A NEW ONE PARAMETER FAMILY OF ARCHIMEDEAN COPULA AND ITS EXTENSIONS

AZAM PIRMORADIAN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2013

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Azam Pirmoradian** (passport No.: **L9523916**)

Registration / Matrix No.: **SHB090006**

Name of Degrees: **Doctor of Philosophy (Ph.D)**

Title of Thesis ("this Work"):

A new one parameter family of Archimedean copula and its extensions

Field of Study: **Applied Statistics**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name:

Designation:

Abstrak

Untuk mencirikan pergantungan risiko yang ekstrem, konsep kebersandaran ekor bagi fungsi taburan bivariat telah diperkenalkan. Kopula Gaussian, sebagai contoh, tidak mempunyai kebersandaran atas maupun bawah - ia menunjukkan ketaksandaran asimptot tanpa mengira korelasi yang mungkin wujud antara pembolehubah. Dalam erti kata lain, nilai ekstrem dalam pembolehubah yang berbeza berlaku secara bebas walaupun terdapat korelasi yang tinggi antara kedua pembolehubah. Konsep kopula bertujuan untuk mengatasi masalah kebersandaran ekor.

Kopula Archimedean membentuk keluarga kopula yang penting yang mempunyai bentuk mudah dengan sifat-sifat seperti associability dan memiliki pelbagai struktur kebersandaran. Khususnya, kopula Archimedean untuk satu set data bivariat boleh dibina dengan mudah oleh fungsi penjana. Secara unik, penjana menentukan kopula Archimedean dan pilihan penjana yang berbeza menghasilkan banyak keluarga kopula. Akibatnya, sifat kebersandaran kopula ini adalah agak mudah untuk dibangunkan kerana mereka mengurangkan kepada hartanah analisis penjana. Kebanyakkan kopula Archimedean dengan keluarga penjana satu parameter, kopula Gumbel atau Clayton sebagai contoh, dapat menjelaskan samada kebersandaran atas atau bawah, tetapi tidak kedua-duanya.

Pembaharuan dalam tesis ini adalah pembinaan sebuah keluarga yang baru Archimedean kopula dengan mengeksploitasi sifat fungsi trigonometri, dengan kelebihan tambahan yang mempunyai hanya satu parameter. Lima kopula trigonometri dibina, dinamakan Cot-, CotII, CSC, CscII dan CscIII-kopula. Hasil dapatan kami menunjukkan kesemua kopula mempunyai sifat kebersandaran positif yang di analisis dengan mempertimbangkan sifat penuaan setiap kopula. Dari segi sifat kebersandaran yang diukur melalui kebersandaran ekor dan Kendall tau, Cot-kopula dan Csc-kopula berupaya untuk menguasai kedua-dua kebersandaran ekor bagi data simetrik dan tak-simetrik. Hasil dapatan juga menunjukkan bahawa Cot-kopula adalah lebih tepat apabila kebersandaran ekor bawah adalah lebih berat daripada kebersandaran ekor atas, dan keadaan yang sebaliknya untuk CSC-kopula. Tidak seperti keluarga ke-12 Archimedean kopula dengan kedua-dua kebersandaran ekor, Cot-dan Csc-kopula mempunyai liputan kebersandaran yang luas. Kelebihan Csc-kopula berbanding Cot-kopula adalah keupayaannya dalam menerangkan hampir keseluruhan kebersandaran dalam [0, 1]. Kami juga melanjutkan kopula trigonometri bivariat kepada kopula multivariat melalui struktur *vine* (menjalar). Dalam perluasan multivariat, Cot-kopula dan Csc-kopula telah dipilih untuk digunakan sebagai blok binaan dalam fungsi pengagihan multivariat. Kelebihan kopula tersebut dalam struktur *vine* adalah disebabkan bilangan parameter yang sedikit yang dapat mengurangkan kesilapan anggaran terutamanya dalam dimensi yang tinggi. Akhirnya kami menunjukkan kaedah yang dibangunkan melalui simulasi dan data-data kewangan dan hidrologi. Dalam aplikasi kewangan, keputusan menunjukkan kelebihan menggunakan Cot-dan Csc-kopula dalam meguasai kebersandaran ekor yang kukuh antara indeks Eropah. Kami berjaya membentuk kebersandaran multivariat antara pasaran Asia melalui struktur C-*vine* kerana terdapat kebersandaran kesemua pasaran ke indeks Singapura.

Abstract

In order to characterize the dependence of extreme risk, the concept of tail dependence for bivariate distribution functions was introduced. The Gaussian copula, for example, does not have upper or lower tail dependence - it shows asymptotic independence regardless of the correlation that may exist between the variables. In other words, the extreme values in different variables occur independently even if there is a high correlation between these variables. The concept of copula aims at overcoming the tail dependence problem.

The Archimedean copulas form an important family of copulas which have a simple form with properties such as associability and possess a variety of dependence structures. Specifically, the Archimedean copula for a bivariate data set can easily be constructed by a generator function. The generator uniquely determines an Archimedean copula and different choices of generator yield many families of copulas. As a consequence, many dependence properties of such copulas are relatively easy to establish because they reduce to analytical properties of the generator. Most of the Archimedean copulas with one-parameter families of generators, the Gumbel or Clayton copula for example, can explain either the upper or lower tail dependence but not both.

The novelty of this thesis is to construct a new Archimedean family of copula by exploiting the properties of trigonometric functions, with an added advantage of having only one parameter. Five trigonometric copulas are constructed, namely the Cot-copula, CotII-copula, Csc-copula, CscII-copula and CscIII-copula. Our results show that these copulas have positive dependence properties which were analyzed by considering the aging properties of the respective copula. In terms of dependence properties measured

by tail dependence and Kendall's tau, the Cot-copula and Csc-copula are able to capture both tail dependences in symmetric and asymmetric data. Our result also shows that Cot-copula is more accurate when the lower tail dependence is heavier than the upper tail dependence, and the opposite applies to Csc-copula. Unlike the 12^{th} family of Archimedean copula with both tail dependences, the Cot- and Csc-copula have wider dependence coverage. The advantage of Csc-copula rather than Cot-copula is its ability in capturing almost complete dependence in [0, 1]. We also extend the bivariate trigonometric copula to multivariate copula via the vine structure. For multivariate extension, the Cot-copula and Csc-copula are selected as building blocks in multivariate distribution function. The advantage of these copulas in vine structure is due to the small number of unknown parameters which reduce the estimation error especially in high dimension. Finally we demonstrate the methods developed in this study through simulation and real financial and hydrological data. In financial applications, the results show the advantage of using Cot- and Csc-copula in capturing strong tail dependences between the European market indexes. We are able to construct the multivariate dependence between the Asian markets via C-vine structure since these markets are dependent on the Singapore market index.

Acknowledgments

I would like to express my deep appreciation and gratitude to my supervisor, Professor Dr. Nor Aishah Hamzah, for the patient guidance and mentorship she provided to me, all the way from when I was first considering applying to the Ph.D program in the ISM, through to completion of this degree.

My many thanks to the Institute of Mathematical Sciences (ISM), Universiti Malaya especially to head of institute, Prof. Dr. Mohd Bin Omar, and former head, Prof. Dr. Ong Seng Huat, for their co-operating.

I would also like to thank Professor Saad Al-Jassabi , Associate Professor Dr. Ahmed Shamiri and Dr. Siraj Uddin for their friendly guidance, thought provoking suggestions, and the support that each of them offered to me over the years.

I gratefully acknowledge the support and generosity of the Bright Spark Units and University of Malaya for their financial support in this research, without which the present study could not have been completed.

I would be remiss if I didn’t acknowledge the innumerable sacrifices made by my parents, Akbar and Effat, all these years. I also thank to my spouse, Hossein, for all his supports which start even before this final degree and I am truly fortunate to have had his support in my life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIC	Akaike Information Criterion
AIC	Akaike Information Criterion
ARCH	Autoregressive Conditional Heteroskedasticity
ARCH	Autoregressive Conditional Heteroskedasticity
CDF	Cumulative Distribution Functions
CDF	Cumulative Distribution Functions
CG	Mixed Clayton-Gumbel
C-vine	Canonical Vine
CvM	Cramer-von Mises
DFR	Decreasing Failure Rate
DFRA	Decreasing Failure Rate in Average
EGARCH	Exponential General Autoregressive Conditional Heteroskedastic
EV	Extreme Values
FCHI	French CAC 40 Index
FTSE	British Ftse 100 Index
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GDAXI	German DAX Index
GOF	Goodness of Fit Test
IFM	Inference Function for Margins
IFR	Increasing Failure Rate
IFRA	Increasing Failure Rate In Average
KLCI	Kuala Lumpur Composite Index
KS	Kolmogorov-Smirnov

LTD	Left-Tail Decreasing
LTI	Left Tail Increasing
MLE	Maximum Likelihood Estimation
NBU	New Better than Used
NKD	Negative K-Dependence
NQD	Negative Quadrant Dependence
NWU	New Worse than Used
PKD	Positive K-Dependent
PML	Pseudo Maximum Likelihood
PQD	Positive Quadrant Dependence
PUOD	Positively Upper Orthant Dependent
RTI	Right-Tail Increasing
SD	Stochastically Decreasing
SI	Stochastically Increasing
SP500	Standard And Poor Index
SSE	Chinese Composite Index
STI	Strait Times Index

Chapter 1: Introduction

This chapter provides an overview of the mandate of the thesis. We begin with motivation and scope of the research together with a brief history of dependence, copula and its application which is the key issue in this research. The objective of the research is delineated in three sections: problem statement, research objective and significant contribution of the study. Finally, different structures of the thesis are outlined, each with a brief description.

1.1 Motivation

Dependence plays an important role in most of the subjects. This is due to the fact that the occurrence of every event may be related to other variables. In financial risk models for example, whether for market or credit, risks are inherently multivariate (McNeil, Frey, & Embrechts, 2005). In measuring risk, an accurate model of dependence is essential to compute the value at risk, expected shortfall and financial contagion. On the other hand, portfolio management deals with the dependence between international financial markets, different classes of assets and currencies (Genest, Gendron, & Bourdeau-Brien, 2009). The importance of dependence structure between random variables in hydrology is also significant; discovering the dependence structure of rainfall variables is required in many water resources projects. A good understanding of the dependence between random variables in various fields of interest allows for proper risk measurement. Thus, the concept of dependence is important and must be clearly understood to both academics and practitioners.

The classical approach to measuring dependence in financial or hydrology studies is based on the multivariate normal distributions or more generally, the elliptical distributions. The mean, covariance matrix and the type of marginal distributions used, are the elements that uniquely determine an elliptical distribution. The advantage of using such distributions lies in its simplicity and analytical manageability with dependence being determined by its correlation matrix. However, empirical research in most areas, including hydrology and finance, suggest that the use of multivariate normal distributions is no longer adequate; the statistical analysis of the distribution of individual asset returns frequently finds fat tails, skewness and other non-normal features which lead to the underestimation of this dependence measure (for example see (Ang & Bekaert, 2002; Ang & Chen, 2002; Bae, Karolyi, & Stulz, 2003; Longin & Solnik, 2001)).

The dependence structure of joint distributions can be described by copula. A copula is a function which binds or 'couples' univariate marginal distributions and the multivariate distribution. By allowing different marginal distributions and a dependence structure which is not solely determined by covariance, the copula is able to generate multivariate distributions with flexible marginals. This is indeed useful in application, because one is often interested in linear combinations of margins from possibly different distributions.

The Archimedean copulas form an important family of copulas. These copulas have a simple form and enjoy certain properties such as being associative and possessing a variety of dependence structures. The Archimedean copula can be specifically and easily constructed by a generator function for a bivariate data set. As a generator uniquely determines an Archimedean copula, different choices of generator yield many families of copulas (see Table 4.1 of Nelsen (Nelsen, 2006), for the list of one-parameter Archimedean copulas). As a result, since many dependence properties of these copulas reduce to analytical properties of the generator, they are fairly easy to establish; see for example, (Genest & MacKay, 1986a, 1986b; Joe, 1997; Muller & Scarsini, 2005) and (Nelsen, 2006). Furthermore, many parametric families of Archimedean copulas which have attractive stochastic features and result in statistically tractable models for continuous data have already been constructed. Some important applications of the Archimedean copulas can be found in studies of marketing, finance (Elizalde, 2006) and hydrology (AghaKouchak, Bárdossy, & Habib, 2010).

Extending bivariate Archimedean copulas to multivariate are not easy tasks. Every single member of this family should be studied separately. This problem can be solved by considering a structure of dependence. Vines are graphical structures that represent joint probability distributions (Kurowicka & Joe, 2010). A special case of vines, called regular vine, can be used successfully to model high-dimensional dependence together with copula. A regular vine is a special case for which all constraints are either two-dimensional or conditional two-dimensional. This structure can be combined with Archimedean copula in constructing multivariate distribution function.

Next, we will briefly cover the scope of the research and review the literature on dependence and copula.

1.2 History of Dependence and Copula

Verbal definition of independence has been made available since the eighteenth century. But the most fascinating definition, according to Keynes (Wilkinson, 1872), was introduced by (Boole, 1854) namely: “Two events are said to be independent when the probability of the happening of either of them is unaffected by our expectation of the occurrence of failure of the other." In 1997, Mosteller and Tukey emphasized on the importance of distinguishing between "dependence" and "exclusive dependence". They mentioned that "y being a dependent on x" means failure of independence, while exclusive dependence means that if x is given, the value of y follows a mathematical formula, $Y = f(X)$ (Mari, Kotz, & ebrary, 2001).

Correlation was started by Francis Galton who is often ascribed with the title of "father of correlation" (Galton, 1886). In 1892 Edgeworth changed the name of "coefficient of co-relation" to "coefficient of correlation" (Edgeworth, 1893) and later in 1986 Pearson derived the analytic product-moment formula, known as Pearson correlation (Chatterjee, 2003). The concept of correlation coefficient was widely accepted across a wide range of statistical fields. It is due to the fact that in social sciences the correlation analysis has been widely used to determine the relationship between the occurrences of economic or social events. Moreover, the role of the correlation within regression analysis is worthless (Dorey & Joubert, 2007; Mari et al., 2001).

