time horizon, also known as the holding period, denotes the period of time over which VaR is considered and it varies with the objective of the risk assessment. For example, on an active trading desk, the time horizon ranges from a few hours up to one trading day and in the context of portfolio theory, the time horizon is typically one trading day. However, for less liquid assets the time horizon can be much longer than one day, depending on the time it takes to liquidate the assets (Dowd and Blake, 2006).

2. The second parameter, the confidence level of $100(1-\alpha)\%$ at which VaR is estimated, represents the probability that the actual loss will fall below the VaR. It typically takes values around 95% - 98% as set, for example, by banks and securities companies but, if the concern lies with extreme risks which have a low probability of occurring, the confidence level can be set above 99% (Dowd and Blake, 2006).

The VaR estimated at a confidence level of $100(1-\alpha)\%$, referred to as $100(1-\alpha)\%$ VaR, is the $\alpha\%$-quantile of the portfolio return over the holding period, such that there is a probability of α of the actual loss falling below the VaR over that time period. Let $H(z)$ denote the distribution function of the portfolio return. Then the $100(1-\alpha)\%$ VaR, termed VaR(α), is defined as $Pr(Z \leq \text{VaR}(\alpha)) = \alpha$, that is $H(VaR) = \alpha$. The VaR of the portfolio is therefore the unique solution of:

$$\text{VaR}(\alpha) = H^{-1}(\alpha)$$

since $H(.)$ is a monotonic increasing function.

3.2.2 Estimation of VaR

A number of ways of estimating VaR have been proposed in the literature with the most popular ones being the covariance approach, the Monte Carlo approach and historical simulation. These methods are briefly discussed in this section following Alexander (2001, Chapter 9), Luciano and Marena (2002) and Dowd (2005b, Chapters 4 and 8).

The covariance approach

The covariance approach, also known as the normal approach, is a parametric method of estimating VaR. It assumes that the joint distribution of the log returns of the two-asset portfolio is bivariate normal and thus that the log returns of the underlying assets are normally distributed. This approach necessarily implies that the dependence structure of the underlying assets is captured by the linear correlation coefficient. However as mentioned before, the distribution of returns is known to deviate from normality and furthermore the use of linear correlation as a dependence measure as outlined in Chapter 2 has several drawbacks. This approach is thus not entirely satisfactory.

The Monte-Carlo approach

Dowd (2005b, p. 209) states that "the idea behind Monte Carlo methods is to simulate repeatedly from the random processes governing the prices or returns of the financial instruments". The simulations yield portfolio returns from which an empirical profit and loss distribution can then be generated and VaR can be computed (Dowd, 2005b).

The first step of the simulation procedure involves selecting a model, often complex, for the returns of the underlying assets in the portfolio and the extraneous factors affecting the returns. Then, using either past observations or judgemental views, all the parameters of interest are estimated and from these estimates, random returns of each asset are simulated and hence simulated portfolio returns are generated (Dowd, 2005b, p. 209). The $100(1-\alpha)\%$ VaR of the two-asset is estimated as the $\alpha\%$-quantile of the simulated profit and loss distribution. However, this approach suffers from model risk such that if the model has been misspecified, it leads to inaccurate estimates of VaR (Luciano and Marena, 2002). Furthermore, as the approach is based on computer modelling, it is computationally intensive.

The historical simulation approach

In this method, past observations are used to generate an empirical distribution of portfolio return and the VaR is obtained as the α%-quantile of this distribution (Alexander, 2001). Therefore this approach deals with the problems encountered in the previous two methods as it makes no assumption about the parametric form of the marginal and joint distributions or the dependence structure of the assets as well as no model assumptions. There are more sophisticated extensions to this approach, for example, in smoothing the density of the portfolio returns and in dealing with fat-tails and correlation (Luciano and Marena, 2002). However, the historical simulation approach essentially assumes that the future returns have the same distribution as that of the past returns and as a consequence, it relies too much on historical observations (Luciano and Marena, 2002).

3.2.3 Limitations of VaR

Good risk measures are those that satisfy the properties of coherence as introduced by Artzner et al. (2002). One of the key properties of coherence is subadditivity (Dowd, 2005a; Dowd and Blake, 2006). Specifically, given two portfolios A and B, subadditivity requires that the VaR of these combined portfolios is not more than, and it can even be less than, the sum of the VaR of portfolios A and B, that is $\text{VaR}(A+B) \leq \text{VaR}(A) + \text{VaR}(B)$ (Dowd and Blake, 2006). However, subadditivity for VaR only holds in general for elliptical distributions such as the normal and the Student's t distributions and in practice most profit and loss distributions are not elliptically distributed. Therefore, VaR is a flawed risk metric as it fails to capture the fundamental notion that risks should decrease or at the very least stay constant through portfolio diversification (Dowd, 2005a). Further criticisms about the use of VaR as a measure of risk exposure can be found in, for example, Dowd and Blake (2006).

In spite of VaR not being subadditive, it is widely used and strongly advocated by various bodies, for example the Basel Committee on Banking supervision as well as financial institutions, as being a key risk benchmark. Even the final Basel Capital Accord which is being implemented in 2007 concentrated only on VaR as a risk measurement tool (Basel-II, 2005). Therefore it is important to investigate the estimation of VaR in a multivariate setting.

3.3 Expected shortfall

As an alternative to VaR, the most popular coherent measure of risk is the expected shortfall (Dowd, 2005a).

3.3.1 Definition of expected shortfall

The expected shortfall (ES) is the "average of the worst $100(1-\alpha)\%$ losses" (Dowd and Blake, 2006, p. 200). It is computed as the expected value of losses that are less than the VaR at the $100(1-\alpha)\%$ confidence level over a certain holding period. That is, for a continuous profit and loss distribution, the $100(1-\alpha)\%$ ES is given by:

$$\mathrm{ES}(\alpha) = \int_{-\infty}^{VaR(\alpha)} z\ h(z)\ dz \tag{3.3}$$

where $h(z)$ is the density function of the portfolio return.

This implies that the ES will in general be less than the VaR. The ES satisfies the subaddivity property as demonstrated by, for example, Dowd and Blake (2006) and indeed is a coherent measure of risk.

3.3.2 Estimation of expected shortfall

Based on the estimations of VaR described in Section 3.2.2, it is very straightforward to compute the ES. If the covariance approach is used, then the ES can be easily computed numerically using integral (3.3) where $h(z)$ is the density function of the appropriate normal distribution. For the Monte Carlo and the historical simulation approaches, once an estimate of VaR is obtained, the ES is then estimated as the average of losses, either simulated losses in the case of the Monte Carlo method or realised losses for the historical simulation, that are less than the VaR estimate.

3.3.3 Limitations of expected shortfall

Even though the expected shortfall is a more attractive risk measure compared to the VaR since it is coherent, it is not as widely used as the VaR. Furthermore, it is "also rarely, if ever, the 'best' coherent risk measure" (Dowd, 2005b, p. 37). This is because it gives an equal weight to all tail losses that are less than the VaR which implies that the user is risk-neutral. However in practice, any rational user is risk averse which implies that the higher the losses, the higher the weighting placed on them should be.

3.4 Computation of VaR and expected shortfall using copulas

Copulas provide flexibility in the choice of both the marginals and the dependence structure of the returns of the underlying assets in a particular portfolio. Therefore they can be used as a basis for modelling the weighted sum of the returns of the risky assets in the portfolio. In this section, the techniques of using the dependence structure between the underlying risky assets, quantified by the copula function, to estimate VaR and subsequently ES will be discussed.

The pairs of log returns, (x_t, y_t), for $t = 1, \ldots, T$, are assumed to be available and independently and identically distributed from a bivariate distribution whose copula function is given by C. That is, the log returns x_t are assumed to be observations from the same random variable X whose marginal distribution function is given by $F(x)$. Similarly, the returns y_t are assumed to be observations from the same random variable Y whose marginal distribution function is given by $G(y)$.

Therefore for a given copula with a known or estimated parameter and marginals, the problem is to find the distribution function $H(z)$, and more specifically to find the α%-quantile, that is $\text{VaR}(\alpha) = H^{-1}(\alpha)$ and compute the average losses that fall below $\text{VaR}(\alpha)$, that is $\text{ES}(\alpha)$, at the $100(1-\alpha)\%$ confidence level.

The first step to estimating VaR consists of selecting the marginal distributions of the log returns of the underlying assets. The choice of the marginals that best fit the data set of historical log returns can be aided by using standard statistical packages. Examples of marginal distributions which are good representations of the distribution of log returns include the Student's distribution and the logLaplace distribution (Hürlimann, 2004). The next step involves selecting a copula function which adequately captures the joint distribution of the underlying assets. Lastly, the joint distribution of the two assets is fitted to their historical log returns, to obtain estimates of the parameters of interest using the MLE, IFM or CML estimation methods described in Chapter 2.

Once the joint distribution has been estimated, the VaR of the portfolio can be computed. There are two main ways of estimating the VaR; one approach is analytically based and uses integration to estimate VaR and the other approach uses Monte Carlo simulation. The details about the analytically based VaR have been obtained from Luciano and Marena (2003) and Cherubini et al. (2004) whilst those of the Monte Carlo simulation have been obtained from Giacomini and Härdle (2005) and Dowd (2005a). The discussion is restricted to the bivariate case even though the same principles apply to the $n-$dimensional case. The computation of VaR by integration is first discussed and then the Monte Carlo method is presented.

3.4.1 Computation of VaR and ES by integration

This method takes advantage of the fact that the distribution function of the portfolio returns Z denoted by $H(z)$ can be written explicitly in terms of the conditional copula function. Specifically, the distribution function of Z is given by (Cherubini et al., 2004, p. 68):

$$\begin{aligned}
H(z) &= Pr(Z \leq z) = Pr(w_1 X + w_2 Y \leq z) \\
&= \int_{-\infty}^{+\infty} Pr\left(X \leq \frac{1}{w_1} z - \frac{w_2}{w_1} y \mid Y = y\right) g(y) dy \\
&= \int_{-\infty}^{+\infty} C_{1|2}\left(F\left(\frac{1}{w_1} z - \frac{w_2}{w_1} y\right), G(y)\right) g(y) dy \qquad (3.4)
\end{aligned}$$

where $C_{1|2}$ is the conditional copula of X given Y and thus $\text{VaR}(\alpha) = H^{-1}(\alpha)$.

In general, a closed-form solution for integral (3.4) does not exist and the expression has to be approximated using numerical methods. Furthermore, the $100(1-\alpha)\%$ VaR corresponds to the value of z for which $H(z) = \alpha$ so that in most cases in order to compute VaR, the numerical integration must be nested into a root-finding routine, for example, a bisection method. Most copulas have a known conditional distribution function, for example, Joe (1997, pp. 146-147) provides a list of conditional distributions, $C_{1|2}$, of some copulas. If the conditional distribution function of a copula is known, then only one integral has to be evaluated as shown in expression (3.4). Otherwise a double integration, which is more awkward, has to be performed as follows (Luciano and Marena, 2003):

$$H(z) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{\frac{1}{w_1} z + \frac{w_2}{w_1} y} c\left(F\left(\frac{1}{w_1} z + \frac{w_2}{w_1} y\right), G(y)\right) f(x) dx \right\} g(y) dy \qquad (3.5)$$

where c is the density function of copula C. Integral (3.5) follows from integral (3.4) as the conditional copula $C_{1|2}(u|v) = \frac{dC}{dv}(u, v)$ (Joe, 1997, p. 146).

The expected shortfall can be analytically computed as follows:

$$\text{ES}(\alpha) = \int_{-\infty}^{VaR(\alpha)} z \; h(z) \; dz. \qquad (3.6)$$

But the analytical computation of Integral (3.6) is very awkward.

3.4.2 Computation of VaR and expected shortfall by Monte Carlo simulation

A more widely used method of computing the VaR of a portfolio using the copula approach is based on Monte Carlo calculations. Furthermore, once an estimate of VaR is generated using this method, the expected shortfall can be obtained in a straightforward way. The algorithm for computing the $100(1-\alpha)\%$ VaR for a portfolio of two risky assets using the copula-based Monte Carlo approach is described as follows:

1. Simulate the returns of the two assets.

(i) Choose the total number of simulations denoted by N and simulate pairs $(u_i, v_i) \in I \times I$, for $i = 1, \ldots, N$ from the fitted copula. The algorithm for doing so can be found in most standard textbooks and articles, for example in Frees and Valdez (1997) and in Cherubini et al. (2004; Chapter 6). MATLAB Version 7.3 (R2006b) has a built-in routine, `copularnd.m`, which simulates observations from the most popular copulas, namely the Frank, Clayton, Gumbel, Gaussian and Student t copulas.

(ii) Pairs of returns are then generated from the marginals obtained in (a) as (x_i, y_i) where $x_i = F^{-1}(u_i)$ and $y_i = G^{-1}(v_i)$, $i = 1, \ldots, N$.

2. Each pair of simulated returns can be used to compute the simulated portfolio return as $z_i = w_1 x_i + w_2 y_i$ where $w_2 = 1 - w_1$ for $i = 1, \ldots, N$ and thus a simulated draw from the profit and loss distribution of the portfolio is obtained.

3. The VaR is computed as the α%-quantile of this profit and loss distribution. Using the VaR estimate, the $100(1-\alpha)$% ES is then estimated as the average of the simulated portfolio losses which are less than the VaR estimate, that is

$$\mathrm{ES}(\alpha) = \frac{\sum_{i=1}^{N} z_i \,\mathrm{I}\left(z_i \leq \mathrm{VaR}(\alpha)\right)}{\sum_{i=1}^{N} \mathrm{I}\left(z_i \leq \mathrm{VaR}(\alpha)\right)} \tag{3.7}$$

where the indicator function $\mathrm{I}(e) = 1$ if the statement e is true and 0 otherwise.

3.5 Computation of VaR bounds

The methods used to estimate VaR which have been discussed in Section 3.4 assume a dependence structure between the underlying assets in the portfolio. Another approach to estimating the VaR of a portfolio without making any assumptions about the dependence structure of the asset returns can be found in, for example, Luciano and Marena (2002), Embrechts et al. (2003) and Cherubini et al. (2004, pp. 84-87). Specifically, instead of estimating the VaR of a portfolio, the VaR bounds are estimated using the theorem that any copula is bounded to lie within the minimum and maximum copulas. In order to calculate the VaR bounds, only the marginal distributions of the underlying assets in the portfolio have to be specified (Luciano and Marena, 2002). The marginals can assume specific parametric forms or they can be estimated nonparametrically using the empirical distributions.

Consider the return of an equally weighted two-asset portfolio. Let the returns of the two risky assets be given by X and Y such that the portfolio return, denoted by Z, is given by $Z = 0.5X + 0.5Y$. Furthermore let $X_1 = 0.5X$ and $Y_1 = 0.5Y$ with marginals denoted by $F(x_1)$ and $G(y_1)$ respectively, assumed to be known or estimated. The portfolio return is then given by $Z = X_1 + Y_1$. As noted in Chapter 2, any copula function $C(u, v)$ lies within the Fréchet-Hoeffding bounds for all u, v

$\in I$. As a result, the distribution function of the portfolio return is also bounded as follows (Luciano and Marena, 2002; Cherubini et al., 2004, p. 85):

$$H_L(z) \leq H(z) \leq H_U(z) \tag{3.8}$$

where $H_L(.)$ and $H_U(.)$ are the distribution functions defined by:

$$\begin{aligned} H_L(z) &= \max_{x \in \mathbb{R}} \{F(x_1) + G(z - x_1) - 1, 0\} \\ H_U(z) &= \min_{x \in \mathbb{R}} \{F(x_1) + G(z - x_1), 1\}. \end{aligned}$$

Further, the VaR of the portfolio at the $100(1-\alpha)\%$ level of confidence is in turn bounded as follows:

$$\text{VaR}_U(\alpha) \leq \text{VaR}(\alpha) \leq \text{VaR}_L(\alpha) \tag{3.9}$$

where $\text{VaR}_U(\alpha)$ and $\text{VaR}_L(\alpha)$ are the VaRs of the distribution functions H_U and H_L respectively, computed at the $100(1-\alpha)\%$ confidence level. This implies that $\text{VaR}_U(\alpha)$ and $\text{VaR}_L(\alpha)$ can be obtained as the $\alpha\%$-quantiles of $H_U(z)$ and $H_L(z)$ respectively.

The lower VaR bound, $\text{VaR}_U(\alpha)$, is of greater interest compared to the upper bound, $\text{VaR}_L(\alpha)$, from a risk exposure point of view as it represents "a 'worst-case scenario' at a given confidence level" (Luciano and Marena, 2002, p. 632). However, the VaR bounds may also be regarded as a benchmark interval within which all copula-based intervals for a VaR estimate must fall so both the lower and the upper bounds are useful.

In order to calculate the quantiles of $H_U(z)$ and $H_L(z)$, their inverses need to be computed. In some cases this can be achieved analytically as shown in an example in Cherubini et al. (2004, p. 85). However for the general case, Williamson and Downs (1990) provide an algorithm to compute the inverses of $H_U(z)$ and $H_L(z)$ for a particular value of α, $0 \leq \alpha \leq 1$, which is given by:

$$\begin{aligned} H_L^{-1}(\alpha) &= \min_{t \in [\alpha, 1]} \left[F_{x_1}^{-1}(t) + G_{y_1}^{-1}(\alpha - t + 1)\right] \\ H_U^{-1}(\alpha) &= \max_{t \in [0, \alpha]} \left[F_{x_1}^{-1}(t) + G_{y_1}^{-1}(\alpha - t)\right] \end{aligned}$$

where $F_{x_1}^{-1}(.)$ and $G_{y_1}^{-1}(.)$ are the inverse distribution functions of X_1 and Y_1 respectively. Furthermore, Williamson and Downs (1990) propose a numerical and efficient way of evaluating the above inverses to any specified degree of accuracy using a discrete maximisation procedure as shown below:

$$H_L^{-1}\left(\frac{i+1}{N}\right) = \min_{j \in [i, N-1]} \left[F_{x_1}^{-1}\left(\frac{j+1}{N}\right) + G_{y_1}^{-1}\left(\frac{i-j+1}{N}\right)\right] \tag{3.10}$$

$$H_U^{-1}\left(\frac{i}{N}\right) = \max_{j \in [0, i]} \left[F_{x_1}^{-1}\left(\frac{j}{N}\right) + G_{y_1}^{-1}\left(\frac{i-j}{N}\right)\right] \tag{3.11}$$

for $0 \leq \frac{i+1}{N}, \frac{i}{N} \leq 1$ where $0 \leq i \leq N$ and i and N can only take integer values.

For example, consider the computation of the VaR bounds at the 95% confidence level. To compute $H_L^{-1}(0.05)$, i and N in expression (3.10) can be set to 4 and 100 respectively, these being the minimum values i and N can take. Similarly, to compute $H_U^{-1}(0.05)$, i and N in expression (3.11) can be set to 5 and 100 respectively. The accuracy of these inverse functions can be increased by increasing N and setting $(i+1) = N\alpha$ and $i = N\alpha$ for H_L^{-1} and H_U^{-1} respectively such that $(i+1)$ and i are integers. For example, a more accurate evaluation of $H_L^{-1}(0.05)$ is obtained by setting i and N to 4999 and 10^5 respectively and to evaluate $H_U^{-1}(0.05)$, i and N are set to 5000 and 10^5 respectively.

According to Luciano and Marena (2002), the calculation of VaR bounds as described above is not exposed to model risk since no assumptions are made about the joint distribution of the assets. In addition, through the work of Williamson and Downs (1990), the methodology is not computationally intensive.

Note that once the VaR bounds have been computed, there is no clear way of estimating the corresponding ES bounds other than using simulation or integration which requires the specification of the copula.

Chapter 4

The simulation study

The aim of this mini-dissertation is to investigate a copula-based Monte Carlo approach to estimating the Value-at-Risk (VaR) and the expected shortfall (ES) of a portfolio of two risky assets using three different estimation techniques. As outlined in Chapter 2, the three most common methods of estimating the copula parameter are maximum likelihood estimation (MLE), inference function for margins (IFM) and canonical maximum likelihood (CML). Each method will yield a different estimate of the parameter of the copula and different estimates of the parameters of the marginals which will in turn yield different estimates of the VaR and the ES risk measures. Therefore, for a particular portfolio, there will be in total three estimates of its VaR and three estimates of its ES. To compare, make inferences and evaluate these estimates, a simulation study has been performed using MATLAB Version 7.3 (R2006b). The simulation study is described in detail in this chapter.

4.1 The objectives of the simulation process

In the simulation study, two copulas have been considered, namely the Frank and the Gaussian copulas with normal and Student's t distributed marginals. The three main objectives of the simulation study are:

1. To investigate the properties of the VaR and the ES estimates obtained from the three estimation methods of the parameter of the copula.
2. To evaluate the performance of the three VaR and the three ES estimates.
3. To compare the VaR and the ES estimates obtained from the Frank and Gaussian copulas with normal and Student's t distributed marginals.

To investigate the properties of the VaR and ES estimates, a single simulated data set has been used and a parametric bootstrap has been performed. On the other hand, the performance of the estimates has been evaluated by simulating repeatedly and examining the bias and the mean squared error (MSE) of the VaR and ES estimates. The third objective has been achieved by repeating the simulation study with a different copula and different marginal distributions.

4.2 Scenario

The portfolio

The portfolio has been assumed to consist of two equally weighted risky assets. The return of the portfolio denoted by the random variable Z is therefore given by $Z = 0.5X + 0.5Y$ where X and Y are the log returns of the two risky assets. The aim of the study was to compute the VaR and the ES of this portfolio from simulated data at the 95% confidence level using the three different estimation techniques of the copula parameter, namely MLE, IFM and CML.

The distribution of the assets

1. As the focus of the mini-dissertation was on the modelling of the dependence structure rather than on the univariate distributions of the risky assets, the marginals of the latter have been assumed to be of the same type, either standard normal or Student's t distributed. The Student's t distribution was taken to have either 10 degrees of freedom or, to represent a more pronounced fat tail, 5 degrees of freedom.

2. Two copulas have been considered one of which is the Gaussian copula and the other an Archimedean copula, namely the Frank copula. Furthermore, the parameters of the two copulas have been assigned values which correspond to a low dependence, with Kendall's tau equal to 0.2139 and a high dependence with Kendall's tau equal to 0.6658. For the Frank copula, this is equivalent to the dependence parameter θ being assigned the values 2 and 10 respectively. For the Gaussian copula, this corresponds to the dependence parameter ρ being assigned the values 0.3297 and 0.8653 respectively.

3. To mimic reality, all the variables of interest, namely the parameter of the copula and the parameters of the marginal distribution functions of the log returns of the two risky assets in the portfolio, have been assumed to be unknown. Specifically, let α_x and α_y represent the parameter(s) of the marginal distributions of X and Y respectively. Therefore for the normally distributed marginals, the parameters $\alpha_x = (\mu_x, \sigma_x^2)$ and $\alpha_y = (\mu_y, \sigma_y^2)$ were assumed to be unknown where μ_x, μ_y are the means and σ_x^2, σ_y^2 are the variances of X and Y respectively. For the Student's t distributed marginals, the parameters $\alpha_x = \nu_x$ and $\alpha_y = \nu_y$ were assumed to be unknown where ν_x, ν_y are the degrees

of freedom of X and Y respectively. Furthermore, the copula parameters θ and ρ were also assumed to be unknown.

Twelve scenarios in terms of copulas and marginals thus arose and these scenarios are numbered $1, 2, \ldots, 12$ as shown in Table 4.1.

	Frank Copula		Gaussian Copula	
	$\theta = 2$	$\theta = 10$	$\rho = 0.3297$	$\rho = 0.8653$
X~N(0,1) Y~N(0,1)	1	2	3	4
X~t_5 Y~t_5	5	6	7	8
X~t_{10} X~t_{10}	9	10	11	12

Table 4.1: The twelve scenarios that were used in the simulation study.

4.3 Procedures

Before describing the full simulation process, the basic procedures that were used in that process are outlined in this section. These are the procedures for simulation from a given copula, estimation of the parameters of the copula and marginals and computation of the VaR and ES.

Procedure 1: Simulation

Suppose that the marginal distribution functions of X and Y, denoted by $F(x)$ and $G(y)$ respectively with corresponding parameters α_x and α_y, as well as the copula function, denoted by C with parameter θ, are given. Then the simulation procedure involved the following steps:

1. Random pairs (u_i, v_i), $i = 1, \ldots, N$, where N is the total number of observations, were simulated from the copula C. This was done in MATLAB using the `copularnd.m` routine which is available in the Statistics Toolbox.

2. The pairs of log returns (x_i, y_i) were then computed as $x_i = F^{-1}(u_i)$ and $y_i = G^{-1}(v_i)$, $i = 1, \ldots, N$, where F^{-1} and G^{-1} are the inverses of the distribution functions of X and Y respectively. MATLAB has built-in routines which evaluate the inverse of a wide selection of distribution functions including those of the normal distribution with the routine `norminv.m` and the Student's t distribution with the routine `tinv.m`. If the pairs (u_i, v_i) were assumed to arise from empirical distributions, the corresponding log returns (x_i, y_i) were

obtained directly by computing the inverse of their empirical distributions, that is $x_i = \tilde{F}^{-1}(u_i)$ and $y_i = \tilde{G}^{-1}(v_i)$ where $\tilde{F}^{-1}$ and $\tilde{G}^{-1}$ are the inverses of the empirical distributions of X and Y respectively.

Procedure 2: Estimation

Suppose that the copula function C, whose density function is given by c, and the marginals, specified by the distribution functions $F(x)$ and $G(y)$ with parameters α_x and α_y respectively, have to be fitted to a given data set which comprises of "real" pairs of log returns (x_i, y_i) for $i = 1, \dots, N$. Furthermore, let the density functions of the log returns of X and Y be denoted by $f(x)$ and $g(y)$ respectively. Then the three estimation methods, namely MLE, IFM and CML were used and the procedures are as follows:

1. The MLE estimates were obtained by a global maximisation of the log-likelihood function given by $\sum_{i=1}^{N} \log[c\left(F(x_i;\alpha_x), G(y_i;\alpha_y);\theta\right)) f(x_i;\alpha_x) g(y_i;\alpha_y)]$ to yield $\hat{\psi}_{MLE} = \left(\hat{\alpha}_{x(MLE)}, \hat{\alpha}_{y(MLE)}, \hat{\theta}_{MLE}\right)'$. The density function of the copula has been evaluated in MATLAB using the `copulapdf.m` routine. MATLAB also has built-in routines to evaluate the density and distribution functions of the normal and the Student's t distributions given by `normpdf.m`, and `normcdf.m` and by `tpdf.m` and `tcdf.m` respectively.

2. The IFM estimates of the parameters, $\hat{\psi}_{IFM} = \left(\hat{\alpha}_{x(IFM)}, \hat{\alpha}_{y(IFM)}, \hat{\theta}_{IFM}\right)'$, have been obtained by firstly separately fitting the marginals to the data x_i and to the data y_i for $i = 1, \dots, N$, and specifically by finding the maximum likelihood estimates of their parameters as $\hat{\alpha}_{x(IFM)}$ and $\hat{\alpha}_{y(IFM)}$ respectively. These estimates were then plugged into the log-likelihood function which is given by $\sum_{i=1}^{N} \log c\left(F\left(x_i;\hat{\alpha}_{x(IFM)}\right), G\left(y_i;\hat{\alpha}_{y(IFM)}\right);\theta\right)$, and the latter was optimised with respect to θ to yield the IFM estimate of the copula parameter.

3. The CML estimate of the parameter of the copula, $\hat{\theta}_{CML}$, was obtained by firstly computing the empirical distributions of the data x_i and of the data y_i, $i = 1, \dots, N$, as $\tilde{F}(x)$ and $\tilde{G}(y)$ respectively and then plugging these into the log-likelihood function given by $\sum_{i=1}^{N} \log[c\{\tilde{F}(x_i), \tilde{G}(y_i)\};\theta]$ and maximising the latter with respect to θ.

The log-likelihood functions were optimised by using the MATLAB minimisation routines, `fmincon.m` or `fminunc.m`, such that the negative of the log-likelihood function was minimised.

Procedure 3: Computation of the VaR and ES

The VaR and ES of the two-asset portfolio at the 95% confidence level were computed by using the copula-based Monte Carlo approach discussed in Section 3.4.2. The parameters of the marginals and the copula function were assumed to be given and the procedure used was as follows:

1. Simulate d pairs (x_i, y_i), $i = 1, \ldots, d$, from the joint distribution function of X and Y, as described in the simulation procedure, that is Procedure 1.

2. Calculate the portfolio log returns as:

$$z_i = 0.5x_i + 0.5y_i \text{ for } i = 1, 2, \ldots, d.$$

3. Compute the VaR and ES at the 95% confidence level as follows:

   (i) Compute the 95% VaR as the 5%-quantile of the profit and loss distribution of z_i using the `quantile.m` routine in MATLAB.

   (ii) Compute the 95% ES as follows:

$$\text{ES} = \frac{\sum_{i=1}^{d} z_i \, \text{I}\left(z_i \leq \text{VaR}\right)}{\sum_{i=1}^{d} \text{I}\left(z_i \leq \text{VaR}\right)}$$

   where VaR is that calculated in (a) and the indicator function $I(e) = 1$ if the statement e is true and 0 otherwise.

The algorithm for the computation of the VaR and ES is summarised as a flowchart in Figure 4.1.

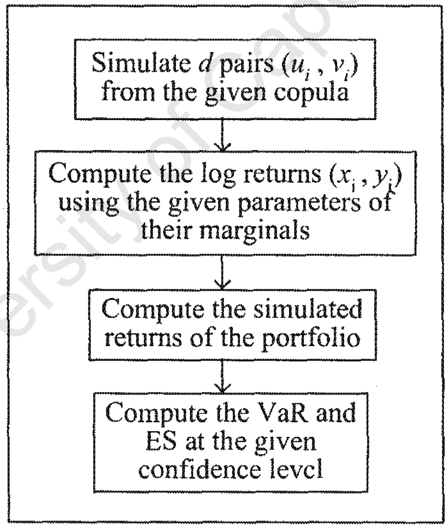

Figure 4.1: Flowchart of the computation of the VaR and ES of the two-asset portfolio.

4.4 Properties of the estimators

The simulation process is illustrated with Scenario 1, for which the marginal distributions of the log returns of the two risky assets were assumed to be standard normal and the dependence structure was modelled through the Frank copula with $\theta = 2$. For the other scenarios, as given in Table 4.1, the same algorithms were used except that the copula and the marginal distributions varied accordingly. These algorithms are presented in three stages, which were the building blocks for all the twelve scenarios.