There is a long history of restriction of the Pearson's correlation outside the Gaussian's (elliptical) framework. The most impressive paper was written by Embrecht which emphasizes on the misunderstanding and confusion about the correlation in risk management (Embrechts, McNeil, & Straumann, 1999, 2002). Moreover, under a strictly growing transformation of variables, the correlation would not be invariant (Ane & Kharoubi, 2003). Also, in case of a given marginal distribution of random variables, none of the linear correlations between -1 and 1 can be achieved through a suitable specification of the joint distribution(Ane & Kharoubi, 2003). Therefore, the lack of robustness associated with Pearson's correlation has given rise to many alternative measures of association (Mosteller & Tukey, 1977; Wilcox, 2005).

There are too many measures of associations to be compared separately. As Jogdeo notes (Jogdeo, 1982), the refractory nature of dependence can enjoy various forms, therefore, in order to control dependence, some specific assumptions should be made. We categorise the different measures of association in this study into four groups, namely concordance, measure of dependence, quadrant and tail dependence (Nelsen, 2006).

The definition of non-parametric measures Kendall's τ and Spearman's ρ are based on concordance and discordance measure. Once "large" values of one tend to be associated with "large" values of the other and "small" values of one with "small" values of the other, a pair of random variables is concordance. In terms of concordance, Kendall's τ (Kruskal, 1958) can be defined as the difference between the number of concordance pairs from discordance pairs divided by number of distinct pairs (Hollander & Wolfe, 1999; Lehmann & D'Abrera, 1975). Spearman's ρ , named after the English psychologist, Spearman, who suggested this measure in 1904 (Spearman, 1904). Spearman's ρ is defined to be proportional to the probability of concordance minus the probability of discordance for a pair of vectors with similar margins, of which one of the vectors enjoys the distribution function H, while the components of the other vector are independent (Kruskal, 1958; Nelsen, 2006).

Measuring monotonic dependence between random variables rather than linear dependence is the first advantage of concordance measure toward correlation. Moreover, monotonic dependence is invariant under increasing transformation of variables. In addition, it has the ability to handle any rank correlation in the interval $[-1, 1]$ for any arbitrary marginal (Ane & Kharoubi, 2003). However, a drawback for this measure is that zero concordance does not imply independence in random variables. There are many examples where there is a zero measure of concordance, but the random variables are not independent (Nelsen, 2006).

The non- parametric measure of association, namely "measure of dependence" is based on distance. According to Lancaster (Lancaster, 1982), dependence relates to the closeness of random variables with the independent and monotone dependence (Jogdeo, 1982; Lancaster, 1982; Nelsen, 2006; Schweizer & Wolff, 1981). More precisely, the affinity of joint distribution function in random variables will be measured by independent or monotone dependence's joint distribution functions.

Tail dependence on the other hand, measures the dependence between the variables in the upper-right quadrant and in the lower-left quadrant on the unit square. This dependence is similar to the dependence concept which is designed to describe how large (or small) values of one variable appear with large values of the other (Nelsen, 2006).

For a couple of random variables, dependence property can be considered a subset of a set of all joint distribution functions. For example, the property of independence corresponds to the subset of all members that have independent joint distribution function (Nelsen, 2006). This is similar to the monotone functional dependence which is related to the Frèchet bound distribution functions. Every dependence structure can be described by the joint distribution which lies between independence and monotone dependence. It is a clear point for advent of copula in dependence concept, since the copula give a more flexible way to construct distribution function (Wolff, 1980).

According to (Durante & Sempi, 2010) the history of copula started with the following problem proposed by Frèchet (Fréchet, 1951): Given the distribution functions F_1 and F_2 of two random variables X_1 and X_2 defined on the same probability space (Ω, f, P), what can be said about the set $\Gamma(F_1, F_2)$ of the bivariate distribution functions whose margins are F_1 and F_2? The set $\Gamma(F_1, F_2)$, Frèchet class, is not empty since the independent distribution function always belongs to $\Gamma(F_1, F_2)$; however, it was not clear which other elements of $\Gamma(F_1, F_2)$ existed.

The profoundest answer to this question was introduced in 1959 by Sklar (Sklar, 1959) with the notion of copula. This concept of copula was extended just in the framework of the theory of probabilistic metric space for about 15 years (Moore & Spruill, 1975; Schweizer & Sklar, 1983; Sklar, 1973).

The copula came into the framework of statistic with the work of Scheweizer and Wolff (Schweizer & Wolff, 1981). The concept was stabilized in this framework later by two reference books which were written by Joe and Nelsen (Joe, 1997; Nelsen, 2006).

Toward the end of 20th century, the discovery of the notion of copulas by researchers in several applied fields like finance increased its popularity. At the same time, the importance of this concept in constructing more flexible multivariate models was demonstrated in different fields like hydrology (Genest & Favre, 2007; Salvadori & De Michele, 2007).

1.3 Objectives of Study

Concept of the problem statment and the objective of study is discussed in this section.

1.3.1 Problem Statement

The concept of tail dependence for bivariate distribution functions was first introduced by Joe, so as to portray the dependence of extreme risk (Joe, 1997). This suggests that lower (upper) tail dependence between two variables (such as two asset returns) exists when the probability of joint negative (positive) extreme events is larger than what we would expect from the marginal distributions. For example, there is no upper or lower tail dependence in the Gaussian copula- it shows asymptotic independence irrespective of the correlation that may exist between the variables. In other words, despite the possibility of a high correlation between variables, the extreme values in different variables happen independently.

For the Archimedean copulas with one- parameter families of generators, the Gumbel copula for example, can explain the upper tail dependence but not the other. Another example of the Archimedean family is the Clayton copula C_θ which has a tendency to be independent between the variables once the parameter $\theta \to 0$. Its tails also show asymmetry, with strong lower tail dependence and relatively weak upper tail dependence. Meanwhile, the dependence in the tails of the Frank copula is relatively weak, which is an indication that this copula is appropriate for data that exhibit weak tail dependence. With the exception of the12th family, most Archimedean copulas introduced in Table 4.1 of Nelsen (Nelsen, 2006) cannot simultaneously explain both tail behaviours observed on financial markets. In order to obtain copulas with bivariate tail dependence measures, many authors construct new copulas which are convex linear combinations of two copulas; examples are Joe-Clayton (Joe, 1997), Gumbel-Clayton (Ane & Kharoubi, 2003) and many more.

Alternatively, the number of parameters in bivariate Archimedean copula is important, especially in the case of multidimensional extension via vine structure. Vines are structures which uses bivariate distribution to construct a multivariate distribution. Through this structure, the dependence properties of the multivariate copula inherit the bivariate ones. As such, if the bivariate copula contains several parameters then the multivariate copula will be more complex. To reduce such complexity, we try to build some bivariate copula with less number of parameters which carry some beneficial dependence properties.

1.3.2 Research Objective

In this thesis our research objectives are as follow:

1) **To document the copula theory in mathematical and statistical literature in the most beneficial way for our research.** This involves searching and collecting the copula theory in mathematics, statistics, finance and hydrology.
2) **To construct new one parameter family of Archimedean copula.** This involves introducing some new bivariate Archimedean copulas with a one-parameter family, which we refer to as trigonometric copulas.
3) **To determine the properties of dependence which are most useful in real word analysis.**
4) **To calculate the dependence properties of constructed copula.**

5) **To compare the performance of the proposed copula with those in the literature, for example Gumbel, Clayton and 12th family of Archimedean copula.** The performance measure will be based on dependence properties, dependence measures and goodness of fit.
6) **To establish multivariate copula according to constructed bivariate copula**. This involves establishing multivariate copula with trigonometric copula according to vine structure.
7) **To verify the properties with real data in finance subject**. This involves application of trigonometric copula in real word application to validate the theoretical part of our research.

1.3.3 Significant Contribution to the Subject

The significance of the study can be divided into two parts: theory and application. The theoretical part is divided into two main contributions:

Theory 1: In this study we first introduce some Archimedean copulas which are built on trigonometric functions. The importance of this family is due to dependence properties of these copulas. Some of them have flexible upper and lower tail dependences with a wide dependence coverage which forms the basic building blocks in multivariate copula.

Theory 2: We construct multivariate copula via vine structure by using trigonometric copula. The advantage of vine structure with trigonometric copula is that it can simultaneously capture the upper and lower tail dependences, with one parametric family. This is important because in multivariate dimension, the estimation error increases with dimensions.

The application contribution is as follow:

Applications: In most of the copula applications in finance and hydrology, bivariate copula are used to find the joint distribution between two random variables. However, in real situation, one event may be related to more than two variables. In such situations, applying multivariate copula instead of bivariate copula can provide more accurate information. To demonstrate the usefulness of the proposed multivariate copula, some datasets from finance are used.

1.4 Thesis Structure

The thesis is structured on six chapters which cover both the theory and application aspects of the research. The flowchart for the thesis is displayed in

Chapter 1 sets the context of the research and motivation, explores the significance of the research together with objectives and the structure of the thesis.

Chapter 2 is divided into two parts: the first part extensively reviews the literature concerning copula in three sections: the first section tracks the concept of copula in mathematic context with exact definition of copula. Some mathematical theories related to copula are also considered. The second section introduces the concept of copula which relates to statistical literatures. An overview of past research concerning estimation and goodness of fit method for copula is summarized. The third section explores different family of copulas where the major focus is on elliptical and Archimedean copula. This chapter ends with the development of trigonometric family of copula, which forms the basis for future chapters.

Chapter 3 focuses on the literature review on the dependence which is employed in the later part of this chapter, to investigate the dependence properties of trigonometric copula. To ease explanation, we divided the chapter into four sections. The first section provides the literature review on dependence concept while the second section studies the dependence properties for all trigonometric copulas. Section three provides an overview on different measures of dependence. Finally, section four computes dependence measures in trigonometric copula.

Chapter 4 focuses on concept of vine structure which is used to construct multivariate trigonometric copula. The chapter ends with a simulation study on multivariate copula in finance datasets.

Chapter 5 presents an analysis of data used in finance. In the first section, we illustrate some application via bivariate copula on seven indices from three continents, Asia, Europe and America. Then, we compare the ability of trigonometric copula in capturing tail dependence with existing one-parameter and two- parameter families of copula. Later and in the second section, we construct multivariate copula via vine structure on four dependence indices. The result of trigonometric multivariate copula is then compared to the optimal choice from the existing copulas.

Chapter 6 presents the conclusions of the research in terms of theory and application and suggests some future research direction.

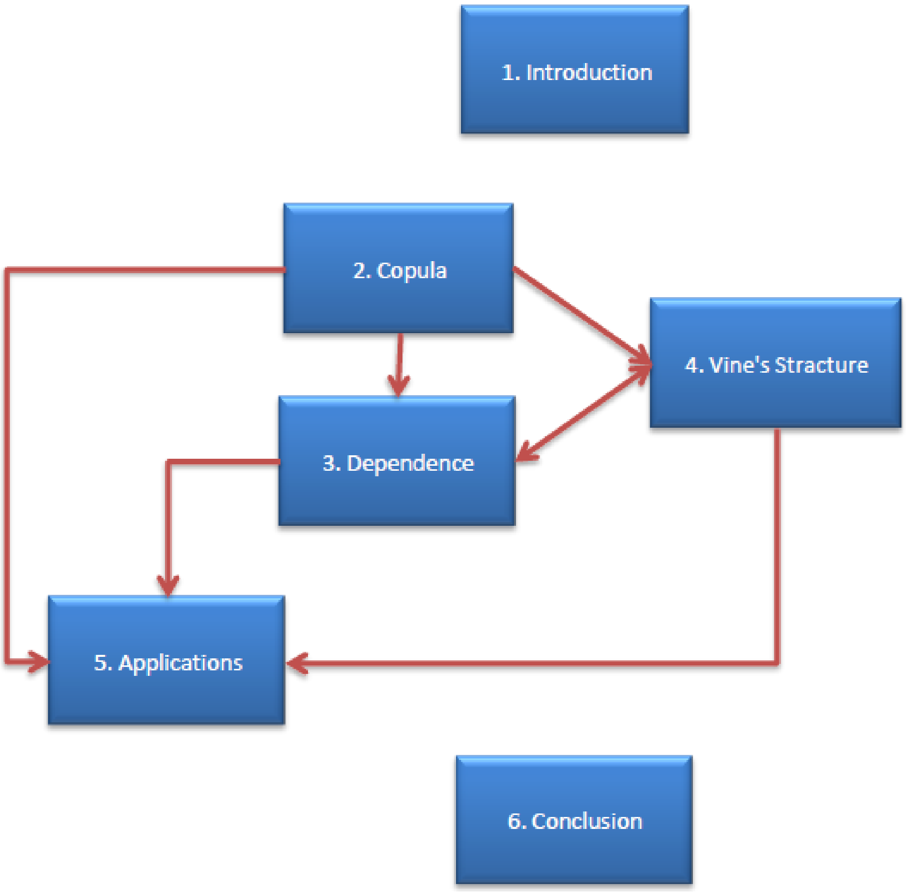


Figure 1.1: Structure of the thesis

Chapter 2: Copula

The main focus of this chapter is to introduce bivariate copula and their properties. It is also organized to introduce the proposed trigonometric copulas for the first time. The first section addresses the copula definition and the important theories in mathematics. The second section covers estimation and goodness of fit for copula. Specifically, we will focus on elliptical and Archimedean family of copulas including trigonometric copula as a sub-set of the Archimedean family.

2.1 The Copula and Its Properties

According to Oxford English Dictionary, the term "Copula" is in fact, a Latin word which means "to fast or fit." Technically, it describes the relation between two things, in our case, the marginal distributions.

To define a copula we start by clarifying the concept of copula in two dimensions. A pair of random variables, X and Y with respective cumulative distribution functions $F(x) = P(X \le x)$, $\forall x \in R$ and $G(y) = P(Y \le y)$, $\forall y \in R$, and a joint distribution function $H(x, y) = P(X \le x, Y \le y)$, $\forall x, y \in R$, on a common probability space are assumed. For some function $C : I^2 \to I$ each pair $(F(X), G(Y))$ in the unit plane I^2 corresponds to the number $z \in I$ given by the relation $z = H(x, y) = C(F(x), G(y))$. The definitions of copula of d random variables generalize the bivariate definition of copula presented above.

Definition 2. 1. A function $C : I^d \to I$ is a *d*-copula if the following properties hold:

(i) For every $j \in \{1,2,...,d\}$, $C(1,...,u_j,1,...,1) = u_j$

(ii) $\forall u_i \in [0,1]$, $C(u_1,...,u_n) = 0$ if at least one of the u_i is zero.

(iii) C is grounded and *d*-increasing.

From this definition, we can claim the fact that a copula is a multivariate distribution function with support in $[0,1]^d$ and uniform margins. The important part of this mathematical object is that they are useful for constructing multivariate distribution function with arbitrary marginals. The following theorem provides support for this statement.

Sklar's Theorem (Sklar, 1959)**:** Let F be a joint distribution function with continuous marginal distribution function F_i for $i = 1,...,n$. Then there exists a unique copula function C, such that:

$$F(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n)). \tag{2.1}$$

On the contrary, if C is a copula and F_i are marginal distribution functions, then F defined above is a joint distribution with margins F_i.

Corollary 1 (Nelsen, 2006): Let $F_i^{-1}(u_i)$ for $i = 1,...,n$ denote the generalized inverses of the uniform marginal distribution function u_i for $i = 1,...,n$. Then for every $(u_1,...,u_n)$ in the unit $n-$cube, there exists a unique copula $C:[0,1]^n$ such that $F\left(F_1^{-1}(u_1),...,F_n^{-1}(u_n)\right) = C(u_1,...,u_n)$, where $F_1^{-1}(u_1) = \inf\,[x_i : F_i(x_i) > u_i]$, for $i = 1,...,n$

.

Taking the derivatives with regards to $(u_1,...,u_n)$ and using the chain rule, the copula density is defined by

$$c(u_1,...,u_n) = \frac{\partial^n C(u_1,...,u_n)}{\partial u_1,...,u_n},$$

The joint probability distribution function may then be recovered as follows:

$$f(x_1,...,x_n) = \frac{\partial^n F(x_1,...,x_n)}{\partial x_1,...,x_n}$$

$$= \frac{\partial^n C(u_1,...,u_n)}{\partial u_1,...,u_n} \frac{\partial F_1(x_1)}{\partial x_1} .... \frac{\partial F_n(x_n)}{\partial x_n} \tag{2. 2}$$

$$= c\left(u_1,...,u_n\right) f_1(x_1)....f_n(x_n)$$

This result shows that it is always possible to identify a joint density function by specifying the respective marginal densities of the random variables and a copula density. Taking this fact into account, it can be claimed that all the information about the dependence structure among random variables exists in the copula.

Some of the copula properties such as being invariant to strictly increasing transformation of the random variables and the ability to measure concordance between random variables are indeed extremely helpful in the dependence study.

Invariance Theorem: Consider n continuous random variables $X_i, \quad i = 1,...,n,$ with copula C. Then, if $h_i(X_i), \quad i = 1,...,n,$ are increasing on the range of

$X_i, i = 1,...,n,$ the random variables $Y_i = h_i(X_i),$ $i = 1,...,n,$ have exactly the same copula C.

It validated from the invariance theorem that the full dependence among random variables is entirely captured by the copula [without considering the shape of marginal distributions]. This is shown in equation (2.2).

Next, we state several useful properties of copula.

Property 1 (Nelsen, 2006): Given a copula C, for all $0 \le u_i \le 1$ and $0 \le v_i \le 1$, $i = 1,...,n$,

$| C(u_1,...,u_n) - C(v_1,...,v_n) | \le | u_1 - v_1 | + ... + | u_n - v_n |$. This reflects that any copula is uniformly continuous.

Property 2 (Nelsen, 2006): Let copula C be an *n*-copula. For almost all $u_i \in [0,1]$ and $i = 1,...,n$, the partial derivative of C with respect to u_i, $i = 1,...,n$ exists and,

$$0 \le \frac{\partial^n C(u_1,...,u_n)}{\partial u_1,...,u_n} \le 1.$$

These two properties indicate that copula enjoys a nice regular condition, means enough diferentiable, which is useful for numerical simulation. The next property provides the boundaries of copulas.

Property 3 (Nelsen, 2006): Given a copula C, for all u_i *for* $i = 1,...,n$, $W(u_1,...,u_n) = \max\{u_1 + ... + u_n - n + 1, 0\} \le | C(u_1,...,u_n) | \le M(u_1,...,u_n) = \min\{u_1,...,u_n\}.$

The upper bound is always a copula and signifies the strongest type of dependence between random variables. However, its lower bound is a copula in merely two dimensions. Those upper and lower copulas known as *Frèchet-Hoeffding bounds* or simply *Frèchet* bounds in the two dimensions with variable (u, v) are represented in Figure 2.1.

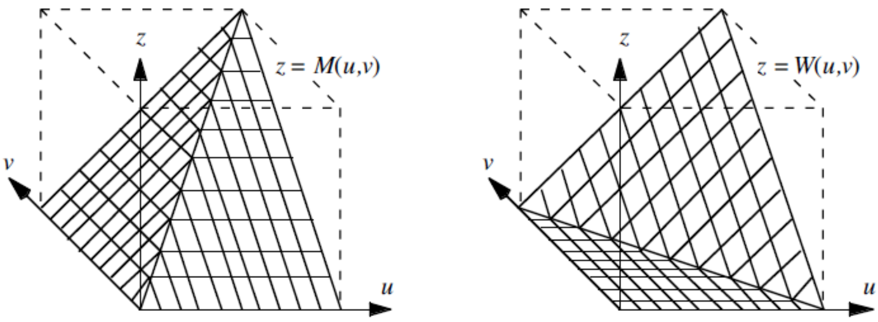


Figure 2.1: Frèchet-Hoeffding bounds

Some basic instances of copulas are as follows:

Independent copula, $\Pi_n(u) = \prod_{i=1}^{n} u_i$, is associated with a random vector $U = (U_1, U_2, ..., U_n)$ whose components are independent and uniformly distributed on $[0,1]$.

The comonotonicity copula, $M_n = \min\{u_1, u_2, ..., u_n\}$, is associated with a random vector $U = (U_1, U_2, ..., U_n)$ whose components are uniformly distributed on $[0,1]$ and is such that equality hold, $U_1 = U_2 = ... = U_n$, almost surely.

The counter monotonicity copula, $W_2 = \max\{u_1 + u_2 - 1, 0\}$, is associated with a random vector $U = (U_1, U_2)$ whose components are uniformly distributed on $[0,1]$ and is such that $U_1 = 1 - U_2$, almost surely.

Frèchet-Mardia copula is defined by a convex linear combination of independent and comonotonicity copula $C_n^{FM}(u) = \alpha\, \Pi_n(u) + (1-\alpha)M_n(u)$. In general every convex linear combination of copula is a copula.

Another important concept in copula is survival copula as define in definition 2.2. However, the subject of survival copula is beyond the scope of this study.

Definition 2.2: Given n random variables X_i, $i = 1,...,n$, with marginal survival distribution $\overline{F}_i$ for $i = 1,...,n$ and joint survival distribution $\overline{F}$, the survival copula $\overline{C}$ is such that:

$$\overline{C}\left(\overline{F}_1(x_1), ..., \overline{F}_n(x_n)\right) = \overline{F}(x_1, ..., x_n).$$

The dual copula C^* of the copula C of $X_i, i = 1,...,n$ is defined by:

$$C^*(u_1, ..., u_n) = 1 - \overline{C}(1-u_1, ..., 1-u_n), \qquad \forall u_1, ..., u_n \in [0,1].$$

2.2 Estimation of Copulas

This section provides two important topics on estimation of copula. The first topic studies and evaluates the most representative approaches in estimating copulas. The focus of the second topics is the problem of model selection and goodness-of-fit test.

Estimation of copulas enjoys a huge body of literature which can be divided into three groups, depending on the methods of estimating the marginal cumulative distribution functions (CDF) and joint CDF. Based on the assumptions made on CDF functions, some functions are estimated (i) parametrically, (ii) semi or (iii) non-parametrically (Charpentier, Fermanian, & Scaillet, 2007; Choroś, Ibragimov, & Permiakova, 2010; Fermanian & Scaillet, 2003; Genest, Ghoudi, & Rivest, 1995). In this section, we summarize the most popular techniques for parametric, semi-parametric and non-parametric methods simultaneously.

2.2.1 Parametric Estimation

Among different methods of parametric estimation of copula, we will focus on Maximum Likelihood Estimation (MLE) and the method of Inference Function for Margins (IFM) since they are the most effective methods.

The log likelihood function of multivariate distribution function $f(x_1,...,x_n)$ of a random sample of identically independent (i.i.d) vectors $x^j = (x_1^j, x_2^j ..., x_n^j), j = 1,2,...,m$ is

$$L = \sum_{j=1}^{m} \log f(x_1^j, x_2^j ..., x_n^j)$$

$$= \sum_{j=1}^{m} \log\left[c\left(F_1(x_1^j),...,F_n(x_1^j)\right)f_1(x_1^j)....f_n(x_1^j)\right] \tag{2.3}$$

$$= \sum_{j=1}^{m} \left[\log c\left(F_1(x_1^j),...,F_n(x_1^j)\right) + \sum_{i=1}^{n} \log f_i(x_i^j) \right]$$

$$= L_C + \sum_{i=1}^{n} L_i$$

L_C is the log likelihood contribution from the dependence structure of joint distribution function and L_i is log likelihood contribution from the margins (Joe, 1997).

Let us further assume that the copula belongs to a family of copulas indexed by a vector of parameters θ, $C(u_1,...,u_n\,;\theta)$. We also assume that $F_i(x_i,\alpha_i)$ $i = 1,2,...,n$ are margins with the corresponding univariate densities $f_i(x_i,\alpha_i)$ with parameters α_i. The maximum likelihood estimation of the model parameters $(\alpha_1,\alpha_2,...,\alpha_n\,;\theta)$ corresponds to simultaneous maximization of log-likelihood L in equation (2.3):

$$\begin{aligned}
&\left(\hat{\alpha}_1^{MLE},\hat{\alpha}_2^{MLE},...,\hat{\alpha}_n^{MLE},\hat{\theta}^{MLE}\right)\\
&= \arg\max_{\alpha_1,\alpha_2,...\alpha_n,\theta} L(\alpha_1,\alpha_2,...,\alpha_n,\theta)\\
&= \arg\max_{\alpha_1,\alpha_2,...\alpha_n,\theta} \sum_{j=1}^{m} \log f(x_1^j,x_2^j...,x_n^j,\alpha_1,\alpha_2,...,\alpha_n,\theta)\\
&= \arg\max_{\alpha_1,\alpha_2,...\alpha_n,\theta} \sum_{j=1}^{m} \log c\left(F_1(x_1^j,\alpha_1),...,F_n(x_n^j,\alpha_n),\theta\right).f_1(x_1^j,\alpha_1)....f_n(x_1^j,\alpha_n)\\
&= \arg\max_{\alpha_1,\alpha_2,...\alpha_n,\theta} \sum_{j=1}^{m} \left[\log c\left(F_1(x_1^j,\alpha_1),...,F_n(x_n^j,\alpha_n),\theta\right) + \sum_{i=1}^{n} \log f_i(x_i^j,\alpha_i) \right]\\
&= \arg\max_{\alpha_1,\alpha_2,...\alpha_n,\theta} \left\{ L_C(\alpha_1,\alpha_2,...,\alpha_n,\theta) + \sum_{i=1}^{n} L_i(\alpha_i) \right\}.
\end{aligned}$$

Since the computation of MLE is time consuming and cannot be done easily especially in high dimensional case, the method of inference function for margins (IFM) has been introduced (Joe, 1997). In the first stage of the IFM, the estimation of the parameters is projected from the log likelihood L_i of each margin under independence assumption. The result from this stage yields the estimation of $\left(\hat{\alpha}_1^{IFM}, \hat{\alpha}_2^{IFM}, \ldots, \hat{\alpha}_n^{IFM}\right)$. In the second stage of IFM, the estimator $\hat{\theta}^{IFM}$ of the copula parameter θ^{IFM} is computed by maximizing the copula likelihood with the margins estimators calculated in the first stage.

The MLE and IFM coincide when the copula is multivariate Gaussian with univariate Normal margins. Both the MLE and IFM estimators are consistent and asymptotical normal under the usual regularity condition (Frees & Valdez, 1998; Joe, 1997; Klugman & Parsa, 1999).

Although the method of MLE estimation optimize all parameters simultaneously, the IFM is more effective when dealing with samples of different length. In such cases, the complete sets of samples are used for the estimation of marginal parameters. While the MLE is asymptotically more efficient than the IFM, the accuracy of IFM is much higher when the sample size is small (Joe, 1997; Patton, 2006).

2.2.2 Semi-Parametric Estimation

Basically, there are two methods of semi-parametric estimations. The first one is based on concordance measure and the second one is a Pseudo maximum likelihood estimation.

The concordance estimation method is a simple method based on non-parametric estimation of parameters which depends merely on the copula. For this procedure, concordance measures like the Kendall's τ and Spearman's ρ, for example, are computed easily, and then parametric family of copulas will be estimated as a function of those estimated quantities.

For example, the Kendall's τ is estimated according to the bivariate sample of size n, $T:\{(x_1, y_1),...,(x_n, y_n)\}$ by using the following formula

$$\hat{\tau} = \frac{2(C-D)}{T(T-1)},$$

where C (resp. D) denote the number of concordant (resp. discordant) pairs, that is, a pair (x_i, y_i) and (x_j, y_j) such that $(x_i - x_j)(y_i - y_j) > 0$, (resp. < 0) (Genest & Favre, 2007; Oakes, 1982).

For the elliptical copula, the Spearman's ρ can be estimated by $\hat{\rho} = \sin\left(\frac{\pi}{2}\hat{\tau}\right)$, based on relation $\tau = \frac{2}{\pi}\arcsin\rho$ (Nelsen, 2006).

The alternative semi-parametric estimation method is based on Pseudo Maximum Likelihood (PML) estimation. This model is similar to the IFM procedures discussed in the parametric method, motivated by density representation and decomposition of log likelihood function of equation (2.3).

For PML, empirical distribution function $\hat{F}_i$ is employed to estimate the margins in first stage. In the second phase, the copula parameters are estimated through the maximization of likelihood function from the dependence structure which is represented by copula function $C(u_1,...,u_n,\theta)$ as

$$\widehat{\theta} = \max_{\theta} \left\{ \sum_{j=1}^{m} \log c\left(\hat{F}_1(x_1^j),...,\hat{F}_n(x_1^j),\theta\right) \right\}.$$

The estimator $\hat{\theta}$ is consistent and asymptotically normal under suitable regularity condition. While concordance measure of estimation is simple and robust, the pseudo likelihood estimation method is more accurate in general (Genest et al., 1995).