4.4.1 Stage 1: Generate a "real" data set

Scenario 1 was assumed to be true and the "real" data set as described by Procedure 1 was generated using the true parameters of the copula and the marginals. That is, $N = 1000$ random pairs (u_i, v_i), $i = 1, \ldots, N$, were simulated from the Frank copula with $\theta = 2$. The pairs of log returns (x_i, y_i) were then computed where $x_i = \Phi^{-1}(u_i)$ and $y_i = \Phi^{-1}(v_i)$, $i = 1, \ldots, N$. Once the real data set has been generated, it was assumed that the true parameters are unknown and they have to be estimated.

4.4.2 Stage 2: Obtaining the VaR and ES estimates

The VaR and ES estimates of the "real" data set were obtained as described in the following procedure:

1. The first step was to estimate the parameters of interest of the "real" data set using the three estimation methods as outlined in Procedure 2. Let $\hat{\theta}_{MLE}$ denote the MLE estimate of the parameter of the copula and let the MLE estimates of the mean and standard deviation of X and Y be denoted by $\hat{\mu}_{x(MLE)}$ and $\hat{\sigma}_{x(MLE)}$ and by $\hat{\mu}_{y(MLE)}$ and $\hat{\sigma}_{y(MLE)}$ respectively. Similarly, let $\hat{\theta}_{IFM}$ denote the IFM estimate of the parameter of the copula and let the IFM estimates of the mean and standard deviation of X and Y be denoted by $\hat{\mu}_{x(IFM)}$ and $\hat{\sigma}_{x(IFM)}$ and by $\hat{\mu}_{y(IFM)}$ and $\hat{\sigma}_{y(IFM)}$ respectively. Lastly, let $\hat{\theta}_{CML}$ denote the CML estimate of the parameter of the copula.

2. Using Procedure 3, each estimate of the copula parameter and the parameters of the marginals were then used to estimate the VaR and the ES of the two-asset portfolio at the 95% level as follows:

   (i) Procedure 1 was used to generate simulated data sets using the parameter estimates of the copula and the marginals. Therefore, $\hat{\theta}_{MLE}$, $\hat{\theta}_{IFM}$ and $\hat{\theta}_{CML}$ were used to separately simulate $d = 1000$ pairs (u_i, v_i), $i = 1, \ldots, d$, from the Frank copula such that three different data sets of (u_i, v_i) were generated. Then for the data sets generated using the parameter estimates $\hat{\theta}_{MLE}$ and $\hat{\theta}_{IFM}$, the pairs of log returns, (x_i, y_i), $i = 1, \ldots, d$, were correspondingly obtained as follows:

$$\begin{aligned} x_i &= \hat{\mu}_{x(MLE)} + \hat{\sigma}_{x(MLE)}\Phi^{-1}(u_i) \text{ and } y_i = \hat{\mu}_{y(MLE)} + \hat{\sigma}_{y(MLE)}\Phi^{-1}(v_i) \\ x_i &= \hat{\mu}_{x(IFM)} + \hat{\sigma}_{x(IFM)}\Phi^{-1}(u_i) \text{ and } y_i = \hat{\mu}_{y(IFM)} + \hat{\sigma}_{y(IFM)}\Phi^{-1}(v_i). \end{aligned}$$

On the other hand, the pairs (u_i, v_i) generated using the copula parameter estimate $\hat{\theta}_{CML}$, were assumed to arise from empirical distributions such that the corresponding log returns (x_i, y_i) were obtained as $x_i = \tilde{F}^{-1}(u_i)$ and $y_i = \tilde{G}^{-1}(v_i)$, for $i = 1, 2, \ldots, d$.

(ii) Then using the three different data sets of (x_i, y_i) obtained from the MLE, IFM and CML parameter estimates, the three estimates of VaR denoted by $\hat{V}_{MLE}$, $\hat{V}_{IFM}$ and $\hat{V}_{CML}$ respectively and the three estimates of ES denoted by $\hat{E}_{MLE}$, $\hat{E}_{IFM}$ and $\hat{E}_{CML}$ respectively were computed at the 95% level as described in Procedure 3.

The algorithm for performing Stage 1 and Stage 2 is summarised as a flowchart in Figure 4.2.

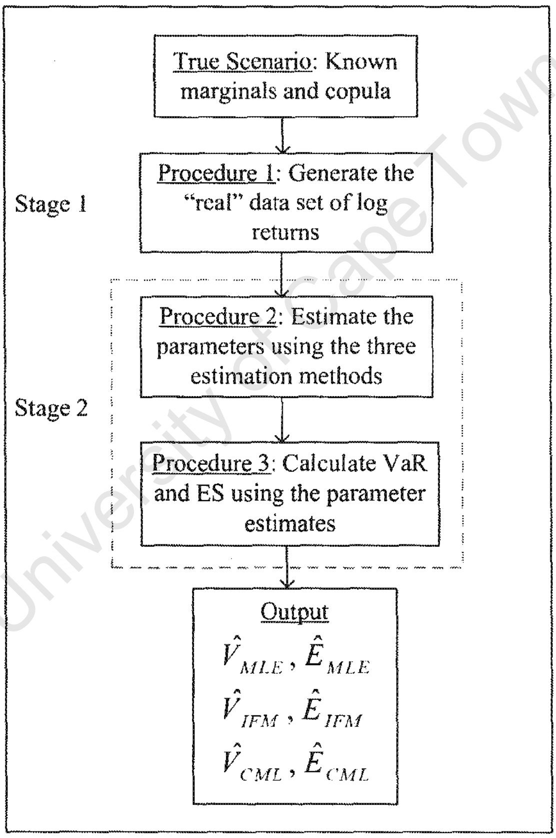

Figure 4.2: Flowchart of the algorithm used to perform Stage 1 and Stage 2.

4.4.3 Stage 3: The bootstrapping process

Once the three estimates of the VaR and the three estimates of the ES risk measures were obtained, their properties, for example the shape of their distribution, bias and confidence intervals, were investigated using the bootstrapping technique.

Let $\hat{V}_{\cdot}^{(b)}$ and $\hat{E}_{\cdot}^{(b)}$, where $\cdot$ represents MLE, IFM, CML, be the b^{th} bootstrap VaR and ES estimates, for $b = 1, \ldots, B$ where B is the number of bootstraps. Given the "real" data set, $\hat{V}_{\cdot}^{(b)}$ and $\hat{E}_{\cdot}^{(b)}$ were obtained as follows:

1. A bootstrap sample (x_i^*, y_i^*), $i = 1, \ldots, N$, with replacement was taken from the "real" data set obtained in Stage 1.
2. The estimates of the parameters of the marginals and the copula were calculated using Procedure 2. The MLE estimates of the marginals of X and Y and the copula are denoted by $\hat{\alpha}_{x(MLE)}^{(b)}$, $\hat{\alpha}_{y(MLE)}^{(b)}$ and $\hat{\theta}_{MLE}^{(b)}$ respectively. Similarly, the IFM estimates of the marginals of X and Y and the copula are denoted by $\hat{\alpha}_{x(IFM)}^{(b)}$, $\hat{\alpha}_{y(IFM)}^{(b)}$ and $\hat{\theta}_{IFM}^{(b)}$ respectively. Lastly, let the CML estimate of the copula be denoted by $\hat{\theta}_{CML}^{(b)}$.
3. Using Procedure 3, $\hat{V}_{\cdot}^{(b)}$ and $\hat{E}_{\cdot}^{(b)}$ were calculated with the estimates obtained from step 2.

The bootstrap samples of the VaR and ES estimates were obtained by repeating the above procedure B times. MATLAB has a built-in bootstrap routine, `bootstrp.m`, which returns a sample of bootstrapped pairs (x_i^*, y_i^*), $i = 1, \ldots, N$, from the specified data set and uses this sample to perform calculations specified by a user-defined function. In this simulation study, the number of bootstrap samples B has been set to 500 given that the algorithm was very computationally intensive and the size of each bootstrap sample N has been set to 1000. The algorithm for obtaining the bootstrap samples of the VaR and ES estimates is summarised as a flowchart in Figure 4.3.

The bootstrap samples were then used to investigate the properties of the VaR and ES estimates as follows:

1. The form of the distribution of the estimates.

   (i) The mean and standard error of the bootstrap sample of the VaR estimates denoted by $\bar{V}_{B(\cdot)}$ and $s_B\left(\hat{V}_{\cdot}\right)$ and those of the ES estimates denoted by $\bar{E}_{B(\cdot)}$ and $s_B(\hat{E}_{\cdot})$ respectively, were computed as follows:

$$\bar{V}_{B(\cdot)} = \frac{\sum_{b=1}^{B} \hat{V}_{\cdot}^{(b)}}{B} \text{ and } s_B(\hat{V}_{\cdot}) = \frac{1}{\sqrt{B}} \sqrt{\frac{\sum_{b=1}^{B} \left(\hat{V}_{\cdot}^{(b)} - \bar{V}_{B(\cdot)}\right)^2}{B-1}}$$

$$\bar{E}_{B(\cdot)} = \frac{\sum_{b=1}^{B} \hat{E}_{\cdot}^{(b)}}{B} \text{ and } s_B(\hat{E}_{\cdot}) = \frac{1}{\sqrt{B}} \sqrt{\frac{\sum_{b=1}^{B} \left(\hat{E}_{\cdot}^{(b)} - \bar{E}_{B(\cdot)}\right)^2}{B-1}}.$$

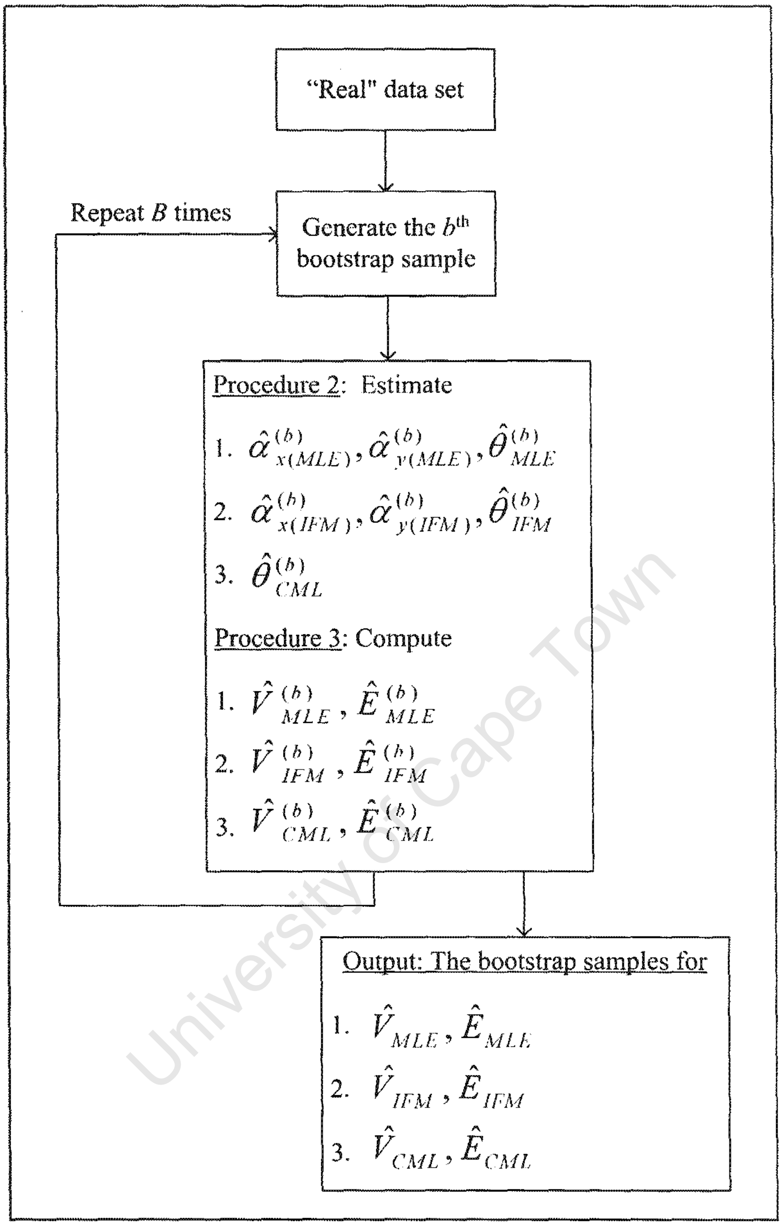

Figure 4.3: Flowchart of the algorithm used to generate the bootstrap samples for the VaR and ES estimates.

(ii) The skewness and kurtosis of the bootstrap sample of the VaR estimates, denoted by $\bar{g}_B(\hat{V}.)$ and $\bar{k}_B(\hat{V}.)$ respectively, and the ES estimates, denoted by $\bar{g}_B(\hat{E}.)$ and $\bar{k}_B(\hat{E}.)$ respectively, were calculated. The routines for computing the two statistics are available in MATLAB under the function names `skewness.m` and `kurtosis.m`.

(iii) The histograms of the bootstrap samples of the VaR and ES estimates were obtained using the MATLAB function `hist.m`.

2. The specifics of the VaR and ES estimates:

   (i) The estimates of the bootstrap bias of the VaR and ES estimates, denoted by bias $(\hat{V}_\cdot)$ and bias $(\hat{E}_\cdot)$, provide a measure of the accuracy of the estimates (Efron and Tibshirani, 1993, p. 125). These were calculated as the difference between the estimates obtained from the "real" data set and the corresponding mean of the bootstrapped sample of the estimates, that is

$$\begin{aligned} \text{bias } (\hat{V}_\cdot) &= \hat{V}_\cdot - \bar{V}_{B(\cdot)} \\ \text{bias } (\hat{E}_\cdot) &= \hat{E}_\cdot - \bar{E}_{B(\cdot)}. \end{aligned}$$

   (ii) The confidence intervals of the VaR and ES estimates were also computed. The 95% confidence interval of the VaR and ES estimates are given by:

$$(C^{-1}(\alpha/2), C^{-1}(1-\alpha/2))$$

   where $C^{-1}(\alpha/2)$ and $C^{-1}(1-\alpha/2)$ are the 2.5%-quantile and the 97.5%-quantile of the corresponding bootstrap VaR and ES samples respectively (Efron and Tibshirani, 1993).

4.4.4 Results

The two stages described in Sections 4.4.1 and 4.4.2 were repeated for the other eleven scenarios and the results of the estimates of the VaR and the ES obtained for all the twelve scenarios are shown in Table 4.2.

	$\hat{V}_{MLE}$	$\hat{V}_{IFM}$	$\hat{V}_{CML}$	$\hat{E}_{MLE}$	$\hat{E}_{IFM}$	$\hat{E}_{CML}$
Scenario 1	-1.260	-1.393	-1.381	-1.585	-1.682	-1.683
Scenario 2	-1.527	-1.700	-1.637	-1.896	-1.950	-1.959
Scenario 3	-1.289	-1.360	-1.357	-1.619	-1.689	-1.577
Scenario 4	-1.480	-1.543	-1.495	-1.959	-1.954	-1.845
Scenario 5	-1.691	-1.737	-1.800	-2.244	-2.356	-2.403
Scenario 6	-1.901	-1.970	-2.081	-2.552	-2.719	-2.748
Scenario 7	-1.657	-1.664	-1.686	-2.183	-2.321	-2.076
Scenario 8	-1.777	-1.843	-1.811	-2.571	-2.577	-2.430
Scenario 9	-1.516	-1.543	-1.565	-1.865	-1.966	-2.004
Scenario 10	-1.681	-1.756	-1.831	-2.151	-2.264	-2.296
Scenario 11	-1.447	-1.484	-1.492	-1.838	-1.931	-1.787
Scenario 12	-1.612	-1.681	-1.637	-2.199	-2.197	-2.099

Table 4.2: The three VaR estimates and the three ES estimates for the twelve scenarios.

The bootstrap procedure described in Section 4.4.3 was performed for all the twelve scenarios and the histograms presented in Figure 4.4 show how the VaR and ES estimates varied across the bootstrap samples for Scenario 1.

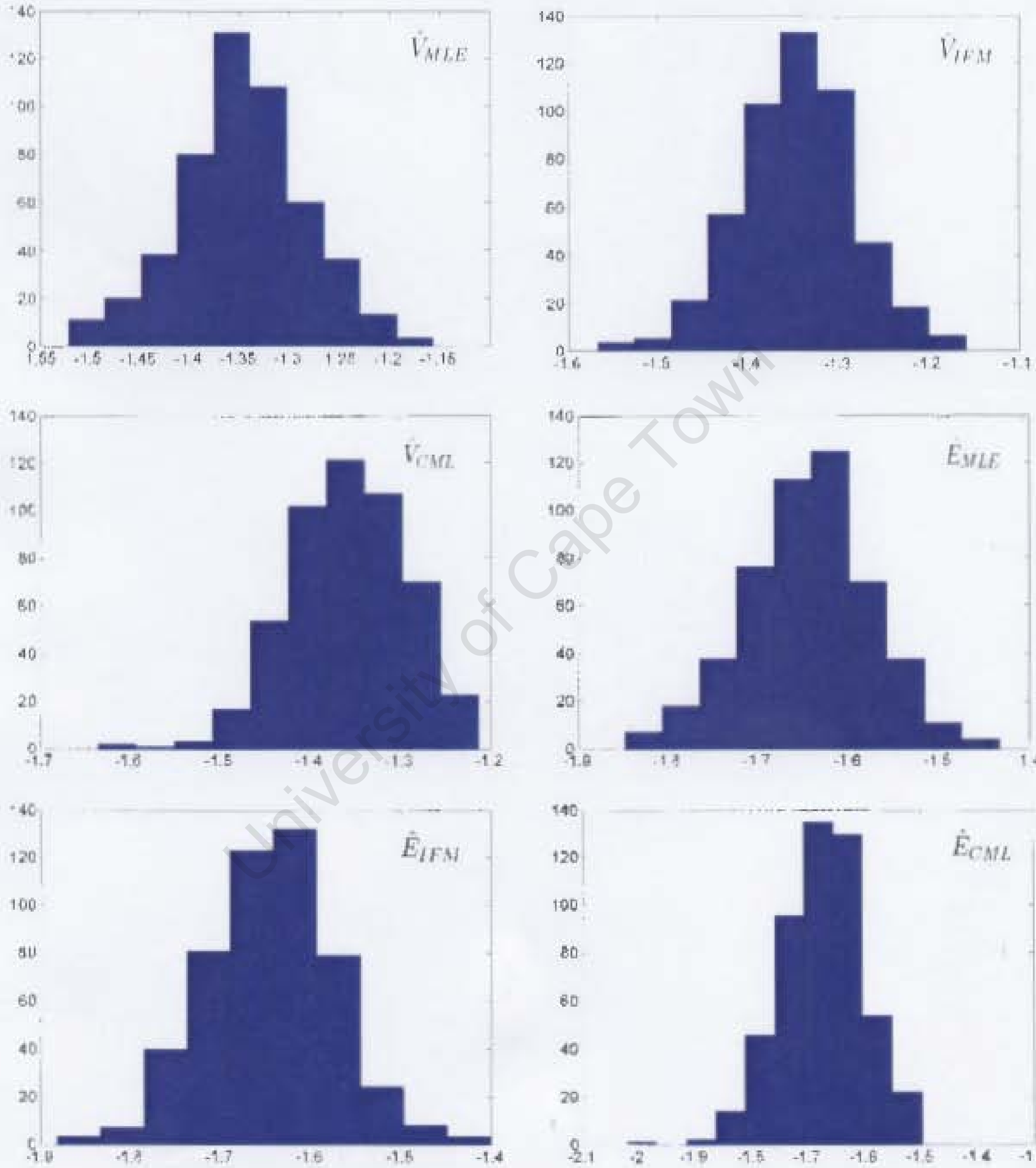

Figure 4.4: Histograms of the bootstrapped VaR and ES estimates obtained for Scenario 1.

The results of the properties of the VaR and the ES estimates obtained for all the twelve scenarios are shown in Table 4.3 and in Table 4.4.

		$\hat{V}_{MLE}$	$\hat{V}_{IFM}$	$\hat{V}_{CML}$
Scenario 1	Mean	-1.347	-1.345	-1.358
X~N(0,1)	Standard error	0.003	0.003	0.003
Y~N(0,1)	Skewness	-0.100	-0.145	-0.364
$\theta = 2$	Kurtosis	3.136	3.370	3.429
	Bias	0.086	-0.048	-0.024
	95% CI	(-1.482,-1.225)	(-1.474,-1.224)	(-1.484,-1.242)
Scenario 2	Mean	-1.603	-1.611	-1.617
X~N(0,1)	Standard error	0.003	0.003	0.003
Y~N(0,1)	Skewness	-0.066	-0.243	-0.364
$\theta = 10$	Kurtosis	3.241	3.296	3.450
	Bias	0.075	-0.089	-0.020
	95% CI	(-1.748,-1.464)	(-1.750,-1.484)	(-1.761,-1.490)
Scenario 3	Mean	-1.276	-1.277	-1.287
X~N(0,1)	Standard error	0.003	0.003	0.003
Y~N(0,1)	Skewness	-0.025	-0.135	0.138
$\rho = 0.3297$	Kurtosis	2.817	3.107	3.059
	Bias	-0.013	-0.083	-0.071
	95% CI	(-1.394,-1.158)	(-1.405,-1.157)	(-1.413,-1.154)
Scenario 4	Mean	-1.503	-1.510	-1.499
X~N(0,1)	Standard error	0.003	0.003	0.004
Y~N(0,1)	Skewness	-0.083	-0.211	-0.158
$\rho = 0.8653$	Kurtosis	2.864	3.211	3.154
	Bias	0.023	-0.032	0.004
	95% CI	(-1.648,-1.357)	(-1.659,-1.378)	(-1.667,-1.347)
Scenario 5	Mean	-1.699	-1.703	-1.748
X~t_5	Standard error	0.004	0.004	0.005
Y~t_5	Skewness	-0.177	-0.237	-0.234
$\theta = 2$	Kurtosis	2.834	3.067	3.303
	Bias	0.009	-0.034	-0.052
	95% CI	(-1.873,-1.536)	(-1.876,-1.556)	(-1.984,-1.539)
Scenario 6	Mean	-1.974	-1.979	-2.035
X~t_5	Standard error	0.005	0.004	0.006
Y~t_5	Skewness	-0.105	-0.252	-0.313
$\theta = 10$	Kurtosis	2.939	3.081	3.166
	Bias	0.073	0.009	-0.047
	95% CI	(-2.169,-1.780)	(-2.157,-1.822)	(-2.330,-1.792)
Scenario 7	Mean	-1.568	-1.566	-1.604
X~t_5	Standard error	0.004	0.003	0.005
Y~t_5	Skewness	-0.200	-0.159	-0.042
$\rho = 0.3297$	Kurtosis	2.731	3.196	3.212
	Bias	-0.090	-0.098	-0.082
	95% CI	(-1.737,-1.415)	(-1.733,-1.420)	(-1.818,-1.397)
Scenario 8	Mean	-1.796	-1.810	-1.819
X~t_5	Standard error	0.004	0.004	0.006
Y~t_5	Skewness	-0.214	-0.187	-0.343
$\rho = 0.8653$	Kurtosis	2.578	2.990	3.310
	Bias	0.019	-0.033	0.008
	95% CI	(-1.975,-1.636)	(-1.986,-1.651)	(-2.108,-1.587)
Scenario 9	Mean	-1.502	-1.503	-1.533
X~t_{10}	Standard error	0.003	0.003	0.004
Y~t_{10}	Skewness	-0.147	-0.126	-0.164
$\theta = 2$	Kurtosis	2.838	2.963	3.416
	Bias	-0.014	-0.040	-0.032
	95% CI	(-1.638,-1.370)	(-1.627,-1.388)	(-1.706,-1.369)
Scenario 10	Mean	-1.750	-1.753	-1.791
X~t_{10}	Standard error	0.003	0.003	0.004
Y~t_{10}	Skewness	-0.031	-0.255	-0.297
$\theta = 10$	Kurtosis	2.900	3.038	3.242
	Bias	0.069	-0.003	-0.040
	95% CI	(-1.897,-1.601)	(-1.894,-1.629)	(-1.991,-1.609)
Scenario 11	Mean	-1.406	-1.403	-1.423
X~t_{10}	Standard error	0.003	0.003	0.004
Y~t_{10}	Skewness	-0.085	-0.084	0.087
$\rho = 0.3297$	Kurtosis	2.610	3.052	3.141
	Bias	-0.041	-0.081	-0.069
	95% CI	(-1.529,-1.278)	(-1.530,-1.287)	(-1.581,-1.260)
Scenario 12	Mean	-1.641	-1.647	-1.644
X~t_{10}	Standard error	0.003	0.003	0.004
Y~t_{10}	Skewness	-0.165	-0.149	-0.235
$\rho = 0.8653$	Kurtosis	2.617	3.016	3.216
	Bias	0.029	-0.034	0.007
	95% CI	(-1.785,-1.513)	(-1.791,-1.518)	(-1.862,-1.460)

Table 4.3: The mean, standard error, skewness, kurtosis and bias of the three VaR estimates for the twelve scenarios obtained from the bootstrap procedure.

		$\hat{E}_{MLE}$	$\hat{E}_{IFM}$	$\hat{E}_{CML}$
Scenario 1	Mean	-1.643	-1.644	-1.675
X~N(0,1)	Standard error	0.003	0.003	0.003
Y~N(0,1)	Skewness	-0.088	0.015	-0.356
$\theta = 2$	Kurtosis	3.145	3.291	3.450
	Bias	0.058	-0.038	-0.009
	95% CI	(-1.792,-1.508)	(-1.777,-1.512)	(-1.824,-1.543)
Scenario 2	Mean	-1.916	-1.925	-1.949
X~N(0,1)	Standard error	0.003	0.004	0.004
Y~N(0,1)	Skewness	-0.072	-0.027	-0.299
$\theta = 10$	Kurtosis	3.208	3.273	3.626
	Bias	0.020	-0.025	-0.011
	95% CI	(-2.075,-1.775)	(-2.072,-1.760)	(-2.111,-1.800)
Scenario 3	Mean	-1.594	-1.593	-1.593
X~N(0,1)	Standard error	0.003	0.003	0.003
Y~N(0,1)	Skewness	-0.066	0.051	0.175
$\rho = 0.3297$	Kurtosis	2.858	3.069	2.760
	Bias	-0.024	-0.096	0.016
	95% CI	(-1.738,-1.460)	(-1.733,-1.443)	(-1.720,-1.455)
Scenario 4	Mean	-1.879	-1.884	-1.874
X~N(0,1)	Standard error	0.004	0.004	0.004
Y~N(0,1)	Skewness	0.005	-0.052	-0.017
$\rho = 0.8653$	Kurtosis	2.886	3.368	2.669
	Bias	-0.080	-0.070	0.028
	95% CI	(-2.046,-1.702)	(-2.052,-1.722)	(-2.052,-1.683)
Scenario 5	Mean	-2.316	-2.313	-2.348
X~t_5	Standard error	0.007	0.006	0.008
Y~t_5	Skewness	-0.664	-0.201	-0.612
$\theta = 2$	Kurtosis	4.532	2.872	4.193
	Bias	0.073	-0.044	-0.055
	95% CI	(-2.631,-2.068)	(-2.584,-2.065)	(-2.700,-2.056)
Scenario 6	Mean	-2.643	-2.654	-2.623
X~t_5	Standard error	0.007	0.006	0.007
Y~t_5	Skewness	-0.694	-0.159	-0.470
$\theta = 10$	Kurtosis	6.053	2.808	3.584
	Bias	0.091	-0.065	-0.125
	95% CI	(-2.958,-2.349)	(-2.945,-2.387)	(-2.970,-2.341)
Scenario 7	Mean	-2.134	-2.127	-2.102
X~t_5	Standard error	0.006	0.005	0.005
Y~t_5	Skewness	-0.365	-0.110	0.150
$\rho = 0.3297$	Kurtosis	3.058	3.470	2.687
	Bias	-0.050	-0.195	0.026
	95% CI	(-2.442,-1.901)	(-2.363,-1.892)	(-2.317,-1.871)
Scenario 8	Mean	-2.432	-2.449	-2.513
X~t_5	Standard error	0.007	0.006	0.008
Y~t_5	Skewness	-0.312	-0.207	-0.104
$\rho = 0.8653$	Kurtosis	2.977	3.170	2.690
	Bias	-0.140	-0.127	0.083
	95% CI	(-2.742,-2.163)	(-2.727,-2.200)	(-2.877,-2.149)
Scenario 9	Mean	-1.919	-1.919	-1.949
X~t_{10}	Standard error	0.004	0.004	0.005
Y~t_{10}	Skewness	-0.308	-0.086	-0.383
$\theta = 2$	Kurtosis	3.114	2.792	3.887
	Bias	0.054	-0.047	-0.055
	95% CI	(-2.109,-1.755)	(-2.084,-1.760)	(-2.160,-1.753)
Scenario 10	Mean	-2.196	-2.205	-2.204
X~t_{10}	Standard error	0.004	0.004	0.005
Y~t_{10}	Skewness	-0.142	-0.050	-0.341
$\theta = 10$	Kurtosis	3.038	2.714	3.450
	Bias	0.045	-0.059	-0.093
	95% CI	(-2.386,-2.006)	(-2.372,-2.040)	(-2.421,-2.018)
Scenario 11	Mean	-1.810	-1.804	-1.807
X~t_{10}	Standard error	0.004	0.004	0.004
Y~t_{10}	Skewness	-0.264	0.107	0.163
$\rho = 0.3297$	Kurtosis	2.953	3.354	2.714
	Bias	-0.027	-0.127	0.021
	95% CI	(-1.991,-1.657)	(-1.960,-1.638)	(-1.970,-1.630)
Scenario 12	Mean	-2.106	-2.110	-2.147
X~t_{10}	Standard error	0.004	0.004	0.006
Y~t_{10}	Skewness	-0.196	-0.122	-0.063
$\rho = 0.8653$	Kurtosis	2.750	3.171	2.679
	Bias	-0.093	-0.086	0.049
	95% CI	(-2.303,-1.936)	(-2.291,-1.940)	(-2.401,-1.886)

Table 4.4: The mean, standard error, skewness, kurtosis and bias of the three ES estimates for the twelve scenarios obtained from the bootstrap procedure.

From the results, the following observations can be made about the VaR and the ES estimates:

1. The VaR and the ES estimates obtained from the three estimation methods are all similar, specially the MLE and the IFM estimates. In fact, the bootstrapped mean of the ES estimates computed using the MLE and IFM methodology for Scenario 9 are identical.