2.2.3 Non-Parametric Estimation

Considering the inverse formula of $F(x_1,...,x_n) = C\left(F_1(x_1),...,F_n(x_n)\right)$, most of the non parametric estimation of copula can be constructed. Suppose $\hat{F}$ is a nonparametric estimation of distribution function F and $\hat{F}_i^{-1}$ for $i = 1,...,n$ are a non parametric estimation of the pseudo-inverses $F_i^{-1}(s) = \{t \mid F_i(t) \geq s\}$ on the univariate margins F_i for $i = 1,...,n$, then empirical estimated copula is given by (Deheuvels, 1981)

$$\hat{C}(u_1,u_2,...,u_n) = \hat{F}\left(\hat{F}_1^{-1}(u_1),\hat{F}_2^{-1}(u_2),...,\hat{F}_n^{-1}(u_n)\right).$$

The problem with this estimation is that even if the marginal distributions are continuous, their empirical distributions are not. Therefore, one cannot determine a unique estimate of copula $\hat{C}$. Following this approach, a unique non-parametric estimator of C defined at T discrete point $\left(\frac{i_1}{T},\frac{i_2}{T},...,\frac{i_n}{T}\right)$ with $i_k \in \{1,2,...,T\}$, would be

$$\hat{C}\left(\frac{i_1}{T},\frac{i_2}{T},...,\frac{i_n}{T}\right) = \frac{1}{T}\sum_{k=1}^{T} 1_{\{x_1(k) \leq x_1(i_1;T),...,x_n(k) \leq x_n(i_n;T)\}} \quad . \tag{2.4}$$

where $x_p(k;T)$ denotes the k^{th} order statistics of the sample. According to Deheuvels (Deheuvels, 1981), any copula which satisfied the equation (2.4) is an empirical copula. The empirical copula is a multivariate distribution function which almost surely converges uniformly to the underlying copula.

Another approach of estimation is to smooth the margins and joint CDFs. To this end, Kernel based approach is the simplest method to employ. Consider a univariate kernel function $K:\Re \to \Re$, $\int K = 1$, and a bandwidth sequence $h_T > 0$, and $h_T \to 0$ when sample size $T \to \infty$. Then the *kth* margin distribution function, $\hat{F}_k(x)$, can be estimated by (Fermanian & Scaillet, 2003)

$$\hat{F}_k(x) = \frac{1}{T}\sum_{i=1}^{T} \kappa\left(\frac{x - X_{ki}}{h}\right),$$

for every real number *x*, where κ is the primitive function of $\kappa(x) = \int_{-\infty}^{x} K$. Similarly, the kernel estimation of joint CDF F can be obtained by $\hat{F}(x) = \frac{1}{T}\sum_{i=1}^{T} \kappa\left(\frac{x - X_i}{h}\right)$, with $n-$dimensional kernel $\kappa(x) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} K$.

Under mild regularity condition(Abdous, Genest, & Rémillard, 2005), the kernel method is asymptotically Guassian: $\hat{C}(u) - C(u) \to N(0, C(u))$.

To sum up this section, we summarize the advantages and disadvantages of different estimation methods. While non-parametric estimators provide a robust and universal way of estimation, they have some drawbacks: from the visual viewpoint, the graphical representation of copula is not pleasant. Moreover, since the copula estimator is not differentiable, it cannot be used directly to derive an estimate of the associated copula density or for optimization purposes. Smoothed estimators are more suitable for graphical use but they suffer from the curse of dimensionality. In other words, when the dimension increases, the complexity of the problem increases exponentially (Charpentier et al., 2007).

The advantages of having a family of copulas which are differentiable in semi-parametric/parametric approach solve the problem of estimation of copula density and optimization problem. However, they can lead to several underestimations when the parametric models for margins/copula are misspecified (Genest et al., 1995; Joe, 1997).

2.3 Goodness of Fit Test

The problem of estimation of copulas is evident when copula C is unknown but is assumed to belong to a specific class of copulas. The problem of goodness of fit is then to test the null hypothesis of $H_0 : C \in \{C_\theta ; \theta \in \Theta\}$ against the alternative $H_1 : C \notin \{C_\theta ; \theta \in \Theta\}$. Although the goodness of fit is relatively new compared to copula estimation, there are some interesting literatures available which measures the goodness of fit (Genest, Remillard, & Beaudoin, 2009; Genest & Rivest, 1993):

1) The method which is developed to test specific dependence structure as normal copula.
2) The methods that can be used for any class of copula but the implementation need some strategic choice.
3) Blanket test that can be used for all classes of copula without considering any strategic choice.

Berg (Berg, 2009) classified the methods into binned approach, multivariate kernel density estimation and dimension reduction approach.

The most important goodness of fit method, considering the classification of Berg and Genest, is considered in this section. We provide some preliminaries and overview of the five methods under their classification.

2.3.1 Preliminary

The goodness-of-fit of ML estimates is usually measured by the log likelihood or some statistics based on it, for example, the AIC. An alternative approach is based on matching the proportions predicted under a specified model (in terms of the expected proportion) to those of the empirical model (observed data). For the goodness-of-fit of the copula, the latter approach leads to the testing of the validity of the null hypothesis $H_0 : C \in C_0$.

Data

Since copula alone is important in concept of goodness of fit for copula, the ideal is not to consider any assumption regarding marginal distributions. One solution to overcome the problem of margins is to carry the test based on rank data. Suppose there are d independent sample $X_1 = (x_{11},...,x_{1n}),...,X_d = (x_{d1},...,x_{dn})$ from n-dimensional random vector X. Then the goodness of fit test for copula is based on pseudo-samples $Z_1 = (z_{11},...,z_{1n}),...,Z_d = (z_{d1},...,z_{dn})$, where

$$Z_j = (z_{j1},...,z_{jn}) = \left(\frac{R_{j1}}{n+1},...,\frac{R_{jn}}{n+1} \right).$$

where R_{ji} is the rank of x_{ji} amongst $(x_{1i},...,x_{di})$. This transformation is often denoted as the empirical marginal transformation. Although the pseudo sample is a sample from the underlying copula, it is no longer independent. To overcome this problem we need to apply bootstrap procedures to achieve reliable P-value estimates.

Rosanblatt's Transformation (Rosenblatt, 1952)

By Rosenblatt transformation, a set of dependent variables with given multivariate distribution function transform to a set of independent variables in $[0,1]$.

Definition: Let $Z=(Z_1,...,Z_n)$, denote a random vector with marginal distributions $F_i(z_i)=P(Z_i \le z_i)$, and conditional distributions

$$F_{i|1,...,i-1}(Z_i \le z_i \mid Z_1 = z_1,...,Z_{i-1} = z_{i-1}), \text{ for } i=1,...,n.$$

The Rosenblatt transformation of Z is defined as $V=(V_1,...,V_n)$, where

$$\begin{aligned}
V_1 &= R_1(Z_1) = P(Z_1 \le z_1) = F_1(z_1),\\
V_2 &= R_2(Z_2) = P(Z_2 \le z_2 \mid Z_1 = z_1) = F_{2|1}(z_2 \mid z_1),\\
&\vdots\\
V_n &= R_n(Z_n) = P(Z_n \le z_n \mid Z_1 = z_1,...,Z_{n-1} = z_{n-1}) = F_{n|1,...,n-1}(z_n \mid z_1,...,z_{d-1}).
\end{aligned}$$

One of the interesting applications of such transformation is multivariate goodness of fit test. The goodness of fit test is based on independence of vectors. When assuming a multivariate distribution function is from a parametric family of copula which is the null hypothesis, the result of transformation should be independent.

2.3.2 Approach 1

The first approach is based on Rosenblatt's transform proposed by Breymann (Breymann, Dias, & Embrechts, 2003) which was generalized later by (Berg & Bakken, 2005). The hypothesis in this approach is that the resulting sample from Rosenblatt's transform $(v_1,...,v_n)$ is a sample from the independent copula, $\Pi_n(v) = v_1\ v_2\ ...\ v_n$. The next step to reduce the $n-$dimensions useing $W_{1j} = \sum_{i=1}^{d} \Gamma\{v_{i,j};\alpha\}, j = \{1,...,n\}$ where Γ is a weight function. Breymann suggest $\Gamma\{v_{i,j};\alpha\} = \Phi^{-1}(v_{i,j})^2$ but Berg (2005) shows that the Anderson-Darling statistics with weight function $\Gamma\{v_{i,j};\alpha\} = |\, v_{i,j} - 0.5\,|$ performs particularly well for testing C_θ from Gaussian null hypothesis.

According to the Berg classification this approach is a dimension reduction approach and it is also the blanket test.

2.3.3 Approach 2 and 3

Approaches based on empirical copula are important since the empirical copula is non-parametric. Therefore, they provide the main objective benchmark for testing the copula goodness of fit test (Genest, Remillard, et al., 2009). We state three statistics which works on empirical copula.

Following Deheuvels (Deheuvels, 1979), the empirical copula is defined as

$$\hat{C}(u) = \frac{1}{d+1}\sum_{j=1}^{d} I\{Z_{j1} \le u_1,...,Z_{jn} \le u_n\},$$

where Z_j is as defined in section (2.2.4) and $u = (u_1,...,u_n) \in [0,1]^n$. An obvious goodness of fit test would then be to measure the distance between the empirical copula $\hat{C}(z)$ and estimated copula $C_{\hat{\theta}}(z)$, that is, $d\,(\hat{C} - C_{\hat{\theta}}) = \sqrt{n}\left(\hat{C}(z) - C_{\hat{\theta}}(z)\right)$. Genest and Remillard (Genest & Rémillard, 2008) considered rank based version of Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS) statistics. These tests are shown to convergence (Genest & Rémillard, 2008).

2.3.4 Approach 4

The combination of Rosenblatt's transform and empirical copula propoces an interesting goodness of fit test which was proposed by (Genest, Remillard, et al., 2009). In this method, the data is first transformed via $V = \Re(Z)$ by the Rosenblatt's transform, then empirical copula $\hat{C}(v)$ is compared with independent copula, $\hat{C}_{\perp}(v)$. Then CvM statistic approach is applied. Genest shows the convergence of this method (Genest, QUESSY, & Rémillard, 2006).

2.3.5 Approach 5

A blanket test based on Kendall's transform which was examined by Genest and Rivest (Genest & Rivest, 1993) and Wang and Wells (Wang & Wells, 2000), can be used as a goodness of fit test to compare the goodness of fit of the estimated copulas.

Let $X_i,\ i \in \{1, 2, ..., d\}\ \ , d \geq 2$ be a random sample. We consider the specific mapping $X \to V = H(X) = C(u_1,...,u_n)$, where $u_j = F_j(X_i)\ \ ,\ j \in \{1,2,...,n\}$ and joint distribution of $u = (u_1,...,u_n)$ is C. This transformation is called Kendall's transform (Barbe, Genest, Ghoudi, & Rémillard, 1996). Now, let K_n be the empirical distribution function based on pseudo-observations $V_1 = \widehat{C}(U_1),...,V_d = \widehat{C}(U_d)$ and $K \in [0, 1]$ the distribution function of the random variable $V = H(X)$. K_n is a consistent estimator of the distribution function K .

Now under the null hypothesis the vector $u = (u_1,..., u_n)$ is distributed as C_θ for some $\theta \in \Phi$, and hence the Kendall's transform has distribution K_θ . Thus, the new null hypothesis is $H_0^{"} : K = \{K_\theta : \theta \in \Phi\}$. Since $H_0^{'} \subset H_0^{"}$, the non rejection of $H_0^{"}$ does not entail the acceptance of null hypothesis. Therefore, tests based on the empirical process $\sqrt{n}\{K_n(t) - K(t)\}$ are not generally consistent (Genest et al., 2006). Acknowledging such limitation, Genest proposed this method by CvM and KS statistics.

(Wang & Wells, 2000) show that the null hypothesises, former and latter, are equal in the case of bivariate Archimedean copula; therefore Archimedean copulas are one of the well-known families of distributions that the method is consistent (Barbe et al., 1996). The distribution function K for the Archimedean copula can be written very simply, using the generator function $\varphi(t)$ for copula $H(Z)$ as follows:

$$K(t) = t + \sum_{i=1}^{d-1} (-1)^i \frac{\varphi^i(t)}{i!} f_{i-1}(t),$$

where $f_0(t) = \dfrac{1}{\varphi^{'}(t)}$ and $f_i(t) = \dfrac{f'_{i-1}(t)}{\varphi^{'}(t)}$.

Numerous metrics or distance measures can be employed as goodness-of-fit statistics to measure the difference between the empirical models and the hypothesis model. The measure used in this study is the L^2 norm distance:

$$d(p,q) = d(q,p) = \left((q_1 - p_1)^2 + (q_2 - p_2)^2 + \ldots + (q_n - p_n)^2\right)^{\frac{1}{2}},$$

where $p = (p_1, p_2, \ldots, p_n)$ and $q = (q_1, q_2, \ldots, q_n)$ are two points in Euclidean *n*-space.

As mentioned by Berg and Genest (Berg, 2009; Genest, Remillard, et al., 2009), it is not an easy task to come up with a specific method as the best goodness of fit statistic. But from the simulation, some interesting result emerge: among those tests which have both KS and CvM, CvM tend to be more powerful. For CvM transform, there is a little different in choosing between the Kendall's transform and empirical based method. Finally we can emphasis on the number of sample in the power of goodness of fit (Genest et al., 2006).

2.4 Family of Copulas

In this study we focus on two important families of copula: elliptical and Archimedean copula. Following this, we construct trigonometric copulas which are a subset of the Archimedean copulas. Although some concepts are defined in multivariate dimension, we will focus on two dimensions for ease of explanation. The extension of these trigonometric copulas to multivariate dimension will be discussed in chapter 4 when the concept of vine structure is introduced.

2.4.1 Elliptical Family

Elliptical copulas are deriven from multivariate elliptical distributions. The advantage of elliptical copulas is due to their numerically synthesizing property which makes them convenient for numerical simulation. Here, the two most significant instances are illustrated, namely, the Gaussian and t-copulas.

Gaussian copula is a natural setting for generalizing Gaussian multivariate distribution to meta- Gaussian distribution. The meta- Gaussian distribution has exactly similar dependence structure as the Gaussian distribution while the different in margins can be arbitrary (Fang, Fang, & Kotz, 2002; Hahn, Wagner, & Pfingsten, 2002).