2. Relative to the bootstrapped mean, the standard error of the VaR and the ES estimates is small, which implies that there is little variability in the bootstrapped VaR and ES estimates.

3. In most cases, the VaR and the ES estimates are slightly negatively skewed but in some cases they are slightly positively skewed as shown in Table 4.3 and Table 4.4 and in Figure 4.4.

4. The kurtosis of all the VaR and ES estimates are close to 3 in all scenarios except in Scenario 6 where the kurtosis of $\hat{E}_{MLE}$ is 6.053.

5. The CML method yielded the least biased VaR estimates in six scenarios and the least biased ES estimates in seven scenarios. The VaR and ES estimates obtained from the MLE method were the least biased in four scenarios and two scenarios respectively. The VaR and ES estimates obtained from the IFM method were the least biased in two scenarios and three scenarios respectively.

The VaR bounds described in Section 3.5 were used here as a benchmark interval to compare the confidence intervals of the VaR estimates obtained from the bootstrapping process. Note that it is not the aim of this mini-dissertation to investigate the properties of the VaR bounds and thus their computations were restricted to the known marginals, that is N(0,1), t_5 and t_{10}. Let $T = X + Y$ such that the portfolio return is given by $Z = T/2$. This implies that:

$$\begin{aligned} Pr(Z \leq z) &= Pr(T \leq 2z) \\ &= H_T(2z) \end{aligned}$$

where $H_T(.)$ is the distribution function of T. Therefore, for a particular value of α, $0 < \alpha < 1$, $z = 0.5H_T^{-1}(\alpha)$, where $H_T^{-1}(.)$ is the inverse of the distribution function of T. Therefore, the 95% VaR bounds of the two-asset portfolio whose return is given by the random variable Z were obtained by computing the 95% VaR bounds of T and then dividing the lower and upper VaR bounds of T by two. The 95% VaR bounds of T were computed by using the Williamson and Downs (1990) method outlined in Section 3.5, as follows:

1. Appropriate values of i and N to be used in Equations (3.10) and (3.11) for $\alpha = 0.05$ were selected.

2. The inverse functions (3.10) and (3.11) were evaluated using the true parameter values of the marginals. Furthermore, instead of $F_{x_1}^{-1}$ and $G_{y_1}^{-1}$, F^{-1} and G^{-1}, which are the inverse distribution function of X and Y, were used.

To obtain a good estimate of the VaR bounds at the 95% confidence level, N was set to 10^5 such that to evaluate $H_L^{-1}\left(\frac{i+1}{N}\right)$ as $H_L^{-1}(0.05)$, $i = 4999$ and to evaluate $H_U^{-1}\left(\frac{i}{N}\right)$ as $H_U^{-1}(0.05)$, $i = 5000$. The VaR bounds obtained using the marginals distributed as N(0,1) are (-1.960, 0.063) and they are used as a benchmark interval for the confidence intervals of the VaR estimates for Scenario 1 to Scenario 4. The VaR bounds obtained using the marginals distributed as t_5 are (-2.571, 0.066) and they are used as a benchmark interval for the confidence intervals of the VaR estimates for Scenario 5 to Scenario 8. The VaR bounds obtained using the marginals distributed as t_{10} are (-2.228, 0.064) and they are used as a benchmark interval for the confidence intervals of the VaR estimates for Scenario 9 to Scenario 12. It can be observed that in all scenarios, the bootstrap confidence intervals fall well inside the corresponding VaR bounds.

Furthermore, comparing the VaR and the ES estimates across the twelve scenarios, the following observations can be made:

1. For the case of the same marginals and the same dependence, as measured by Kendall's tau, the VaR and the ES estimates obtained from the Gaussian copula are smaller, in absolute terms, than those obtained using the Frank copula.

2. Whilst keeping the copula and the marginals constant and increasing the dependence between the two risky assets, as measured by Kendall's tau, the VaR and the ES estimates increase.

3. For the case of the same copula and same copula parameter but different marginals, different estimates of the VaR and the ES estimates are obtained. Specifically, the estimates of the VaR and ES obtained using the t_5 marginals are larger, in absolute value, than the estimates obtained using the t_{10} marginals, which are in turn greater than the estimates obtained using the standard normal marginals.

4.5 The performance of the the VaR and ES estimates

4.5.1 The procedure used

For a particular scenario, in order to investigate the performance of the VaR and ES estimates, Stage 1 and Stage 2 described in Sections 4.4.1 and 4.4.2 were repeated

S times. Specifically, for each repetition of the simulation process, a "real" data set was generated and the VaR and ES estimates of this data set were estimated using the MLE, IFM and CML estimation methods. After each repetition of the simulation process, the VaR and ES estimates were compared to the true VaR and the true ES, denoted by V_{true} and E_{true}. The true VaR and the true ES were computed using Procedure 3 and the true parameters of the copula and the marginals. To obtain a good estimate of V_{true} and E_{true}, d was set to 10^6 in Procedure 3. Let $\hat{\alpha}^{(s)}_{x(\cdot)}$, $\hat{\alpha}^{(s)}_{y(\cdot)}$ and $\hat{\theta}^{(s)}_{\cdot}$ denote the estimates of the marginals of X, Y and the copula function respectively and let $\hat{V}^{(s)}_{\cdot}$ and $\hat{E}^{(s)}_{\cdot}$ denote the VaR and ES estimate respectively, obtained for the s^{th} simulation, $s = 1, \ldots, S$, where $\cdot$ represents MLE, IFM or CML. The algorithm for repeating the simulation the process S times is illustrated as a flowchart in Figure 4.5.

For a particular scenario, the performance of the generic estimators $\hat{V}_{\cdot}$ and $\hat{E}_{\cdot}$ were evaluated as follows:

1. The bias was computed as follows:

$$\text{bias}(\hat{V}_{\cdot}) = \frac{\sum_{s=1}^{S}\left(\hat{V}^{(s)}_{\cdot} - V_{true}\right)}{S} \tag{4.1}$$

$$\text{bias}(\hat{E}_{\cdot}) = \frac{\sum_{s=1}^{S}\left(\hat{E}^{(s)}_{\cdot} - E_{true}\right)}{S} \tag{4.2}$$

2. The mean square error (MSE) was computed as follows:

$$\text{MSE}(\hat{V}_{\cdot}) = \frac{\sum_{s=1}^{S}\left(\hat{V}^{(s)}_{\cdot} - V_{true}\right)^2}{S} \tag{4.3}$$

$$\text{MSE}(\hat{E}_{\cdot}) = \frac{\sum_{s=1}^{S}\left(\hat{E}^{(s)}_{\cdot} - E_{true}\right)^2}{S} \tag{4.4}$$

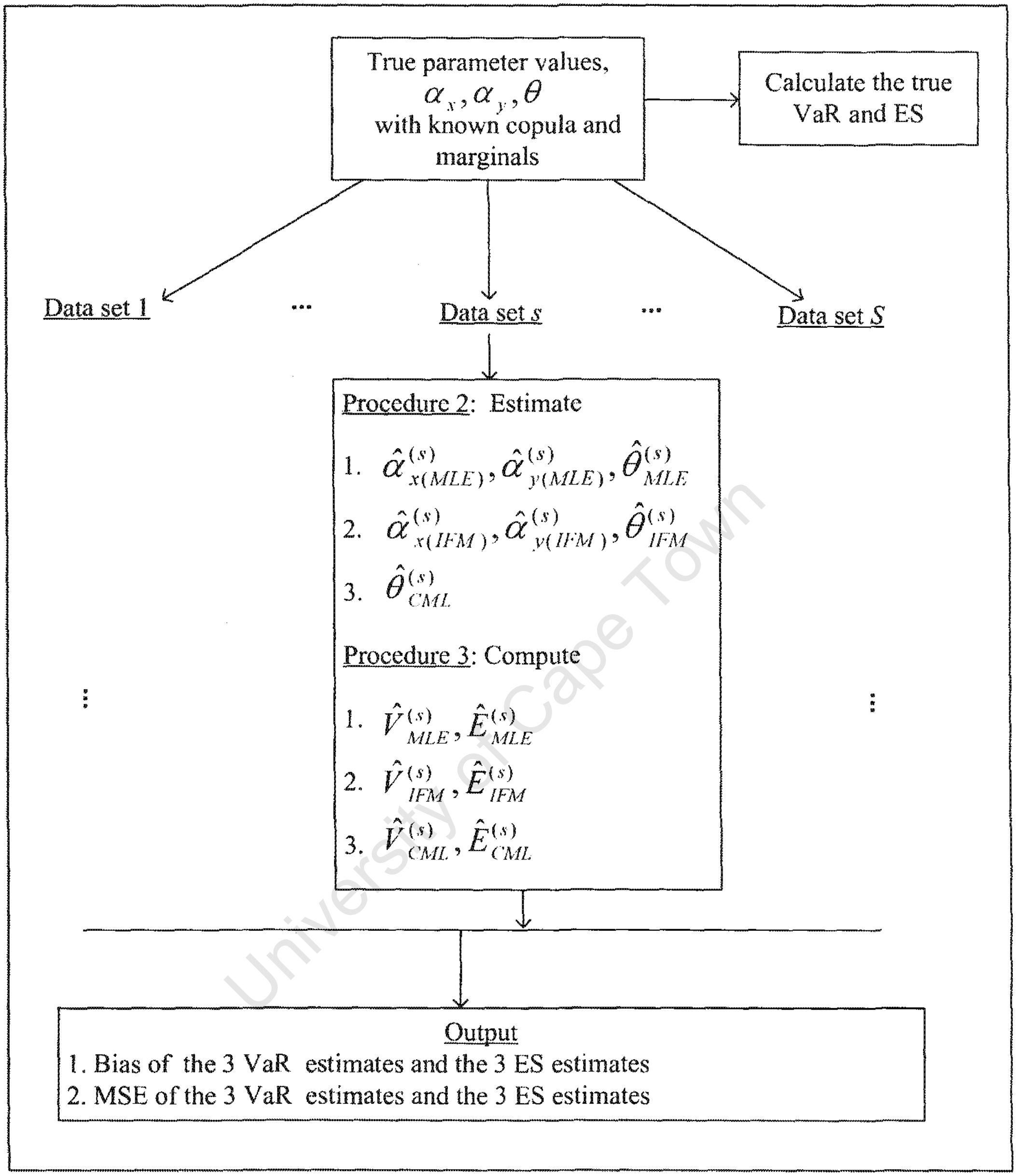

Figure 4.5: Flowchart of the algorithm used for running the simulation process S times.

4.5.2 Results

Given that the simulation process was very computationally intensive, S was set to 500 and the results obtained for all the twelve scenarios are shown in Table 4.5.

		$\hat{V}_{MLE}$	$\hat{V}_{IFM}$	$\hat{V}_{CML}$	$\hat{E}_{MLE}$	$\hat{E}_{IFM}$	$\hat{E}_{CML}$
Scenario 1	Bias	0.0030	0.0023	0.0011	0.0079	0.0094	0.0071
	MSE	0.0038	0.0040	0.0033	0.0044	0.0049	0.0045
Scenario 2	Bias	0.0054	0.0006	0.0004	0.0098	0.0088	0.0066
	MSE	0.0045	0.0048	0.0048	0.0051	0.0056	0.0058
Scenario 3	Bias	-0.0068	-0.0055	-0.0072	0.0619	0.0615	0.0614
	MSE	0.0045	0.0046	0.0040	0.0060	0.0060	0.0059
Scenario 4	Bias	-0.0085	-0.0046	-0.0055	0.0712	0.0741	0.0816
	MSE	0.0059	0.0067	0.0069	0.0080	0.0086	0.0104
Scenario 5	Bias	0.0051	0.0025	-0.0014	0.0082	0.0161	0.0120
	MSE	0.0077	0.0078	0.0088	0.0200	0.0196	0.0226
Scenario 6	Bias	0.0064	0.0040	-0.0018	0.0153	0.0171	0.0114
	MSE	0.0100	0.0099	0.0138	0.0259	0.0246	0.0313
Scenario 7	Bias	0.0720	0.0753	0.0792	0.1255	0.1251	0.1352
	MSE	0.0084	0.0091	0.0100	0.0246	0.0254	0.0284
Scenario 8	Bias	0.0863	0.0958	0.1059	0.1554	0.1702	0.1906
	MSE	0.0116	0.0145	0.0180	0.0394	0.0459	0.0599
Scenario 9	Bias	0.0041	0.0032	0.0005	0.0097	0.0139	0.0088
	MSE	0.0046	0.0048	0.0051	0.0078	0.0081	0.0089
Scenario 10	Bias	0.0062	0.0027	-0.0002	0.0128	0.0137	0.0084
	MSE	0.0057	0.0056	0.0079	0.0100	0.0097	0.0125
Scenario 11	Bias	0.0567	0.0614	0.0625	0.0812	0.0819	0.0865
	MSE	0.0052	0.0058	0.0061	0.0101	0.0108	0.0117
Scenario 12	Bias	0.0664	0.0737	0.0830	0.0964	0.1073	0.1214
	MSE	0.0068	0.0086	0.0110	0.0149	0.0180	0.0233

Table 4.5: The bias and MSE of the VaR and ES estimates obtained by repeating the simulation process $S = 500$ times.

The following observations can be made from the results:

1. The MLE methodology yielded the least biased VaR and ES estimates, but nevertheless the bias of the VaR and ES estimates obtained from IFM and CML methodologies are very similar to the bias of the VaR and ES estimates computed using the MLE estimation method.

2. Similarly, the MSE of the VaR and the ES estimates obtained from the MLE method are the smallest. However they do not differ much from the MSE of the VaR and the ES estimates obtained from the IFM and CML methods.

Chapter 5

Risk measures for a real data set

In this chapter, the copula-based Monte Carlo approach of estimating the Value-at-Risk (VaR) and the expected shortfall (ES) of a portfolio of two risky assets is illustrated with a set of real data. The portfolio consists of one stock, Anglo American Plc (Anglo) and a market index, the Top 40 Index (TOP 40). MATLAB Version 7.3 (R2006b) has been used to fit the data set and generate the results. The structure of the chapter is as follows. In the first section, the data set is described and some summary statistics are presented. The subsequent section outlines the procedure of selecting the marginals as well as a copula that best fits the data, followed by the estimation of the best-fitting copula and the estimation of the VaR and the ES of the portfolio. The properties of the VaR and the ES estimates are then investigated, using the bootstrapping technique, and the results are discussed. The next section involves the estimation of the VaR bounds of the portfolio and finally, the last section of the chapter describes the backtesting procedure used to evaluate the performance of the VaR estimates generated from the fitted models.

5.1 Description of the portfolio and the data

Five years of daily closing prices of Anglo and the TOP 40 from the 11^{th} March 2002 to the 7^{th} March 2007, obtained from the database of Datastream, have been used such that there is a total of 1247 observations. The time span includes the period of high volatility which was experienced by both the domestic and international markets in 2002. The remainder of the time span includes the period of the bull market which started in the second quarter of 2003, with some periods of volatility and minor market corrections (Laubscher, 2007).

Let $p_{1,t}$ and $p_{2,t}$ represent the closing prices of Anglo and the TOP 40 respectively at some time t. If the random variables X and Y represent the log returns of Anglo and the TOP 40 respectively, then the log returns of the two risky assets at time t

are given by

$$x_t = ln\left(\frac{p_{1,t}}{p_{1,t-1}}\right)$$
$$y_t = ln\left(\frac{p_{2,t}}{p_{2,t-1}}\right)$$

for $t = 1, \ldots, T$ where $T = 1246$ is the total sample size of log returns. The pairs of log returns (x_t, y_t) were assumed to be identically and independently distributed from a bivariate distribution, $H(x, y)$. Furthermore, each asset was assumed to be equally weighted in the portfolio such that the return of the portfolio at time t, denoted by z_t, is given by

$$z_t = 0.5x_t + 0.5y_t \quad \text{for } t = 1, \ldots, T. \tag{5.1}$$

Some descriptive statistics of the log returns of the two assets, which have been computed in the MATLAB Statistics Toolbox, are presented in Table 5.1 and the histograms of the log returns of Anglo and the TOP 40 for the time period $t = 1, \ldots, 1246$, are shown in Figure 5.1. Note that the sample means of the log returns of both time series have small positive values. The coefficients of variation show that there is variability in the log returns of Anglo and the TOP 40 relative to their mean for the time period under consideration. Also, the log returns of Anglo varies more than the TOP 40 as shown by the standard errors and the coefficients of variation. The time series of the log returns of Anglo and the TOP 40 are both slightly negatively skewed, as shown in their histograms in Figure 5.1 and in Table 5.1. From Table 5.1, it can also be noted that both time series exhibit excess kurtosis which implies that they are fat-tailed. The Jarque-Bera test of normality rejects the null hypothesis that the time series of the log returns of the two assets are normally distributed since the test statistics, displayed in Table 5.1, are greater than the critical value of 5.943 at the 5% significance level .

	Anglo	**TOP 40**
Mean	0.000444	0.000591
Standard error	0.000589	0.000342
Coefficient of variation	46.8080	20.4248
Skewness	-0.0003	-0.1555
Kurtosis	3.9644	5.1889
Jarque-Bera Statistic	48.2944	253.5376

Table 5.1: Some descriptive statistics of the log returns of Anglo and the TOP 40.

The plot of the log returns of the TOP 40 against the log returns of Anglo, that is y_t against x_t for $t = 1, \ldots, 1246$, is displayed in Figure 5.2. The scatter plot shows that there is a very high linear correlation between the two assets. In fact, Pearson's correlation coefficient is 0.8333, Kendall's tau is 0.6351 and Spearman's rho is 0.8239. Furthermore it is noted that extreme values in the log returns of Anglo occurred simultaneously with extreme values in the log returns of the TOP 40.

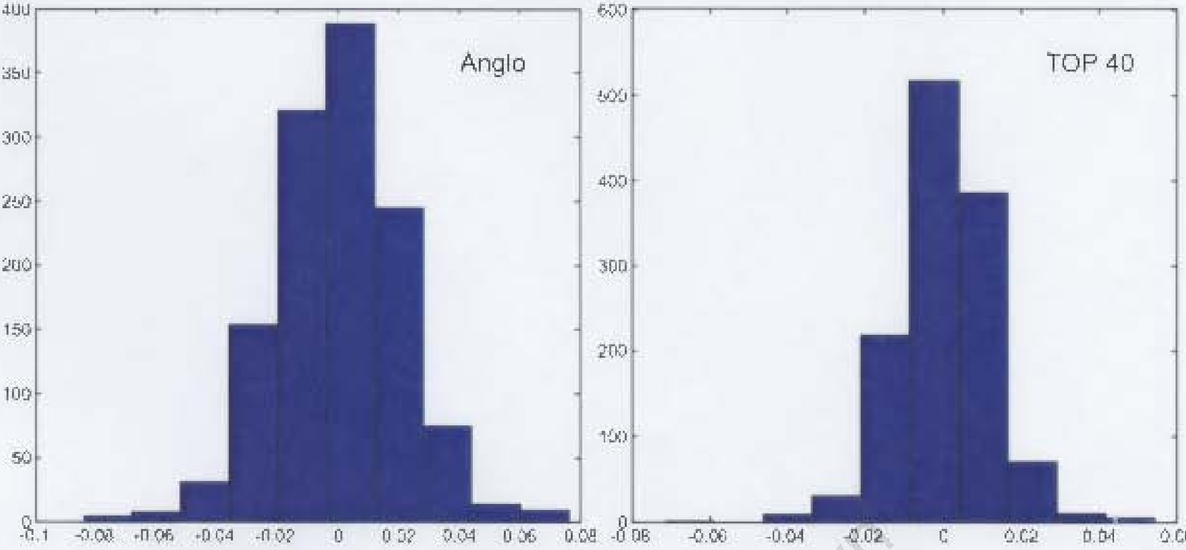

Figure 5.1: The histograms of the log returns of Anglo and the TOP 40.

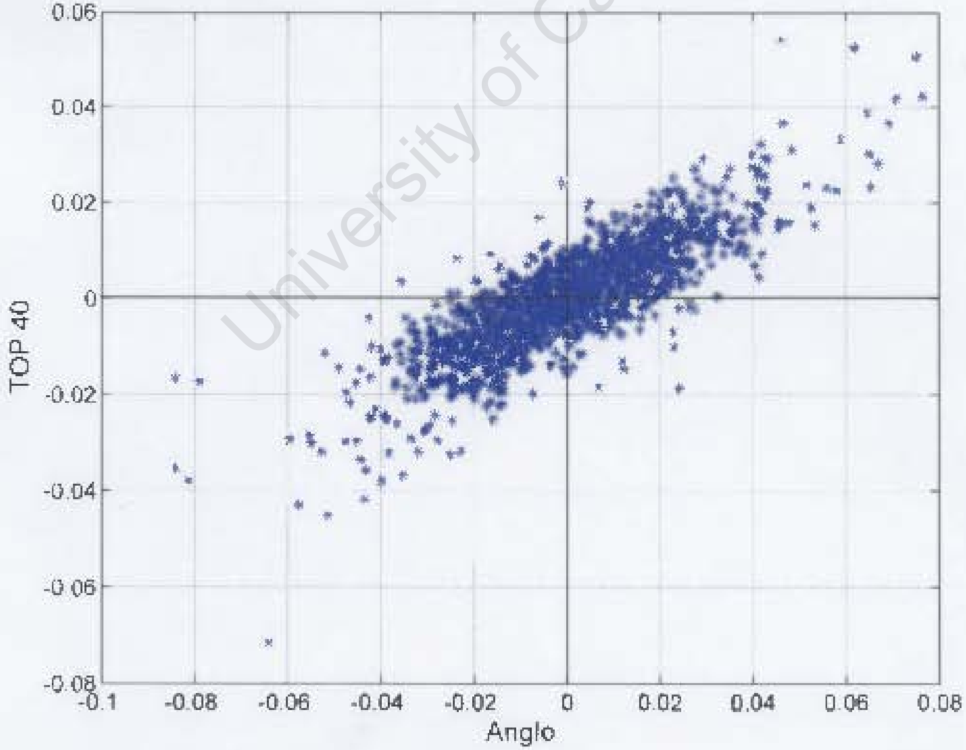

Figure 5.2: The scatter plot of the log returns of the TOP 40 against the log returns of Anglo.

The aim of the study is to forecast the one-day ahead VaR and the one-day ahead expected shortfall of the portfolio at time $t = 1246$ denoted by V_{1247} and E_{1247} respectively using copula. Therefore, the log returns (x_t, y_t), $t = 1, \ldots, 1246$, were used to compute estimates of V_{1247} and E_{1247} denoted by $\hat{V}_{1247}$ and $\hat{E}_{1247}$ at the 95% confidence level using the copula-based Monte Carlo approach.

5.2 Selection of the marginals and the copula

Before estimating the VaR and the ES of the portfolio, a bivariate distribution which adequately describes the data set had be specified. This was done by specifying the marginal distributions of the log returns of Anglo and the TOP 40 as well as a copula function which connects the two risky assets to form the bivariate distribution.

As the time series of the log returns of Anglo and the TOP 40 exhibited fat tails, a Student's t distribution should have provided an appropriate fit. However, the Student's t distribution did not fit the data adequately as both time series have a non-zero mean and a standard deviation which is smaller than one. As a result, if the Student's t distribution is used as the marginal of the log returns of Anglo and that of the TOP 40, the nonlinear optimisation algorithm for estimating the degrees of freedom of the marginals and the parameter of the copula presents problems in convergence. Therefore the more conservative t-location scale distribution was used to model the marginal distribution of the two risky assets. Note that a random variable X which follows a t-location scale distribution, written as $X \sim t(\mu, \sigma, \nu)$, where μ is the location parameter, $\sigma > 0$ is the scale parameter and $\nu > 0$ is the shape parameter, is such that $\left(\dfrac{X-\mu}{\sigma}\right)$ has a Student's t distribution with ν degrees of freedom. The t-location scale distribution was therefore used as the marginals of the log returns of Anglo and the TOP 40.

Given the wide selection of well-documented copulas that are available in the literature, the problem of selecting a suitable copula which would adequately capture the dependence structure between the log returns of Anglo and the TOP 40 had to be addressed. Because of their desirable tail dependence properties and their simplicity, the three Archimedean copulas introduced in Chapter 2, namely the Frank, Gumbel and Clayton copulas, were considered. The method introduced by Genest and Rivest (1993), which is discussed in Chapter 2, was used to select the best-fitting Archimedean copula amongst these three copulas.

Firstly the distribution function $K(z)$, which is defined in Equation (2.34), was estimated non-parametrically to yield $\hat{K}(z)$ as outlined in Section 2.6.4. Kendall's tau is given by 0.6351 and thus using the `copulaparam.m` routine in MATLAB, the parameter estimates of the Clayton, Frank and Gumbel copulas were obtained as 3.482, 8.950 and 2.741 respectively. Once the parameter associated with each copula was estimated, $K(z)$ was estimated parametrically for each copula as discussed in

Section 2.6.4. The plots of the parametric estimates of $K(z)$ for the Clayton, Frank and Gumbel copulas denoted by $K_C(z)$, $K_F(z)$ and $K_G(z)$ respectively as well as that of the empirical estimate of $K(z)$, denoted by $K_h(z)$, plotted against z for $0 \leq z \leq 1$, are shown in Figure 5.3.

It is clear from the graph in Figure 5.3 that the parametric estimate of $K(z)$ for the Frank copula, represented by the brown line, is comparatively closest to the nonparametric estimate of $K(z)$, represented by the green line. To verify this observation, the sum of the squared differences between the parametric and the nonparametric estimate of $K(z)$ for all three copulas was evaluated across all values of z for $0 \leq z \leq 1$, using Equation (2.38) in Chapter 2. The results obtained are 0.1007, 0.0222 and 0.0781 for the Clayton, Frank and Gumbel copulas respectively which confirm that the Frank copula was the best fitting copula. The Frank copula was therefore used to model the dependence between the log returns of Anglo and the TOP 40.

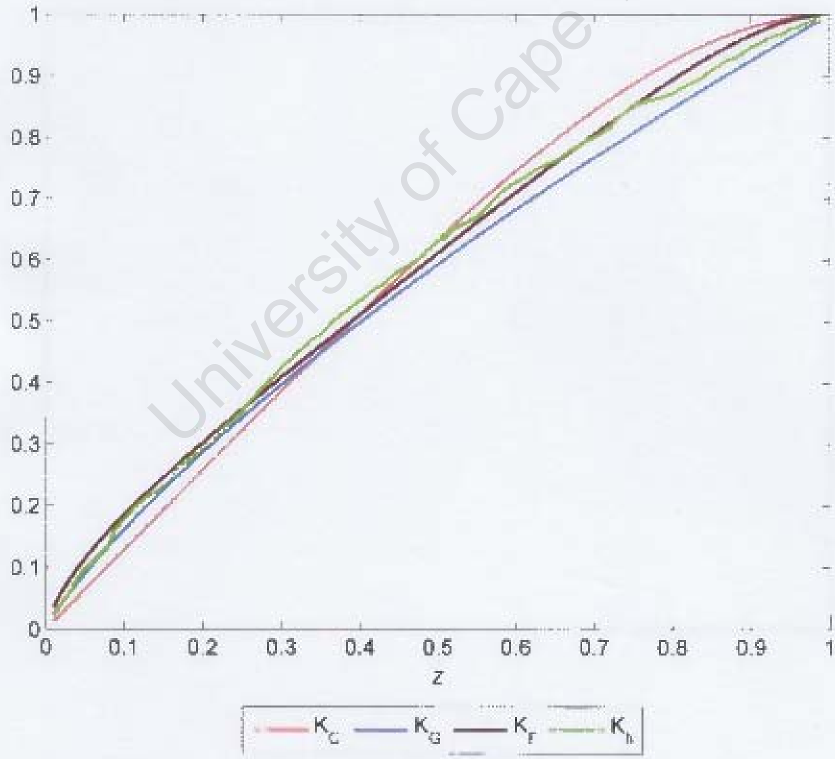

Figure 5.3: Plot of the parametric and nonparametric estimates of $K(z)$ against the corresponding z values for the Clayton, Frank and Gumbel copulas.

5.3 Estimating the risk measures of the portfolio

Assume now that the marginals of the log returns of Anglo and the TOP 40 are given by the t-location scale distribution and that the Frank copula captures their dependence structure. Then, the one-day ahead VaR and the one-day ahead ES of the two-asset portfolio have to be estimated at time $t = 1246$, at the 95% confidence level, by using the copula-based Monte Carlo approach. Therefore V_{1247} and E_{1247} were estimated using the whole data set of log returns of Anglo and the TOP 40, (x_t, y_t), for $t = 1, \ldots, 1246$. However, before estimating the Value-at-Risk and the expected shortfall of the portfolio, the parameters of the marginals and the copula have to be estimated.

5.3.1 Estimation of the marginals and the copula

The three estimation methods discussed in this mini-dissertation, namely maximum likelihood estimation (MLE), inference function for margins (IFM) and canonical maximum likelihood (CML), were used to estimate the parameters of the marginals and the copula function.

"Standardising" the log returns

As already noted, the marginals of the log returns of the two risky assets were assumed to have a t-location scale distribution. However, MATLAB has built-in routines to evaluate the distribution function, the density function and the inverse distribution function for the Student's t distribution but not for the t-location scale distribution. Therefore, before performing the MLE and the IFM estimation procedures, the log returns of Anglo, x_t, and the log returns of the TOP 40, y_t, $t = 1, \ldots, 1246$, were "standardised" such that a t-distribution could be fitted to the resulting data sets. Note that this did not affect the CML methodology as empirical distributions were used to estimate the marginals.

To "standardise" the log returns of Anglo and the TOP 40, the location and scale parameters of the two risky assets, denoted by μ_x, σ_x and by μ_y, σ_y respectively, have to be estimated by separately fitting a t-location scale distribution to x_t and to y_t, $t = 1, \ldots, 1246$. MATLAB has a routine, `mle.m`, which returns the maximum likelihood estimates for a wide range of distributions including the t-location scale distribution. The parameter estimates of the mean and standard deviation of x_t and y_t denoted by $\hat{\mu}_x, \hat{\sigma}_x$ and by $\hat{\mu}_y, \hat{\sigma}_y$ respectively are shown in Table 5.2. The fit of the t-location scale distribution to the log returns of Anglo and the TOP 40 is displayed in Figure 5.4. It can be observed the the distribution of the log returns of Anglo has fatter tails than that of the log returns of the TOP 40. The estimates of the location and the scale parameters were then used to adjust the log returns x_t and y_t to "standardised" values $x_{(adj)t}$ and $y_{(adj)t}$, $t = 1, \ldots, 1246$, such that a

Log returns of Anglo		Log returns of the TOP 40	
$\hat{\mu}_x$	0.000409	$\hat{\mu}_y$	0.000738
$\hat{\sigma}_x$	0.018521	$\hat{\sigma}_y$	0.010036
$\hat{\nu}_x$	9.734614	$\hat{\nu}_y$	6.543564

Table 5.2: The estimates of the location, scale and degrees of freedom of the distribution of the log returns of Anglo and the TOP 40.