To define a Gaussian copula we assume that Φ is a standard normal distribution function while $\Phi_{\rho,n}$ is $n-$dimensional standard Gaussian distribution with correlation matrix ρ. Then, the Gaussian $n-$copula with correlation matrix ρ is

$$C_{\rho,n}(u_1,...,u_n)=\Phi_{\rho,n}\left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_1)\right),$$

where density is

$$\begin{aligned} c_{\rho,n}(u_1,...,u_n) &= \frac{\partial^n C_{\rho,n}(u_1,...,u_n)}{\partial u_1,...,u_n} \\ &= \frac{1}{\sqrt{\det \rho}}\exp\left(\frac{-1}{2}y^t(u)(\rho^{-1}-I_n)y(u)\right), \end{aligned}$$

with $y^t(u)=\left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_1)\right)$ and I_n is identity matrix.

Like Gaussian copula, t-copula is derived from the Student multivariate distribution. In the form of meta-elliptical distribution t-copula have precisely similar dependence structure as the t- student distribution with arbitrary margins.

Suppose $T_{\rho,n,v}$ is $n-$dimensional student distribution function with v degrees of freedom and a shape matrix ρ. The t-copula which corresponds to t- student distribution function is

$$C_{\rho,n,v}(u_1,\ldots,u_n) = T_{\rho,n,v}\left(T^{-1}(u_1),\ldots,T^{-1}(u_1)\right),$$

where T is univariate student distribution with v degrees of freedom, and its copula density,

$$c_{\rho,n,v}(u_1,\ldots,u_n) = \frac{\partial^n C_{\rho,n,v}(u_1,\ldots,u_n)}{\partial u_1,\ldots,u_n}$$

$$= \frac{1}{\sqrt{\det \rho}} \times \frac{\Gamma\left(\frac{v+n}{2}\right)\left[\Gamma\left(\frac{v}{2}\right)\right]^{n-1}}{\left[\Gamma\left(\frac{v+1}{2}\right)\right]^n} \times \frac{\prod_{k=1}^n \left(1+\frac{y_k^2}{v}\right)^{\frac{v+1}{2}}}{\left(1+\frac{y^t \rho y}{v}\right)^{\frac{v+n}{2}}},$$

here $y^t(u) = \left(T^{-1}(u_1),\ldots,T^{-1}(u_1)\right)$.

The explanation of copula parameter depends on two parameters: the shape ρ, and the degrees of freedom v. A precise estimation of degree of freedom v is fairly difficult and has an effect on the estimation of the shape ρ. So, the student's copula might be more challenging to employ than Gaussian copula.

2.4.2 Construction of Archimedean Copula

In this part, one of the important classes of copula which is called Archimedean family of copula will be introduced. The importance of these copulas comes from the fact that they can be constructed easily and also some of the important copulas belong to this family. Also they possess some nice properties which is derived from their generator functions (Nelsen, 2006). We demonstrate the definition and the properties of Archimedean copula in two dimensions. Then we consider the family of Archimedean copula with one- and two-parameters.

A bivariate Archimedean copula C can be generated by considering a class Φ of functions $\varphi:(0,1]\rightarrow[0,\infty)$ which are continuous, strictly decreasing, convex, and for which $\varphi(1)=0$. This copula can be constructed, based on its generator φ, as follows:

$$C(u,v)=\varphi^{[-1]}(\varphi(u)+\varphi(v)), \quad 0\le u,v\le 1,$$

More generally we can extend this formula to include several dimensions, that is,

$$C^n(u_1,u_2,\ldots,u_n)=\varphi^{[-1]}(\varphi_1(u_1),\varphi_2(u_2),\ldots,\varphi_n(u_n)), \quad 0\le u_i\le 1,$$

$$\varphi^{[-1]}=\begin{cases}\varphi^{-1}(t) & 0\le t\le\varphi(0),\\ 0 & \varphi(0)\le t\le\infty.\end{cases}$$

where $\varphi^{[-1]}$ is the pseudo-inverse of the continuous and strictly decreasing function φ with domain $\varphi^{[-1]}=[0,\infty)$ and range $\varphi^{[-1]}=[0,1]$.

The pseudo-inverse $\varphi^{[-1]}$ is equal to the usual inverse function φ^{-1} if $\varphi(0)=\infty$. An important subclass of Φ, as noted by Nelsen (Nelsen, 2006) includes those elements φ which has two continuous derivatives with $\varphi'(t)\le 0$ and $\varphi''(t)\ge 0$ for $t\in(0,1)$.

The family of Archimedean copula can be constructed by one or more generator function. A single generator function can be constructed according to the following theorem:

Theorem: Let φ in Φ and let α and β be positive real numbers and define $\varphi_{\alpha,1}(t)=\varphi(t^{\alpha})$ and $\varphi_{1,\beta}(t)=[\varphi(t)]^{\beta}$ (Nelsen, 2006).

1) If $\beta\ge 1$, then $\varphi_{1,\beta}$ is an element of Φ.

2) If $\alpha \in (0,1]$, then $\varphi_{\alpha,1}$ is an element of Φ.

3) If φ is twice differentiable and $t\varphi'$ is non-decreasing on $(0,1)$, then $\varphi_{\alpha,1}(t)$ is an element of Φ for all $\alpha > 0$.

The family of Archimedean copula which was built on the $\varphi_{\alpha,1}(t) = \varphi(t^{\alpha})$ structure is called interior power family and those which are built on the $\varphi_{1,\beta}(t) = [\varphi(t)]^{\beta}$ structure is called the exterior family of copula. Following this, we introduce two members, Clayton and Gumbel, of this family (Nelsen, 2006).

The Clayton copula is given by

$$C^{c}(u,v) = \left[u^{-\theta} + v^{-\theta} - 1\right]^{\frac{-1}{\theta}}$$

The Clayton copula has a generator, $\varphi = \left(t^{-\theta} - 1\right)$ while $\varphi^{[-1]}(t) = \varphi^{-1}(t) = \left(t + 1\right)^{\frac{-1}{\theta}}$

where $\theta > 0$. It is completely monotonic; when $\theta \to 0$, $C^{c}(u,v) = uv$ and when $\theta \to \infty$, the upper Frèched-Hoefding bound is attained.

The Gumbel copula which belongs to the Extreme Values (EV) family (Nelsen, 2006) is expressed as

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

Its generator is $\varphi = (-\ln t)^{\theta}$ while $\varphi^{[-1]} = \varphi^{-1} = \exp\left(-t^{\frac{1}{\theta}}\right)$. The parameter θ controls the strength of dependence; $\theta = 1$, implies $C^{G}(u,v) = uv$, which reflects independence; and as $\theta \to \infty$, it yields a perfect dependence.

Two parametric family of Archimedean copula can be constructed easily by combining the interior and exterior structure for a specific generator function. For example by considering generator function $\varphi_{\alpha,\beta}(t) = (1-t^{\alpha})^{\beta}$, we have the following two-parametric Archimedean copulas:

$$C_{\alpha,\beta}(u,v) = \max\left\{\left[\ 1-\left[(1-u^{\alpha})^{\beta}+(1-v^{\alpha})^{\beta}\right]^{\frac{1}{\beta}}\right]^{\frac{1}{\alpha}}, 0\right\}$$

An example of a two- parameter family of Archimedean copula is the Joe-Clayton copula that belongs to the family of BB7 (Joe, 1997)and is expressed as

$$C^{JC}{}_{\alpha,\beta}(u,v) = 1-\left(\left[1-\left[\left(1-(1-u)^{\alpha}\right)^{-\beta}+\left(1-(1-v)^{\alpha}\right)^{-\beta}-1\right]^{\frac{-1}{\beta}}\right]^{\frac{1}{\alpha}}\right)$$

where $\varphi_{\alpha,\beta}(t) = \left[1-(1-t)^{\alpha}\right]^{-\beta}$ and $\varphi^{-1}{}_{\alpha,\beta}(t) = 1-\left[1-(1+t)^{\frac{-1}{\alpha}}\right]^{\frac{1}{\beta}}$.

The convex combination of two generator functions also can be a generator function to shape an Archimedean copula. For example, for $\pi \in [0,1]$ and two Archimedean copulas, namely, Clayton C^{C} and Gumbel C^{G}, we define

$$C^{GC}(u,v) = \pi\, C^{C}(u,v) + (1-\pi)\, C^{G}(u,v)$$

We denote this equation as a mixed Clayton-Gumbel (CG) copula. The features of these copulas can be derived from those of C^C and C^G.

2.4.2.1 Trigonometric Family

Finally, we end this chapter by introducing five new copulas from Archimedean family, called trigonometric copula, since the generator function is based on the trigonometric functions. We propose two new generators based on cotangent and another three based on cosecant of trigonometric function. The dependence properties of these copulas will be discussed later in Chapter 3.

Cot-Copula

Based on cotangent function we first define the generator as

$$\varphi(t) = \cot^{\theta}\left(\frac{\pi}{2}t\right), \qquad \theta \geq 1. \tag{2. 5}$$

The condition $\theta \geq 1$ in equation (2.5) guarantees the following properties of the generator function $\varphi(t)$

(i) $\varphi(1) = \cot^{\theta}\left(\frac{\pi}{2}\right) = 0$

(ii) $\varphi'(t) = -\theta\frac{\pi}{2} \cot^{\theta-1}\left(\frac{\pi}{2}t\right)\left(1+\cot^2\left(\frac{\pi}{2}t\right)\right) \leq 0 \quad \Leftrightarrow \quad \theta \geq 0$

(iii) $\varphi''(t) = \theta\left(\frac{\pi}{2}\right)^2 \cot^{\theta-2}\left(\frac{\pi}{2}t\right) \left(1+\cot^2\left(\frac{\pi}{2}t\right)\right)\left((\theta-1)+(\theta+1)\cot^2\left(\frac{\pi}{2}t\right)\right) \geq 0 \quad \Leftrightarrow \quad \theta \geq 1.$

In addition, $\varphi(0) = \lim_{t\to 0} \cot^{\theta}\left(\frac{\pi}{2}t\right) = \infty$, which suffices to guarantee that the strict inverse exists, that is,

$$\varphi^{[-1]}(t) = \varphi^{-1}(t) = \frac{2}{\pi} arc\cot\left(t^{\frac{1}{\theta}}\right) \tag{2.6}$$

From (2.5) and (2.6) the corresponding copula, called Cot-copula, is then defined as,

$$C(u,v) = \frac{2}{\pi} arc\cot\left(\cot^{\theta}\left(\frac{\pi}{2}u\right) + \cot^{\theta}\left(\frac{\pi}{2}v\right)\right)^{\frac{1}{\theta}} \qquad \theta \geq 1.$$

with density function given by

$$c(u,v) = \left(\cot^{\theta}\left(\frac{\pi}{2}u\right) + \cot^{\theta}\left(\frac{\pi}{2}v\right)\right)^{\frac{1}{\theta}-2}\left[\left(1+\left(\cot^{\theta}\left(\frac{\pi}{2}u\right)+\cot^{\theta}\left(\frac{\pi}{2}v\right)\right)^{\frac{2}{\theta}}\right)(1+\theta)-2\right]$$
$$\times\cot^{\theta-1}\left(\frac{\pi}{2}u\right)\cot^{\theta-1}\left(\frac{\pi}{2}v\right)\csc^{2}\left(\frac{\pi}{2}u\right)\csc^{2}\left(\frac{\pi}{2}v\right).$$

CotII-Copula

An alternative generator is defined by

$$\varphi(t) = \cot\left(\frac{\pi}{2}t^{\theta}\right), \quad \theta \geq 0.$$

The following properties for a generator function are satisfied by conditioning $\theta \geq 0$.

(i) $\varphi(1)=0$

(ii) $\varphi'(t)=-\theta\frac{\pi}{2}t^{\theta-1}\csc^2\left(\frac{\pi}{2}t^{\theta}\right)\leq 0 \quad \Leftrightarrow \quad \theta\geq 0$

(iii) $\varphi''(t)=\theta\frac{\pi}{2}t^{\theta-2}\csc^2\left(\frac{\pi}{2}t^{\theta}\right)\left[\theta-1+2\pi t^{\theta}\cot\left(\frac{\pi}{2}t\right)\right]\geq 0 \Leftrightarrow \quad \theta\geq 0$

In addition, the strict inverse exists due to the fact that $\varphi(0)=\lim_{t\to 0}\cot\left(\frac{\pi}{2}t^{\theta}\right)=\infty$, thus the inverse function of this generator can be expressed as $\varphi^{-1}(t)=\left(\frac{2}{\pi}arc\cot(t)\right)^{\frac{1}{\theta}}$.

The corresponding copula, which is called CotII-copula, is defined as:

$$C(u,v)=\left[\frac{2}{\pi}arc\cot\left(\cot\left(\frac{\pi}{2}u^{\theta}\right)+\cot\left(\frac{\pi}{2}v^{\theta}\right)\right)\right]^{\frac{1}{\theta}}+1, \qquad \theta\geq 0$$

with density function given by

$$\begin{aligned}c(u,v)=&\left(u^{\theta-1}\csc^2\left(\frac{\pi}{2}u^{\theta}\right)\right)\left(v^{\theta-1}\csc^2\left(\frac{\pi}{2}v^{\theta}\right)\right)\\&\times\left(arc\cot\left(\cot\left(\frac{\pi}{2}u^{\theta}\right)+\cot\left(\frac{\pi}{2}v^{\theta}\right)\right)\right)^{\frac{1}{\theta}-2}\left(\frac{2}{\theta}-2\right)\\&+2\left(\cot\left(\frac{\pi}{2}u^{\theta}\right)+\cot\left(\frac{\pi}{2}v^{\theta}\right)\right)\left(\frac{2}{\pi}arc\cot\left(\cot\left(\frac{\pi}{2}u^{\theta}\right)+\cot\left(\frac{\pi}{2}v^{\theta}\right)\right)\right).\end{aligned}$$

Csc-Copula

Now, based on cosecant function the third generator is defined as:

$$\varphi(t)=\left(\csc\left(\frac{\pi}{2}t\right)-1\right)^{\theta},\quad \theta\ge 0.5.$$

The following properties of a generator function is satisfied by conditioning $\theta \ge 0.5$.