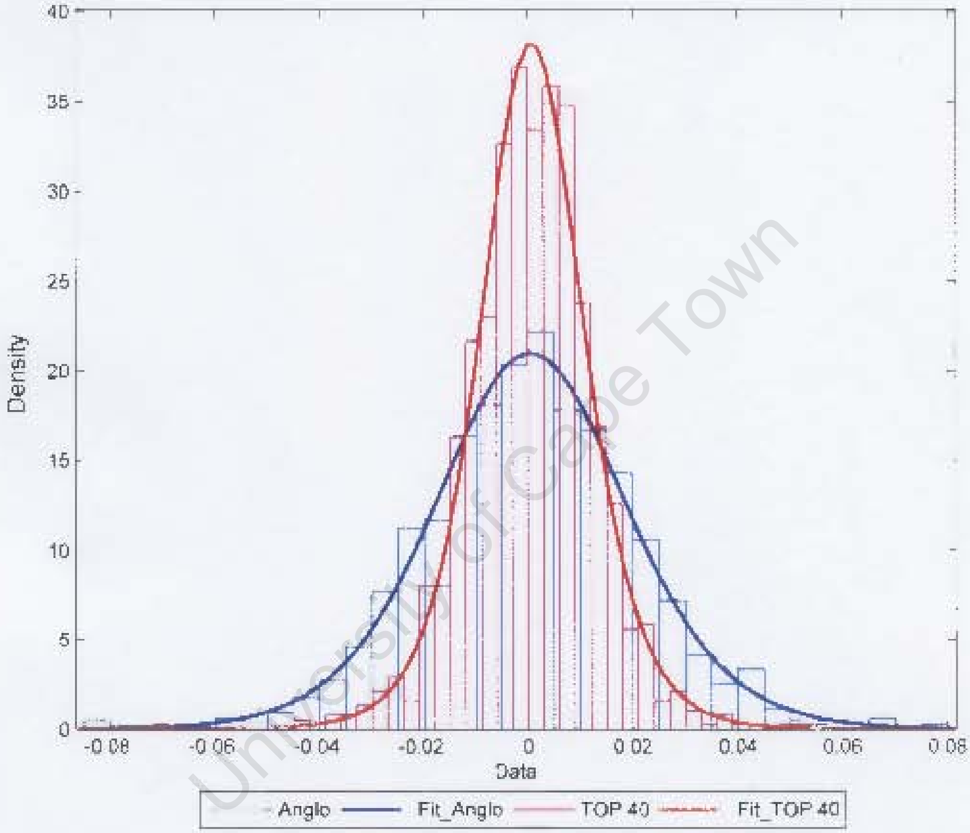

Figure 5.4: The fit of the t-location scale distribution to the log returns of Anglo and the TOP 40.

Student's t distribution could then be fitted to the "standardised" data sets, that is

$$\begin{aligned} x_{(adj)t} &= \frac{x_t - \hat{\mu}_x}{\hat{\sigma}_x} \\ y_{(adj)t} &= \frac{y_t - \hat{\mu}_y}{\hat{\sigma}_x} \end{aligned}$$

where the values of $\hat{\mu}_x, \hat{\sigma}_x, \hat{\mu}_y$ and $\hat{\sigma}_y$ are given in Table 5.2.

Fitting the marginals and the copula

For the MLE and IFM methodologies, the Frank copula was fitted to the "standardised" log returns of Anglo and the TOP 40 whose marginals were given by the Student's t distribution with ν_x and ν_y degrees of freedom respectively. For the CML methodology, the Frank copula was fitted to the log returns of Anglo and the TOP 40, whose marginals were assumed to be empirically distributed. The three estimation procedures are outlined as follows:

1. The MLE estimates of the parameters $\hat{\psi}_{MLE} = \left(\hat{\nu}_{x(MLE)}, \hat{\nu}_{y(MLE)}, \hat{\theta}_{MLE}\right)'$ were obtained by a global maximisation of the log-likelihood function given by $\sum_{t=1}^{1246} \log\left[c\left(F\left(x_{(adj)t};\nu_x\right), G\left(y_{(adj)t};\nu_y\right);\theta\right) f\left(x_{(adj)t};\nu_x\right) g\left(y_{(adj)t};\nu_y\right)\right]$ where $F(.)$ and $G(.)$ are the Student's t distribution functions for the log returns of Anglo and the TOP 40 respectively.

2. The IFM estimates of the parameters, $\hat{\psi}_{IFM} = \left(\hat{\nu}_{x(IFM)}, \hat{\nu}_{y(IFM)}, \hat{\theta}_{IFM}\right)'$, were obtained by first separately fitting a Student's t distribution to the data $x_{(adj)t}$ and $y_{(adj)t}$, $t = 1, \ldots, 1246$, and specifically by finding the maximum likelihood estimates of their degrees of freedom as $\hat{\nu}_{x(IFM)}$ and $\hat{\nu}_{y(IFM)}$ respectively, which were taken to be unknown. These estimates were then plugged into the log-likelihood function, $\sum_{t=1}^{1246} \log c\left(F\left(x_{(adj)t};\hat{\nu}_{x(IFM)}\right), G\left(y_{(adj)t};\hat{\nu}_{y(IFM)}\right);\theta\right)$ and the latter was optimised with respect to θ to yield the IFM estimate of the copula parameter.

3. The CML estimate of the parameter of the copula, $\hat{\theta}_{CML}$, was obtained by firstly computing the empirical distributions of the data x_t and y_t, $t = 1, \ldots, 1246$, given by $\tilde{F}(x_t)$ and $\tilde{G}(x_t)$ respectively and then plugging those into the log-likelihood function, $\sum_{t=1}^{1246} \log[c\{\tilde{F}(x_t), \tilde{G}(y_t)\};\theta]$ and maximising the latter with respect to θ.

The parameter estimates of the marginals and the copula obtained from the MLE, IFM and CML estimation methods are shown in Table 5.3. Note that the parameter estimates computed using the three estimation methods are very similar specifically the MLE and the IFM parameter estimates.

	MLE	IFM	CML
$\hat{\nu}_x$	9.694	9.734	-
$\hat{\nu}_y$	6.466	6.543	-
$\hat{\theta}$	8.813	8.808	8.845

Table 5.3: The MLE, IFM and CML parameter estimates.

5.3.2 Estimation of VaR and expected shortfall

The VaR and the expected shortfall of the two-asset portfolio were computed, at the 95% confidence level, by using the copula-based Monte Carlo approach as discussed in Section 4.3. The parameter estimates obtained from the MLE, IFM and CML methodologies were used to estimate the two risk measures in a stepwise procedure as follows:

1. Simulate 1000 pairs of log returns of Anglo and the TOP 40, that is (x_i, y_i), $i = 1, \ldots, 1000$, as follows:

   (i) Using the MLE and IFM estimates of the copula parameter, $\hat{\theta}_{MLE}$ and $\hat{\theta}_{IFM}$, two data sets consisting of 1000 pairs (u_i, v_i), $i = 1, \ldots, 1000$, were generated from the Frank copula using the MATLAB `copularnd.m` routine. The pairs of adjusted log returns of Anglo and the TOP 40, $(x_{(adj)i}, y_{(adj)i})$, were then computed. That is, the adjusted log returns were computed as $x_{(adj)i} = F^{-1}(u_i)$ and $y_{(adj)i} = G^{-1}(v_i)$, $i = 1, \ldots, 1000$, where $F^{-1}(.)$ and $G^{-1}(.)$ are the inverses of the Student's t distribution function, using the corresponding degrees of freedom obtained from the MLE and IFM methodologies. The pairs of "standardised" log returns, $(x_{(adj)i}, y_{(adj)i})$, were adjusted back into the pairs of log returns, (x_i, y_i), as follows:

   $$x_i = (x_{(adj)i} \times \hat{\sigma}_x) + \hat{\mu}_x$$
   $$y_i = (y_{(adj)i} \times \hat{\sigma}_y) + \hat{\mu}_y$$

   where the values of the parameter estimates $\hat{\sigma}_x, \hat{\mu}_x, \sigma_y$ and $\hat{\mu}_y$ are given in Table 5.2.

   (ii) The CML estimate of the copula parameter, $\hat{\theta}_{CML}$, was used to generate 1000 pairs (u_i, v_i), $i = 1, \ldots, 1000$, from the Frank copula. To obtain the pairs of log returns, (x_i, y_i), the inverse of the empirical distribution was used, that is $x_i = \tilde{F}^{-1}(u_i)$ and $y_i = \tilde{G}^{-1}(v_i)$, $i = 1, \ldots, 1000$, where $\tilde{F}^{-1}(.)$ and $\tilde{G}^{-1}(.)$ are the inverses of the empirical distribution of the log returns of Anglo and the TOP 40 respectively.

2. For each of the three simulated data sets of log returns of Anglo and the TOP 40,the portfolio return was calculated as follows:

   $$z_i = 0.5x_i + 0.5y_i, \text{ for } i = 1, \ldots, 1000.$$

3. The three estimates of the 95% VaR were obtained as the 5%-quantile of the corresponding profit and loss distribution, that is the empirical distribution of z_t, of the portfolio. The 95% ES estimates were computed as the average of the losses that fall below the corresponding VaR estimates.

The MLE, IFM and CML estimates of the one-day ahead 95% VaR of the two-asset portfolio denoted by $\hat{V}_{(MLE)1247}$, $\hat{V}_{(IFM)1247}$ and $\hat{V}_{(CML)1247}$, obtained using 1246 observations of log returns, are -0.0270, -0.0237 and -0.0241 respectively. The corresponding MLE, IFM and CML estimates of the one-day ahead expected shortfall denoted by $\hat{E}_{(MLE)1247}$, $\hat{E}_{(IFM)1247}$ and $\hat{E}_{(CML)1247}$, computed at the 95% confidence level, are -0.0331, -0.0309 and -0.0320 respectively.

5.3.3 Properties of the VaR and the ES estimates

The bootstrapping technique was used to investigate the properties of the VaR and the ES estimates, specifically by generating B bootstrap samples of VaR and ES estimates. This was done by sampling $N = 1246$ pairs of log returns of Anglo and the TOP 40 with replacement from the pairs of observed values (x_t, y_t), $t = 1, \ldots, 1246$. Then the VaR and the ES of this bootstrapped data set were estimated by following the procedures described in Sections 5.3.1 and 5.3.2. This was repeated for $B = 500$ such bootstrapped data sets to yield the bootstrapped samples of VaR estimates and ES estimates.

Some descriptive statistics of the bootstrapped samples of the VaR and ES estimates, as defined in Section 4.4.3, are presented in Table 5.4. It can be observed that the means of the bootstrapped samples obtained from the three estimation methods are all similar. The VaR and ES estimates computed by the MLE methodology vary the least, as reflected by the standard errors and the coefficients of variation. All VaR and ES estimates generated by the three estimation methods are slightly negatively skewed, as shown in their histograms in Figure 5.5 and in the skewness coefficients in Table 5.4. The coefficients of kurtosis are close to 3. The Jarque-Bera test of normality does not reject the null hypothesis that the VaR and the ES estimates are normally distributed for all estimates except for $\hat{E}_{(MLE)1247}$ for which the test statistic, displayed in Table 5.1, exceeds the critical value of 5.8581 at the 5% significance level. The biases of all the VaR and ES estimates are small relative to the estimates.

5.4 Computing the VaR bounds

Using the Williamson and Downs (1990) method as outlined in Section 3.5, the VaR bounds of the two-asset portfolio at time $t = 1246$ were also computed, at the 95% confidence level, using the 1246 observations of log returns. The bounds provide a worst-case scenario and were thus used as a benchmark to compare the confidence intervals of the VaR estimates obtained using the bootstrapping technique.

The VaR bounds were computed by either assuming that the marginal distributions of the "standardised" log returns of Anglo, $x_{(adj)t}$, and those of the TOP 40, $y_{(adj)t}$,

		$\hat{V}_{(MLE)1247}$	$\hat{V}_{(IFM)1247}$	$\hat{V}_{(CML)1247}$
VaR	Mean	-0.024668	-0.024607	-0.024597
	Standard error	0.000056	0.000058	0.000061
	Coefficient of variation	-0.050745	-0.052875	-0.055164
	Skewness	-0.031015	-0.107570	-0.113076
	Kurtosis	2.886979	2.934120	3.108366
	Jarque-Bera statistic	0.346280	1.054699	1.310173
	Bias	-0.002306	0.000894	0.000476
	95% CI	(-0.027145,-0.022199)	(-0.027106,-0.022147)	(-0.027318,-0.021914)
		$\hat{E}_{(MLE)1247}$	$\hat{E}_{(IFM)1247}$	$\hat{E}_{(CML)1247}$
ES	Mean	-0.031507	-0.031361	-0.031662
	Standard error	0.000075	0.000077	0.000085
	Coefficient of variation	-0.053420	-0.054987	-0.060224
	Skewness	-0.309091	-0.209901	-0.228034
	Kurtosis	3.067769	2.920765	3.142652
	Jarque-Bera Statistic	8.057105	3.802330	4.757267
	Bias	-0.001574	0.000463	-0.000339
	95% CI	(-0.034880,-0.028548)	(-0.034887,-0.028208)	(-0.035627,-0.028084)

Table 5.4: The descriptive statistics of the bootstrapped sample of the VaR and ES estimates.

$t = 1, \ldots, 1246$, are Student's t distributed or by assuming that the log returns were empirically distributed. For the parametric estimates of the VaR bounds, the MLE and IFM estimates of the degrees of freedom of the "standardised" log returns of Anglo and the TOP 40 were used and the resulting VaR bounds are referred to as the MLE and the IFM bounds respectively. For the MLE bounds, the estimated degrees of freedom of the marginals of $x_{(adj)t}$ and $y_{(adj)t}$ are given by $\hat{\nu}_{x(MLE)}$ and $\hat{\nu}_{y(MLE)}$ respectively. For the IFM bounds, the estimated degrees of freedom of the marginals of $x_{(adj)t}$ and $y_{(adj)t}$ are given by $\hat{\nu}_{x(IFM)}$ and $\hat{\nu}_{y(IFM)}$ respectively. The MLE and the IFM bounds were then computed at the 95% confidence level by evaluating the inverse of the two distribution functions, $H_L(.)$ and $H_U(.)$, at $\alpha = 0.05$, as shown by Equation (3.10) and Equation (3.11) respectively. However, to "unstandardise" the adjusted log returns of Anglo and the TOP 40 back to their corresponding log returns, instead of evaluating $F^{-1}(.)$ and $G^{-1}(.)$ in Equation (3.10) and in Equation (3.11), $\left(F^{-1}_{x(adj)}(.)\hat{\sigma}_x + \hat{\mu}_x\right)$ and $\left(G^{-1}_{y(adj)}(.)\hat{\sigma}_y + \hat{\mu}_y\right)$ were evaluated. Furthermore, the variable N in Equation (3.10) and Equation (3.11) was set to 10^5 such that to evaluate $H_L^{-1}\left(\frac{i+1}{N}\right)$ as $H_L^{-1}(0.05)$, $i = 4999$ and to evaluate $H_U^{-1}\left(\frac{i}{N}\right)$ as $H_U^{-1}(0.05)$, $i = 5000$. The values of the parameter estimates $\hat{\mu}_x, \hat{\sigma}_x, \hat{\mu}_y$ and $\hat{\sigma}_y$ are given in Table 5.2 and $F^{-1}_{x(adj)}(.)$ and $G^{-1}_{y(adj)}(.)$ are the inverse of the Student's t distribution of the "standardised" log returns of Anglo and the TOP 40.

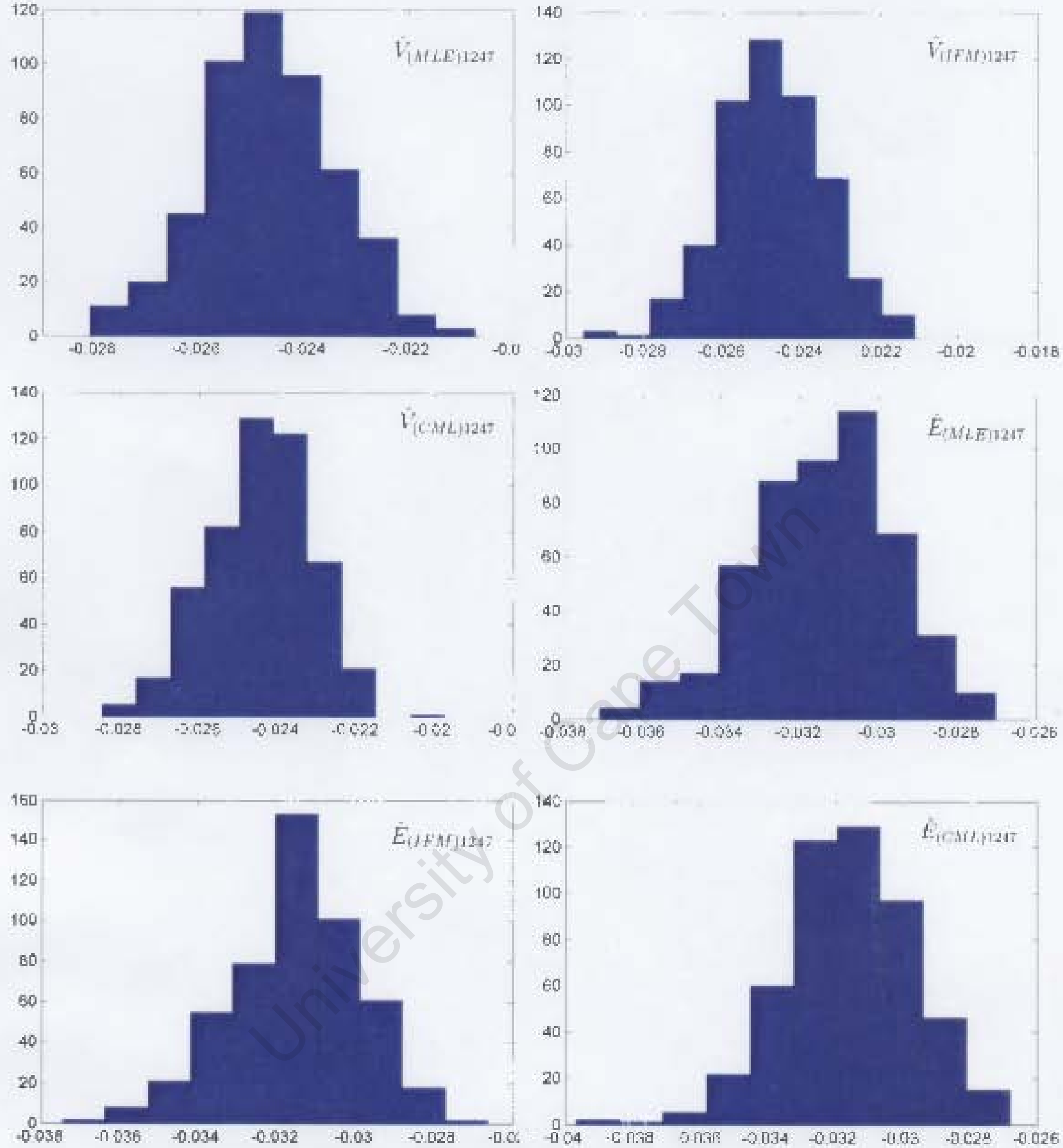

Figure 5.5: Histograms of the bootstrapped samples of the VaR and ES estimates obtained from the MLE, IFM and CML methods.

The VaR bounds which were computed by assuming that the log returns of Anglo and the TOP 40 are empirically distributed are referred to as the CML bounds. The CML bounds were computed at the 95% confidence level by evaluating the inverse of $H_L(.)$ and $H_U(.)$ at $\alpha = 0.05$, using the same values of i and N as specified above. However, instead of evaluating $F^{-1}(.)$ and $G^{-1}(.)$ in Equation (3.10) and in Equation (3.11) respectively, the inverse of the empirical distribution of log returns of Anglo and the TOP 40, that is $\hat{F}_x^{-1}(.)$ and $\hat{F}_y^{-1}(.)$ respectively, were evaluated.

The MLE, IFM and CML bounds which were estimated at the 95% confidence level are presented in Table 5.5. Note that the MLE and the IFM bounds are very similar

	VaR Bounds
MLE	(-0.032001,-0.002587)
IFM	(-0.031953,-0.002599)
CML	(-0.031445,-0.002501)

Table 5.5: The MLE, IFM and CML bounds which were estimated at the 95% confidence level.

and that the 95% confidence intervals of $\hat{V}_{(MLE)1247}$, $\hat{V}_{(IFM)1247}$ and $\hat{V}_{(CML)1247}$ fall within the corresponding MLE, IFM and CML bounds.

5.5 Backtesting the VaR estimates

Once the one-day ahead VaR was estimated at the 95% confidence level, its performance was evaluated using the backtesting technique. In this study, a rolling window of 1000 observations was used to estimate the one-day ahead VaR, V_{T^*+1}, where $T^* = 1000, \ldots, 1245$. The VaR estimates, $\hat{V}_{T^*+1}$, were then compared with the actual profit or loss that the portfolio incurred at time $(T^* + 1)$, which was obtained as follows:

$$z_{T^*+1} = 0.5x_{T^*+1} + 0.5y_{T^*+1}, \text{ for } T^* = 1000, \ldots, 1245.$$

The performance of the expected shortfall estimate cannot be evaluated using the backtesting technique since comparing the mean of the losses that fall below the VaR estimate to the actual loss cannot be interpreted in a meaningful way.

It was assumed that the Frank copula is the best-fitting copula and the marginals of the log returns of Anglo and the TOP 40 were the t-location scale distribution for any moving window of 1000 observations in the sample in order to compute $\hat{V}_{T^*+1}$ for $T^* = 1000, \ldots, 1245$. Therefore for a given T^*:

1. A moving window of 1000 observations was used to estimate V_{T^*+1} using the MLE, IFM and CML estimation methods as discussed in Sections 5.3.1 and 5.3.2. That is, the pairs of log returns of Anglo and the TOP 40, (x_t, y_t), $t = T^* - 999, \ldots, T^*$ were used to estimate the parameters of the copula and the marginals, which were then used to compute $\hat{V}_{T^*+1}$. Therefore for each estimation method, there is a total of 246 VaR estimates, $\hat{V}_{1001}$ to $\hat{V}_{1246}$.

2. Once the estimate, $\hat{V}_{T^*+1}$, was computed, a check was performed to see if the observed portfolio return was less than the VaR estimate, that is if $z_{T^*+1} < \hat{V}_{T^*+1}$.

This stepwise procedure was repeated for $T^* = 1000, \ldots, 1245$, and for each value of T^* the percentage of observed losses which exceeded, in absolute terms, the VaR estimates obtained from the three estimation methods were computed. The percentage of exceedance of the VaR estimates obtained from the MLE, IFM and CML methodologies are shown in Table 5.6.

	MLE	IFM	CML
Exceedance Level	10.204%	11.020%	11.429%

Table 5.6: The exceedance level of the VaR estimates obtained from the MLE , IFM and CML methods.

The results obtained are disappointing as the percentage of exceedance lies in the region of approximately 10% to 11% rather than close to the 5% nominal level. However, it can be noted that the VaR estimates obtained from the MLE method yielded the lowest exceedance level, which is closest to the 5% nominal level, followed by the VaR estimates obtained from the IFM method, followed by those obtained from the CML method. Possible reasons as to why the exceedance levels are not close to 5% are that the size of the rolling window is not big enough or that the Frank copula is not a good fit for the rolling window. Another reason could be that the number of simulations $d = 1000$ of log returns of Anglo and the TOP 40 are not enough to yield accurate estimates of the VaR and ES.

Chapter 6

Conclusions

Aims achieved

In this mini-dissertation, the background to copulas and their suitability as a measure of dependence have been explored and summarised. Common families of copulas have been introduced together with three methods of estimating copulas, namely maximum likelihood estimation (MLE), inference function for margins (IFM) and canonical maximum likelihood (CML). Furthermore, a procedure for selecting the best-fitting Archimedean copula has been reviewed.

Two popular measures of risk, namely Value-at-Risk (VaR) and expected shortfall (ES) have been introduced. The most popular methods of computing these risk measures and the limitations of the risk measures have been noted. Specifically, the computation of Value-at-Risk and expected shortfall using copulas has been explored as well as the computation of the VaR bounds.

A simulation study has been performed to compare the properties and performance of the VaR and ES estimates obtained using the three estimation methods of the copula parameter. In the simulation study, two copulas have been considered, namely the Frank and the Gaussian copulas with normal and Student's t distributed marginals. The results show that the VaR and ES estimates obtained from the three estimation methods are very similar, in particular the estimates obtained from the MLE and IFM estimation methods. In terms of performance, the MLE estimation method is the preferred estimation method. However the biases and MSE of the VaR and ES estimates obtained from the IFM method are almost identical and in some cases smaller than those of the estimates obtained from the MLE method. The CML method also yielded satisfactory results and the estimates were easily computed and were similar to the MLE estimates. However the simulation results also showed that using the CML method may come at the cost of the performance of the estimates.

A real data set has been used to illustrate the copula-based approach of estimating the VaR and the ES of a portfolio of two equally weighted risky assets, namely Anglo and the Top 40 Index. The properties of the VaR and ES estimates have been investigated and the VaR bounds of the portfolio have been computed. The backtesting technique has been used to evaluate the performance of the VaR estimates generated from the fitted models. However, the results obtained were somewhat disappointing in that the percentage of exceedance levels of the VaR estimates were not close to the 5% nominal level.

Recommendations and future work

The following recommendations can be made:

1. Given that the MLE methodology is very computationally intensive and the attendant nonlinear optimisation routine can easily run into convergence problems, the IFM methodology provides a robust alterative to the MLE.

2. In practice, the marginals of the returns of assets are generally unknown and as a result, the VaR and ES estimates obtained from the MLE and IFM methods can suffer from model misspecification. Furthermore, as demonstrated by the real data set, it can be computationally difficult to fit the marginals to the data. Therefore, in practice the CML estimation method provides a robust alternative to the MLE and IFM methodologies, but it may come at the cost of the performance of the VaR and ES estimates.

Concerning future work, the following suggestions which are beyond the scope of this mini-dissertation can be made:

1. Further research needs to be done on the backtesting procedure. For example, the number of simulations for estimating the VaR or the size of the rolling window can be increased to investigate whether there is an improvement in the results.

2. The coverage of the confidence intervals provided by the bootstrap samples of the MLE, IFM and CML estimates can be computed to further investigate the performance of the three estimation methods.

3. The robustness of the VaR and ES estimates to misspecification of both the marginals and the copula function can be investigated.

Bibliography

Alexander, C. (2001). *Market Models: A Guide to Financial Data Analysis*, John Wiley & Sons Ltd.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (2002). *Coherent Measures of Risk. In: Dempster M.A.H, Risk Management Value at Risk and Beyond*, Cambridge University Press, 145-175.

Basel-II (2005). International convergence of capital measurement and capital standards: A revised framework. http://www.bis.org, *Technical Report* .

Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula Methods in Finance*, John Wiley & Sons Ltd.

Dowd, K. (2005a). Copulas and coherence, *The Journal of Portfolio Management, 32, 123-127* .

Dowd, K. (2005b). *Measuring Market Risk*, Second Edition, John Wiley & Sons Ltd.

Dowd, K. and Blake, D. (2006). After VaR: The theory, estimation, and insurance applications of quantile-based risk measures, *The Journal of Risk and Insurance, 73, 193-229* .

Embrechts, P., Höing, A. and Juri, A. (2003). Using copulae to bound the value-at-risk for funtions of dependent risks, *Finance and Stochastics, 7, 145-167* .

Embrechts, P., Lindskog, F. and McNeil, A. (2001). Modelling dependence with copulas and applications to risk management, *Department of Mathematics ETHZ, CH-8092 Zurich* .

Embrechts, P., McNeil, A. J. and Straumann, D. (2002). *Correlation and Dependence in Risk Management: Properties and Pitfalls. In: Dempster M.A.H, Risk Management Value at Risk and Beyond*, Cambridge University Press, 176-223.

Frees, E. W. and Valdez, E. A. (1997). Understanding relationships using copulas, *North American Actuarial Journal, 2, 1-25* .

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas, *Journal of the American Statistical Association, 88, 1034-1043* .

Giacomini, E. and Härdle, W. (2005). Value-at-risk calculations with time varying copulae, *Humboldt-Universität zu Berlin* .

Hürlimann, W. (2004). Fitting bivariate cumulative returns with copulas, *Computational Statistics & Data Analysis, 45, 355-372* .

Joe, H. (1997). *Multivariate Models and Dependence Concepts*, Chapman & Hall.

J.P.Morgan (1995). Riskmetrics technical manual, J.P. Morgan Bank, New York, *Technical Report* .

Kim, G., Silvapulle, M. J. and Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas, *Computational Statistics & Data Analysis, 51, 3418-3432* .

Laubscher, J. (2007). The economic implications of financial market volatility, *Technical Report* .

Luciano, E. and Marena, M. (2002). Portfolio value at risk bounds, *International Transactions in Operations Research, 9, 629-641* .

Luciano, E. and Marena, M. (2003). Copulae as a new tool in financial modelling, *Operational Research: An International Journal, 2, 139-155* .

Nelsen, R. (2006). *An Introduction to Copulas*, Second Edition, Springer.

Patton, A. J. (2006a). Copula-based models for financial time series, *Technical Report, London School of Economics* .

Patton, A. J. (2006b). Modelling assymetric exchange rate dependence, *International Economic Review, 47, 527-556* .

Roch, O. and Alegre, A. (2006). Testing the bivariate distribution of daily equity returns using copulas. an application to the spanish stock market, *Computational Statistics & Data Analysis, 51, 1312-1329* .

Williamson, R. C. and Downs, T. (1990). Probabilistic arithmetic. I. numerical methods for calculating convolutions and dependency bounds, *International Journal of Approximate Reasoning, 4, 89-158* .