(i) $\varphi(1)=0$

(ii) $\varphi'(t)=-\theta\frac{\pi}{2}\left(\csc\left(\frac{\pi}{2}t\right)-1\right)^{\theta-1}\csc\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right)\le 0 \quad \Leftrightarrow \quad \theta\ge 0$

(iii) $\varphi''(t)=\theta\frac{\pi^2}{4}\left(\csc\left(\frac{\pi}{2}t\right)-1\right)^{\theta-2}\csc\left(\frac{\pi}{2}t\right)$

$$\times\left[\csc\left(\frac{\pi}{2}t\right)\left(\theta\cot^2\left(\frac{\pi}{2}t\right)+\csc^2\left(\frac{\pi}{2}t\right)\right)-\left(\cot^2\left(\frac{\pi}{2}t\right)+\csc^2\left(\frac{\pi}{2}t\right)\right)\right]\ge 0 \quad \Leftrightarrow \quad \theta\ge 0.5$$

In addition, the strict inverse exists since $\varphi(0)=\lim_{t\to 0}\left(\csc\left(\frac{\pi}{2}t\right)-1\right)^{\theta}=\infty$.

Therefore, the inverse function of this generator is:

$$\varphi^{-1}(t)=\frac{2}{\pi}arc\csc\left(t^{\frac{1}{\theta}}+1\right)$$

The corresponding copula, called Csc-copula, is then defined by the following function,

$$C(u,v)=\frac{2}{\pi}arc\csc\left(\left(\left(\csc\left(\frac{\pi}{2}u\right)-1\right)^{\theta}+\left(\csc\left(\frac{\pi}{2}v\right)-1\right)^{\theta}\right)^{\frac{1}{\theta}}+1\right)\quad \theta\ge 0.5.$$

with density function given by

$$c(u,v)=\left(\frac{2}{\theta\pi}df_u df_v (A)^{\frac{1}{\theta}-2}\left(B^2-1\right)^{\frac{-1}{2}}\right)\left(\frac{(\theta-1)}{\theta B}+\frac{A^{\frac{1}{\theta}}}{\theta}\left(B^{-2}+\left(B^2-1\right)^{-1}\right)\right)$$

where:

$$A=\left(\csc\left(\frac{\pi}{2}u\right)-1\right)^{\theta}+\left(\csc\left(\frac{\pi}{2}v\right)-1\right)^{\theta},$$

$$B=\left(\left(\csc\left(\frac{\pi}{2}u\right)-1\right)^{\theta}+\left(\csc\left(\frac{\pi}{2}v\right)-1\right)^{\theta}\right)^{\frac{1}{\theta}+1},$$

$$df_u=-\frac{\theta\pi}{2}\left(\csc\left(\frac{\pi}{2}u\right)-1\right)^{\theta-1}\csc\left(\frac{\pi}{2}u\right)\cot\left(\frac{\pi}{2}u\right),$$

$$df_v=-\frac{\theta\pi}{2}\left(\csc\left(\frac{\pi}{2}v\right)-1\right)^{\theta-1}\csc\left(\frac{\pi}{2}v\right)\cot\left(\frac{\pi}{2}v\right).$$

CscII-Copula

The fourth generator is defined by $\varphi(t)=\csc\left(\frac{\pi}{2}t^{\theta}\right)-1,\quad \theta\geq 0$, when $\theta\geq 0$, the following properties of generator function is satisfied by condition $\theta\geq 0$

(i) $\varphi(1)=0$

(ii) $\varphi'(t)=-\frac{\pi\theta}{2}t^{\theta-1}\csc\left(\frac{\pi}{2}t^{\theta}\right)\cot\left(\frac{\pi}{2}t^{\theta}\right)\leq 0 \quad \Leftrightarrow \quad \theta\geq 0$

(iii) $$\varphi''(t)=\left(-\frac{\pi\theta}{2}\right)t^{\theta-2}\csc\left(\frac{\pi}{2}t^{\theta}\right)$$
$$\times\left[(\theta-1)\cot\left(\frac{\pi}{2}t^{\theta}\right)-\frac{\pi\theta}{2}t^{\theta}\left(\cot^{2}\left(\frac{\pi}{2}t^{\theta}\right)+\csc^{2}\left(\frac{\pi}{2}t^{\theta}\right)\right)\right]\geq 0 \quad\Leftrightarrow\quad \theta\geq 0$$

Similar to the previous case, the strict inverse exists since $\varphi(0)=\lim_{t\to 0}\csc(\frac{\pi}{2}t^{\theta})-1=\infty$. Therefore, the inverse function of this generator can be expressed as $\varphi^{-1}(t)=\left(\frac{2}{\pi}a\csc(t+1)\right)^{\frac{1}{\theta}}$.

The corresponding copula which is called CscII-copula is then defined by the following function

$$C(u,v)=\left(\frac{2}{\pi}a\csc\left(\csc\left(\frac{\pi}{2}u^{\theta}\right)+\csc\left(\frac{\pi}{2}v^{\theta}\right)-1\right)\right)^{\frac{1}{\theta}} \qquad \theta\geq 0$$

with density function, given by

$$c(u,v)=\left(\frac{2}{\theta\pi}\,df_u\ \ df_v\ \ B^{\frac{1}{\theta}-2}\left[\left(\frac{1}{\theta}-1\right)\frac{2}{\pi\,A^2\,\left(A^2-1\right)}+\frac{B}{A^2\sqrt{A^2-1)}}\right]\frac{1}{\left(A^2-1\right)}\right),$$

where:

$$A=\left(\csc\left(\frac{\pi}{2}u^{\theta}\right)-1\right)+\left(\csc\left(\frac{\pi}{2}v^{\theta}\right)-1\right)+1,$$

$$B=\frac{2}{\pi}arc\csc(A),$$

$$df_u = -\frac{\theta\pi}{2}(u)^{\theta-1}\csc\left(\frac{\pi}{2}u^{\theta}\right)\cot\left(\frac{\pi}{2}u^{\theta}\right),$$

$$df_v = -\frac{\theta\pi}{2}(v)^{\theta-1}\csc\left(\frac{\pi}{2}v^{\theta}\right)\cot\left(\frac{\pi}{2}v^{\theta}\right).$$

CscIII-Copula

Finally we propose another generator according to the cosecant function, defined by $\varphi(t) = \csc^{\theta}\left(\frac{\pi}{2}t\right) - 1, \quad \theta \geq 0.$

The following properties of generator function is satisfied when $\theta \geq 0$,

(i) $\varphi(1) = 0$

(ii) $\varphi'(t) = -\theta\frac{\pi}{2}\csc^{\theta}\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right) \leq 0 \quad \Leftrightarrow \theta \geq 0$

(iii) $\varphi''(t) = \theta\left(\frac{\pi}{2}\right)^2 \csc^{\theta}\left(\frac{\pi}{2}t\right)\left[\theta\cot\left(\frac{\pi}{2}t\right) + \csc^2\left(\frac{\pi}{2}t\right)\right] \quad \Leftrightarrow \theta \geq 0$

As in previous cases, the strict inverse exists since $\varphi(0) = \lim_{t\to 0}\csc^{\theta}\left(\frac{\pi}{2}t\right) - 1 = \infty$. Therefore, the inverse function of this generator is shown to be:

$$\varphi^{-1}(t) = \left(\frac{2}{\pi}arc\csc\left(t+1\right)\right)^{\frac{1}{\theta}}.$$

The corresponding copula, CscII-copula, which is then defined by the following function

$$C(u,v) = \frac{2}{\pi} arc\csc\left(\csc^{\theta}\left(\frac{\pi}{2}u\right) + \csc^{\theta}\left(\frac{\pi}{2}v\right) - 1\right)^{\frac{1}{\theta}} \qquad \theta \geq 0,$$

with density function given by

$$c(u,v) = \left(\frac{2B^{-\frac{1}{2}}}{\theta\,\pi\,A}\; df_u\; df_v\right)\left(\frac{1}{A} + \frac{A^{\frac{2}{\theta}-1}}{\theta\,B}\right),$$

where:

$$A = \left(\csc\left(\frac{\pi}{2}u\right)^{\theta} - 1\right) + \left(\csc\left(\frac{\pi}{2}v\right)^{\theta} - 1\right) + 1,$$

$$B = A^{\frac{2}{\theta}} - 1,$$

$$df_u = -\frac{\theta\pi}{2}\left(\csc\left(\frac{\pi}{2}u\right)\right)^{\theta} \cot\left(\frac{\pi}{2}u\right),$$

$$df_v = -\frac{\theta\pi}{2}\left(\csc\left(\frac{\pi}{2}v\right)\right)^{\theta} \cot\left(\frac{\pi}{2}v\right).$$

Considering the definition of the interior and exterior Archimedean copula, one can see the Cot and Csc copula are exterior copula while CotII and CscII are interior copulas. Finally, we can define interior copula by Gumbel generator as

$$\varphi(t) = -\ln\left(t^{\theta}\right) \qquad \theta \geq 0$$

$$c(u,v) = \left(\exp\left[\ln\left(u^{\theta}\right) + \ln\left(v^{\theta}\right)\right]\right)^{\frac{1}{\theta}} \qquad \theta \geq 0$$

Note that CscIII is neither interior nor exterior Archimedean copula. The trigonometric copulas are summarized in Table 2.1.

Table 2.1: Trigonometric Copulas

Copula name	Generator function	Copula formula	Parameter
Cot	$\varphi(t) = \cot^{\theta}\left(\frac{\pi}{2}t\right)$	$C(u,v) = \frac{2}{\pi} arc\cot\left(\cot^{\theta}\left(\frac{\pi}{2}u\right) + \cot^{\theta}\left(\frac{\pi}{2}v\right)\right)^{\frac{1}{\theta}}$	$\theta \geq 1$
Cot II	$\varphi(t) = \cot\left(\frac{\pi}{2}t^{\theta}\right)$	$C(u,v) = \left[\frac{2}{\pi} arc\cot\left(\cot\left(\frac{\pi}{2}u^{\theta}\right) + \cot\left(\frac{\pi}{2}v^{\theta}\right)\right)\right]^{\frac{1}{\theta}} + 1$	$\theta \geq 1$
Csc	$\varphi(t) = \left(\csc\left(\frac{\pi}{2}t\right) - 1\right)^{\theta}$	$C(u,v) = \frac{2}{\pi} arc\csc\left(\left(\csc\left(\frac{\pi}{2}u\right) - 1\right)^{\theta} + \left(\csc\left(\frac{\pi}{2}v\right) - 1\right)^{\theta}\right)^{\frac{1}{\theta}} + 1$	$\theta \geq 0.5$
CscII	$\varphi(t) = \csc\left(\frac{\pi}{2}t^{\theta}\right) - 1$	$C(u,v) = \left[\frac{2}{\pi} arc\csc\left(\csc\left(\frac{\pi}{2}u^{\theta}\right) + \csc\left(\frac{\pi}{2}v^{\theta}\right) - 1\right)\right]^{\frac{1}{\theta}}$	$\theta \geq 0$
CscIII	$\varphi(t) = \csc^{\theta}\left(\frac{\pi}{2}t\right) - 1$	$C(u,v) = \frac{2}{\pi} arc\csc\left(\csc^{\theta}\left(\frac{\pi}{2}u\right) + \csc^{\theta}\left(\frac{\pi}{2}v\right) - 1\right)^{\frac{1}{\theta}}$	$\theta \geq 0$

Chapter 3: Dependence and Trigonometric Copula

3.1 Introduction

In this chapter, the concept of dependence is presented and the properties of trigonometric copulas proposed in the previous section are studied. We further consider several important measures of dependences along with the dependence measure for trigonometric copulas. We end the chapter with simulation results using both data from symmetric and asymmetric distribution to compare specifically the ability of tail dependence measure for trigonometric copula.

3.2 Dependence Concept

In this section the most important theories of the consept of depndnece is reviewed.

3.2.1 Theory

Dependence relations between two random variables are important in determining the strength of their association or relationship. The initial concept of dependence was introduced by Karl Pearson by defining the measure of strength of linear relationship between two random variables (Balakrishnan & Lai, 2009; Joe, 1997; Nelsen, 2006).

Technically, the best way of presenting dependence between random variables is to define independence as a unique concept. Stochastically independence entails X and Y being completely useless in predicting one another. Using this approach, we next define the concept of dependence.

At a glance, if random variable X is a function of Y and Y is function of X, each of these random variables can be predicted from the other which contradicts with independence.

If there is a function b such that $\Pr[Y = b(X)] = 1$ then random variable Y is said to be *completely dependent* on X. If the function b is a one-to-one function, then X and Y are *mutually completely dependence*. The notion of mutual completely dependence is an antithesis of stochastic independence.

Kimeldorf and Sampson (1980) construct a pair of mutually completely dependent random variables, with uniform distribution function that converge to a pair of independent random variables. Therefore, mutual complete dependence is not a perfect opposite of independence. The concept of *monotonically dependent* is defined when b in $\Pr[Y = b(X)] = 1$ is a strict monotone function. More specifically, if b is an increasing (decreasing) function, we say random variables are increasingly (decreasingly), dependent. The necessary and sufficient condition that X and Y are increasingly (decreasingly) monotonically dependent is that its joint distribution function of random variables are Frèchet bounds (Kimeldorf, May, & Sampson, 1980) and ((Kimeldorf & Sampson, 1978).

X and Y are *functionally dependent*, if either $X = a(Y)$ or $Y = b(X)$ for some function a and b, and if $a(X) = b(Y)$, then X and Y are *implicitly dependent*. Therefore, the different notions of total dependence in decreasing order of strength are as follows:

- Linear dependence,

- Monotone dependence,
- Mutually completely dependence,
- Functional dependence,
- Implicit dependence.

The second concept of dependence which is introduced in this chapter is *positive dependence*. Positive dependence means that large values of Y tend to accompany large value of X, and similarly small values of Y tend to accompany small value of X. By the same principle, negative dependence between two random variables means large value of Y tend to accompany small value of X and vice versa (Harris, 1970).

Kimeldorf and Sampson (Kimeldorf & Sampson, 1987) define condition of positive dependence concept on joint distribution function H of X and Y as follow:

- $H \in F^{+} \Rightarrow H(x, y) \geq F(x)\ G(y)$ for all x and y.
- If $H(x, y) \in F^{+}$, so does $H^{+}(x, y)$.
- If $H(x, y) \in F^{+}$, so does $H_0(x, y) = F(x)\ G(y)$.
- If $(X, Y) \in F^{+}$, so does $(\varphi(X), Y) \in F^{+}$, where φ is any increasing function.
- If $(X, Y) \in F^{+}$, so does (Y, X).
- If $(X, Y) \in F^{+}$, so does $(-X, -Y)$.
- If H_n converges to H in distribution, then $H \in F^{+}$.

where F^{+} is a subfamily of distributions satisfying positive dependency. Recall that $H^{+}(x, y) = \min\ (F(x),\ G(y))$ and $H^{-}(x, y) = \max\ (0,\ F(x) + G(y) - 1)$ are the upper and lower Frèchet bounds, where $F(x)$ and $G(y)$ are the marginal distributions of X and Y, respectively. We list positive dependence concepts accordingly.

3.2.2 Positive Quadrant Dependence

Two random variables X, Y are Positive Quadrant Dependence (PQD) if and only if:

$$P(X \geq x)\, P(Y \geq y) \leq P(X \geq x, Y \geq y) \qquad \forall\, all \;\; x, y \tag{3.1}$$

or equivalently, if

$$P(X \leq x)\, P(Y \leq y) \leq P(X \leq x, Y \leq y) \qquad \forall\, all \;\; x, y \tag{3.2}$$

For (3.1) and (3.2), every increasing function a and b defined on the real line R implies that $\operatorname{cov}(a(X), b(Y)) \geq 0$ (Lehmann, 1966).

The PQD cannot be extended to multivariate dimension since equation (3.1) and (3.2) are equivalent only in two dimensions. In case of multivariate dimension, the random variables are said to be Positively Upper Orthant Dependent (PUOD) if

$$\prod_{i=1}^{n} P(X_i > x_i) \leq P(X_1 > x, X_2 > x, \ldots X_n > x_n).$$

3.2.3 Left-Tail Decreasing (LTD) and Right-Tail Increasing (RTI)

Consider two random variables denoted by X and Y respectively. A random variable Y is Left-Tail Decreasing (LTD) in X, denoted by $LTD(Y \mid X)$, if $P(Y \leq y \mid X \leq x)$ is decreasing in x for all y (Balakrishnan & Lai, 2009; Joe, 1997; Nelsen, 2006). That is

$$P(Y \leq y \mid X \leq x') \leq P(Y \leq y \mid X \leq x) \qquad \forall\, x < x' \quad \forall\, y;$$

The random variable X is Left-Tail Decreasing (LTD) in random variable Y, denoted by $LTD(X \mid Y)$, if $P(X \le x \mid Y \le y)$ is decreasing in y for all x. That is

$$P(X \le x \mid Y \le y') \le P(X \le x \mid Y \le y) \qquad \forall\, y < y' \quad \forall\, x.$$

Similarly Y is Right-Tail Increasing (RTI) in random variable X, denoted by $RTI(Y \mid X)$, if $P(Y > y \mid X > x)$ is increasing in x for all y:

$$P(Y > y \mid X > x') \le P(Y > y \mid X > x) \qquad \forall\, x' < x \quad \forall\, y;$$

Likewise, X is Right-Tail Decreasing (RTI) in random variable Y, denoted by $RTI(X \mid Y)$, if $P(X \le x \mid Y \le y)$ is decreasing in y for all x:

$$P(X > x \mid Y > y') \le P(X > x \mid Y > y) \qquad \forall\, y' < y \quad \forall\, x.$$

Suppose Y is RTI in X then

$$P(Y > y \mid X > x') \le P(Y > y \mid X > x) \qquad \forall x' < x \quad \forall y.$$

When $x' \to -\infty$ results $P(X > x)\, P(Y > y) \le P(X > x, Y > y)$. Hence $P(Y > y) \le P(Y > y \mid X > x)$, which means RTI is PQD, and similarly LTD implies PQD.

As additional tool to identify the property of LTD/RTI of copula (Avérous & Dortet-Bernadet, 2004) offered a link between the LTD/RTI dependence consept of a generator function and its alpha family assiaoated with. Based on the properties, if copula C_φ with generator function φ is LTD/RTI then the alpha family, C_{φ_α} associated with the φ is LTD/RTI.

3.2.4 Stochastically Increasing

A random variable Y is *Stochastically Increasing (SI)* which is also called *Positive Regression Dependent* in X, $SI(Y \mid X)$, if $P(Y > y \mid X = x)$ is increasing in x for all y:

$$P(Y > y \mid X = x') \leq P(Y > y \mid X = x) \qquad \forall x' \prec x \quad \forall y.$$

Likewise X is $SI(X \mid Y)$ in Y, if $P(X > x \mid Y = y)$ is increasing in y for all x

$$P(X > x \mid Y = y') \leq P(X > x \mid Y = y) \qquad \forall y' \prec y \quad \forall x.$$

3.2.5 Positive K-Dependent

The fourth concept of dependence is based on the probability integral transformation which is studied by Genest and Rivest (Genest & Rivest, 1993). Let K be the distribution function of a random variable V, which is a transformation of two random variables X and Y, via the copula C, $V = C(X,Y)$. For the Archimedean copula C_φ, the corresponding function K is denoted by K_φ in order to characterize the copula. This function is defined by:

$$K_\varphi(v) = v - \lambda(v) = v - \frac{\varphi(v)}{\varphi'(v^+)}, \qquad \forall 0 \leq v \leq 1.$$

where the $\varphi'(v^+)$ denotes the first derivative of φ at v. Copula C is Positive K-Dependent (PKD) if and only if $K(v) \leq K_{\varphi_0}(v)$ for all $0 \leq v \leq 1$. $K_{\varphi_0}(v)$ corresponds to independence distribution function by $\varphi_0(x) = -\ln(x)$ (Avérous & Dortet-Bernadet, 2004).

Generally, we state the following chain of implication for various concepts of dependence

$$\begin{array}{ccccc} SI(Y\mid X) \Rightarrow RTI(Y\mid X) \Rightarrow & PKD & \Leftarrow LTD(X\mid Y) & \Leftarrow & SI(X\mid Y). \\ & \Downarrow & & & \\ & PQD & & & \end{array} \tag{3.3}$$

A similar nesting feature is also valid for Negative Quadrant Dependence (NQD), Left Tail Increasing (LTI), Stochastically Decreasing (SD) and Negative K-Dependence (NKD) which is define by reversing the inequalities in the preceding definitions.

$$\begin{array}{ccccc} SD(Y\mid X) \Rightarrow RTD(Y\mid X) \Rightarrow & NKD & \Leftarrow LTI(X\mid Y) & \Leftarrow & SD(X\mid Y). \\ & \Downarrow & & & \\ & NQD & & & \end{array} \tag{3.4}$$

3.2.6 Aging Properties and Archimedean Copula Dependence

Verifying the distribution function with positive or negative dependence is not straight forward. Fortunately, in the case of Archimedean copula there is a connection between dependence and aging properties of their generator function which help us to find out the dependence properties of Archimedean copula in a simple way. These results also depend on the following notion of aging (Barlow & Proschan, 1975):

Let $F_{\varphi}(t)$ denotes univariate cumulative distribution function for each generator function φ which defined by $F_{\varphi}(t) = 1 - \varphi^{-1}(t) \quad \forall t \geq 0$. Let $\overline{F_{\varphi}}(t) = 1 - F_{\varphi}(t)$ then,

- Increasing Failure Rate (IFR): F_{φ} is IFR if $-\ln \overline{F}_{\varphi}$ is a convex function.

- Increasing Failure Rate in Average (IFRA): F_φ is IFRA if $-\ln \overline{F}_\varphi$ is a star-shaped function; i.e. $-\ln \overline{F}_\varphi(x)/x$ is increasing in x.
- New Better than Used (NBU): F_φ is NBU if $-\ln \overline{F}_\varphi$ is a super additive function; i.e.; $-\ln \overline{F}_\varphi(x+y) \geq -\ln \overline{F}_\varphi(x) - \ln \overline{F}_\varphi(y) \quad \forall \; x, y \geq 0.$

The corresponding negative concept of Decreasing Failure Rate (DFR), Decreasing Failure Rate in Average (DFRA), and New Worse than Used (NWU) can be defined mutatis mutandis. Recall also that a life distribution F_φ with density function f_φ is said to be *strongly unimodal* if $-\ln f_\varphi$ is a convex. These properties imply that F_φ is IFR.

Now, let Φ denote the set of continues generator function of Archimedean copula and Φ^* denote the subset of Φ of generators that are differentiable on (0, 1). From (Avérous & Dortet-Bernadet, 2004) we have the following results:

Proposition 3.1:

1) Let $\varphi \in \Phi^*$. Then:

   a) C_φ is PQD $\Leftrightarrow F_\varphi$ is NWU.

   b) C_φ is NQD $\Leftrightarrow F_\varphi$ is NBU.

2) Let $\varphi \in \Phi^*$. If f_φ denote the density function of F_φ then:

   a) C_φ is SD $\Leftrightarrow F_\varphi$ is strongly unimodal.

   b) C_φ is SI $\Leftrightarrow \ln f_\varphi$ is convex.

3) Let $\varphi \in \Phi^*$. Then,

a) C_φ is PKD $\Leftrightarrow F_\varphi$ is DFRA.

b) C_φ is NKD $\Leftrightarrow F_\varphi$ is IFRA.

4) Let $\varphi \in \Phi^*$. Then,

a) C_φ is LTD $\Leftrightarrow F_\varphi$ is DFR.

b) C_φ is LTI $\Leftrightarrow F_\varphi$ is IFR.

3.3 Trigonometric Dependence Properties

We first study the dependence properties of trigonometric copula by employing the chain properties (3.3) and (3.4) of the trigonometric copula. By starting with the strongest dependence properties which is SI/SD, we continue to check the other properties until one of these properties satisfied. The results compare to the other wellknown parametric Archimedean copula such as the 12th family, Gumbel and Clayton.

3.3.1 Cot Copula

According to Proposition 3.1 Cot copula is SI if and only if $\ln f_\varphi$ is a convex function, where f_φ denotes the density function of F_φ. For Cot copula, f_φ is given by,

$$f_\varphi = \frac{2t^{\frac{1}{\theta}-1}}{\pi\theta(1+t^{\frac{2}{\theta}})}.$$

Now, according to the basic concept in calculus, the convesity of $\ln f_{\varphi}$ is determined by specifying the sign of $g(t)$ which defined as:

$$g(t) = \ln(f_{\varphi}) = \ln \frac{2t^{\frac{1}{\theta}-1}}{\pi\theta(1+t^{\frac{2}{\theta}})}$$

In general to specify the sign we calculate the second derivative. For this,

$$g'(t) = \frac{(\frac{1}{\theta}-1)-(\frac{1}{\theta}+1)t^{\frac{2}{\theta}}}{t(1+t^{\frac{2}{\theta}})},$$

$$g''(t) = \frac{-\left(\frac{1}{\theta}+1\right)\left(\frac{2}{\theta}\right)t^{\frac{2}{\theta}}\left(1+t^{\frac{2}{\theta}}\right)-\left[\left(1+t^{\frac{2}{\theta}}\right)+\left(\frac{2}{\theta}\right)t^{\frac{2}{\theta}}\right]\left[\frac{1}{\theta}\left(1-t^{\frac{2}{\theta}}\right)-\left(1+t^{\frac{2}{\theta}}\right)\right]}{\left(t\left(1+t^{\frac{2}{\theta}}\right)\right)^2}. \qquad (3.5)$$

$g(t)$ is not a convex or concave function because when $\theta = 1$ equation (3.5) is

$$g''(t) = \frac{-4t^2\left(1+t^2\right)-\left[\left(1+t^2\right)+2t^2\right]\left[\left(1-t^2\right)-\left(1+t^2\right)\right]}{\left(t\left(1+t^2\right)\right)^2} = \frac{2t^2\left(t^2-1\right)}{\left(t\left(1+t^2\right)\right)^2},$$

which has a root at $t=1$. Our result shows that for the case $\theta \geq 1.2$, the Cot- copula is stochastically Increasing (SI).

To verify if Cot copula is Left Tail Decreasing (LTD), we need to check if $\ln(\overline{F})$ is a convex function. Let $F(t) = 1-\varphi^{-1}(t)$ and $\overline{F}(t) = 1 - F(t) \quad \forall\, 0 \leq t \leq 1$. Then

$$\overline{F}(t) = \frac{2}{\pi} arc\cot\left(t^{\frac{1}{\theta}}\right)$$

Properties 1: Cot copula is Left Tail Decreasing (LTD).

Cot copula is LTD if $g(t)=Ln\left(\frac{2}{\pi}arc\cot\left(t^{\frac{1}{\theta}}\right)\right)$ be a convex function. To ease the calculation we check the convexity of $\overline{F}(t)$. Then since $\overline{F}(t)=\exp\left(g(t)\right)$ convexity of $\overline{F}(t)$ implies convexity of $g(t)$.

$$\overline{F}'(t)=\frac{-2}{\pi\,\theta}\left(\frac{t^{\frac{1}{\theta}-1}}{1+t^{\frac{2}{\theta}}}\right)$$

$$\overline{F}''(t)=\frac{-2}{\pi\,\theta}\left(\frac{\left(\frac{1}{\theta}-1\right)t^{\frac{1}{\theta}-2}-\left(\frac{1}{\theta}+1\right)t^{\frac{3}{\theta}-2}}{\left(1+t^{\frac{2}{\theta}}\right)^2}\right)$$

$$=\frac{2}{\pi\,\theta}\left(\frac{\left(1-\frac{1}{\theta}\right)t^{\frac{1}{\theta}-2}+\left(\frac{1}{\theta}+1\right)t^{\frac{3}{\theta}-2}}{\left(1+t^{\frac{2}{\theta}}\right)^2}\right)$$

$$=\left(\frac{2\,t^{-2}}{\pi\,\theta\left(1+t^{\frac{2}{\theta}}\right)^2}\right)\left(\left(1-\frac{1}{\theta}\right)t^{\frac{1}{\theta}}+\left(\frac{1}{\theta}+1\right)t^{\frac{3}{\theta}}\right)$$

$\overline{F}''(t)$ is always positive if $\left(\left(1-\frac{1}{\theta}\right)t^{\frac{1}{\theta}}+\left(\frac{1}{\theta}+1\right)t^{\frac{3}{\theta}}\right)\geq 0$ which implies positivity of $\left(\left(1-\frac{1}{\theta}\right)+\left(\frac{1}{\theta}+1\right)t^{\frac{2}{\theta}}\right)\geq 0$.

To investigate whether the value $\left(\left(1-\frac{1}{\theta}\right)+\left(\frac{1}{\theta}+1\right)t^{\frac{2}{\theta}}\right)$ is positive we find the roots:

$$t^{\frac{2}{\theta}}=\frac{\left(\frac{1}{\theta}-1\right)}{\left(\frac{1}{\theta}+1\right)}.$$

Since $\theta \geq 1$ therefore $\frac{1}{\theta}-1 \leq 0$, it proves that the

$\left(\left(1-\frac{1}{\theta}\right)+\left(\frac{1}{\theta}+1\right)t^{\frac{2}{\theta}}\right) \geq 0 \quad \forall t \geq 0, \theta \geq 1$. This implies that $g(t)$ is a convex function.

Following the chain properties of (3.3), it is obvious that Cot- copula has the weaker dependence properties PKD and PQD.

3.3.2 Cot II Copula

Since the SI/SD in Archimedean copula implies the other positive dependence concept, we first start with this concept. For Cot II:

$$F(t)=1-\left(\frac{2}{\pi}arc\cot(t)\right)^{\frac{1}{\theta}}, \qquad \forall\, 0 \leq t \leq 1$$

$$f(t)=\frac{2}{\theta\pi}\left(\frac{2}{\pi}arc\cot(t)\right)^{\frac{1}{\theta}-1}\left(\frac{1}{1+t^2}\right), \qquad \forall\, 0 \leq t \leq 1, \theta \geq 1.$$

By assuming that $g(t)=\ln\left[\frac{2}{\theta\pi}\left(\frac{2}{\pi}arc\cot(t)\right)^{\frac{1}{\theta}-1}\left(\frac{1}{1+t^2}\right)\right]$ we have

$$g'(t) = \left(1 - \frac{1}{\theta}\right)\left(arc\cot(t)\right)^{-1} - \frac{2t}{1+t^2}, \qquad \forall 0 \le t \le 1\,, \theta \ge 1$$

$$g''(t) = \left(1 - \frac{1}{\theta}\right)\left(\frac{1}{1+t^2}\right)\left(arc\cot(t)\right)^{-2} + \frac{2t^2 - 2}{\left(1+t^2\right)^2}, \quad \forall 0 \le t \le 1\,, \theta \ge 1 \qquad (3.6)$$

Since the sign of $g''(t)$ in equation (3.6) is not absolute positive or absolute negative, the original function $g(t)$ it is not convex or concave (SI/SD) function. The result of simulation study with Matlab program is available in appendix.

The next strong dependence property is LTD or LTI. Since Cot II copula is an alpha family of generator $\varphi(t) = \cot\left(\frac{\pi}{2}t\right)$, LTD property of this generator can imply the LTD of Cot II copula.

In similar manner as Cot-copula we define $\overline{F}(t) = \frac{2}{\pi} arc\cot(t)$ for generator $\varphi(t)$. Then we check the convexity of function $g(t) = Ln\left(\frac{2}{\pi} arc\cot(t)\right)$ as follow:

$$g'(t) = \frac{-1}{\left(1+t^2\right) arc\cot(t)},$$

$$g''(t) = \left(1+t^2\right)^{-2}\left(arc\cot\,(t)\,\right)^{-2}\left(2t \;\; arc\cot(t) - 1\right).$$

Since the value of $\left(2t\, arc\cot(t) - 1\right)$ is equal to zero at $t \cong 0.4297, g''$ is not a convex/concave function. Which means Cot II is not LTD/LTI.

The next property in the chain dependence properties is PKD/NKD which is indicated by IFRA/DFRA according to proposition 1. Considering the aging properties, copula is IFRA/DFRA if the function $g(t)$ in equation (3.7) be increasing/decreasing in t.

$$g(t) = -\frac{\ln\left(\frac{2}{\pi} arc\cot(t)\right)^{\frac{1}{\theta}}}{t}, \tag{3.7}$$

It can be simply check by the sign of first derivatives

$$g'(t) = \frac{\left(\frac{t}{\theta\left(1+t^2\right) arc\cot(t)}\right) + \ln\left(\frac{2}{\pi} arc\cot(t)\right)^{\frac{1}{\theta}}}{t^2} \tag{3.8}$$

From the results in appendix $g'(t)$ in equation (3.8) is not positive or negative. Therefore Cot II is not PKD/NKD.

Final step in chain property is the PQD/NQD which is matched by NWU/NBU. F is NBU if $-\ln\overline{F}$ is a super additive function; i.e.;

$$-\ln\overline{F}(x+y) \geq -\ln\overline{F}(x) - \ln\overline{F}(y) \qquad \forall\ x, y \geq 0,$$

$$-\ln\overline{F}(x+y) + \ln\overline{F}(x) + \ln\overline{F}(y) \geq 0 \qquad \forall\ x, y \geq 0,$$

$$\ln\frac{\overline{F}(x)\overline{F}(y)}{\overline{F}(x+y)} \geq 0 \qquad \forall\ x, y \geq 0,$$

$$\frac{\overline{F}(x)\overline{F}(y)}{\overline{F}(x+y)} \geq 1 \qquad \forall\ x, y \geq 0,$$

In case of CotII we have:

$$\overline{F}(y) = \left(\frac{2}{\pi} arc\cot(y)\right)^{\frac{1}{\theta}}, \quad \overline{F}(x) = \left(\frac{2}{\pi} arc\cot(x)\right)^{\frac{1}{\theta}}, \quad \overline{F}(x+y) = \left(\frac{2}{\pi} arc\cot(x+y)\right)^{\frac{1}{\theta}}$$

therefore

$0 < \overline{F}(x), \overline{F}(y), \overline{F}(x+y) \le 1$ which implies that $\dfrac{\overline{F}(x)\overline{F}(y)}{\overline{F}(x+y)} \ge 0 \quad \forall \; x, y \ge 0,$

therefore CotII is not NQD or PQD.

3.3.3 Csc Copula

Properties 2: Csc copula is Stochastically Increasing (SI).

We start with the strongest dependence properties, SI/SD. For special case of Csc copula with inverse generator $\varphi^{-1}(t) = \dfrac{2}{\pi} arc\csc\left(t^{\frac{1}{\theta}} + 1\right)$, F function is

$$F(t) = 1 - \left(\frac{2}{\pi} arc\csc\left(t^{\frac{1}{\theta}} + 1\right)\right), \qquad \forall t \ge 0$$

with following derivative

$$f(t) = \frac{2}{\theta\pi}\left(\frac{t^{\frac{1}{2\theta}-1}}{|\,t^{\frac{1}{\theta}} + 1\,|\sqrt{\left(t^{\frac{1}{\theta}} + 2\right)}}\right) \qquad \forall t \ge 0\,, \theta \ge 0.5.$$

We define $g(t)$ as

$$g(t) = \ln\frac{2}{\theta\pi}\left(\frac{t^{\frac{1}{2\theta}-1}}{|\,t^{\frac{1}{\theta}} + 1\,|\sqrt{\left(t^{\frac{1}{\theta}} + 2\right)}}\right). \tag{3.9}$$

To ease the calculation we check the convexity of $f(t)$. Then since $f(t)=\exp\left(g(t)\right)$ convexity of $f(t)$ implies convexity of $g(t)$.

$$f'(t)=\left(\frac{2}{\pi\,\theta^2}\right)\left(A^{\frac{1}{2}}\,t^{\frac{1}{\theta}-2}\right)\left(\frac{(1-\theta)\left(t^{\frac{1}{\theta}}+1\right)-\left[t^{\frac{1}{\theta}}+t^{\frac{1}{\theta}}\left(t^{\frac{1}{\theta}}+1\right)^2 A^{\frac{-1}{2}}\right]}{\left(t^{\frac{2}{\theta}}+1\right)^2 A}\right)$$

$$=\left(\frac{2}{\pi\,\theta^2}\right)\left(A^{\frac{1}{2}}\,t^{\frac{1}{\theta}-2}\right)\left(\frac{(1-\theta)\left(t^{\frac{1}{\theta}}+1\right)-\left[t^{\frac{1}{\theta}}+t^{\frac{1}{\theta}}\left(t^{\frac{1}{\theta}}+1\right)^2 A^{\frac{-1}{2}}\right]}{\left(t^{\frac{2}{\theta}}+1\right)^2 A}\right)$$

$$=\left(\frac{2}{\pi\,\theta^2}\right)\left(A^{\frac{1}{2}}\,t^{\frac{1}{\theta}-2}\right)\left(\frac{(1-\theta)-t^{\frac{1}{\theta}}\left[\theta+\left(t^{\frac{1}{\theta}}+1\right)^2 A^{\frac{-1}{2}}\right]}{\left(t^{\frac{2}{\theta}}+1\right)^2 A}\right)$$

$$=\left(\frac{-2}{\pi\,\theta^2}\right)\left(t^{\frac{1}{\theta}-2}\right)\left(\frac{(\theta-1)\,A^{\frac{1}{2}}+t^{\frac{1}{\theta}}\left[\theta\,A^{\frac{1}{2}}+\left(t^{\frac{1}{\theta}}+1\right)^2\right]}{\left(t^{\frac{2}{\theta}}+1\right)^2 A}\right)$$

$$=\left(\frac{-2}{\pi\,\theta^2}\right)\left(t^{\frac{1}{\theta}-2}\right)\left(\frac{(\theta-1)\sqrt{1+2t^{\frac{-1}{\theta}}}+\theta\,A^{\frac{1}{2}}+\left(t^{\frac{1}{\theta}}+1\right)^2}{\left(t^{\frac{2}{\theta}}+1\right)^2\left(t^{\frac{1}{\theta}}+2\right)}\right)$$

where

$$A=\left(t^{\frac{2}{\theta}}+2t^{\frac{1}{\theta}}\right)$$

$$f''(t)=\left(\frac{-2}{\pi\,\theta^2}\right)\left[\left(\frac{1}{\theta}-2\right)\left(t^{\frac{1}{\theta}-3}\right)\left(\frac{(\theta-1)\sqrt{1+2t^{\frac{-1}{\theta}}}+\theta\,A^{\frac{1}{2}}+\left(t^{\frac{1}{\theta}}+1\right)^2}{\left(t^{\frac{2}{\theta}}+1\right)^2\left(t^{\frac{1}{\theta}}+2\right)}\right)+T'\left(t^{\frac{1}{\theta}-2}\right)\right]$$

where

$$T=\left(\frac{(\theta-1)\sqrt{1+2t^{\frac{-1}{\theta}}}+\theta\,A^{\frac{1}{2}}+\left(t^{\frac{1}{\theta}}+1\right)^2}{\left(t^{\frac{2}{\theta}}+1\right)^2\left(t^{\frac{1}{\theta}}+2\right)}\right)$$

and

$$T'=\left(\frac{\left(\left(\frac{1-\theta}{\theta}\right)\left(1+2t^{\frac{-1}{\theta}}\right)^{\frac{-1}{2}}t^{\frac{-1}{\theta}-1}+\frac{\theta}{2}A'A^{-\frac{1}{2}}+\frac{2}{\theta}\left(t^{\frac{1}{\theta}}+1\right)\left(t^{\frac{1}{\theta}-1}\right)\right)\left(t^{\frac{2}{\theta}}+1\right)^2\left(t^{\frac{1}{\theta}}+2\right)}{\left(t^{\frac{2}{\theta}}+1\right)^4\left(t^{\frac{1}{\theta}}+2\right)^2}\right)$$

$$-\left(\frac{\frac{4}{\theta}t^{\frac{2}{\theta}-1}\left(t^{\frac{2}{\theta}}+1\right)\left(t^{\frac{1}{\theta}}+2\right)+\frac{1}{\theta}t^{\frac{1}{\theta}-1}\left(t^{\frac{2}{\theta}}+1\right)^2}{\left(t^{\frac{2}{\theta}}+1\right)^4\left(t^{\frac{1}{\theta}}+2\right)^2}\right).$$

The value of $f''(t)$ is always positive for $\forall t\geq 0\,,\theta>0.5$. The test of convexity can also be done graphically. According to the Figure 3.1 the function $g(t)$ in equation

(3.9) is convex. Therefore, it is SI function. According to the chain properties, it is also LTD, PKD and PQD as well. Program in MATLAB can be found in appendix.

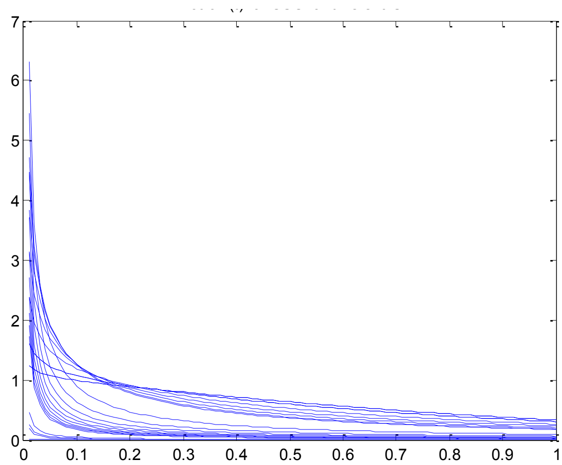


Figure 3.1: $g(t)$ plot of Csc-copula for different value of $\theta > 0.5$

3.3.4 CscII Copula

To investigate whether CscII copula has the strongest property of dependence, SI/SD, we check the function $g(t) = \ln(f(t))$ to be convex/concave:

$$\varphi^{-1}(t) = \left(\frac{2}{\pi} a\csc(t+1)\right)^{\frac{1}{\theta}},$$

$$F(t) = 1 - \left(\frac{2}{\pi} arc\csc(t+1)\right)^{\frac{1}{\theta}}, \qquad \forall t \geq 0$$

$$f(t) = \frac{4}{\theta\pi^2}\left[\frac{(arc\csc(t+1))^{\frac{1}{\theta}-1}}{|t+1|\sqrt{(t^2+2t)}}\right], \qquad \forall t \geq 0$$

Therefore $g(t)$ is defined as

$$g(t) = \ln\left(\frac{4}{\theta\pi^2}\left[\frac{(arc\csc(t+1))^{\frac{1}{\theta}-1}}{|t+1|\sqrt{(t^2+2t)}}\right]\right). \tag{3.10}$$

Simulation result of $g(t)$ in equation (3.10) for different parameter value $\theta \geq 0$, and $t \geq 0$ proves that $g(t)$ is a convex function. Result of this simulation can be found in Figure 3.2 the program in MATLAB can be found in appendix.

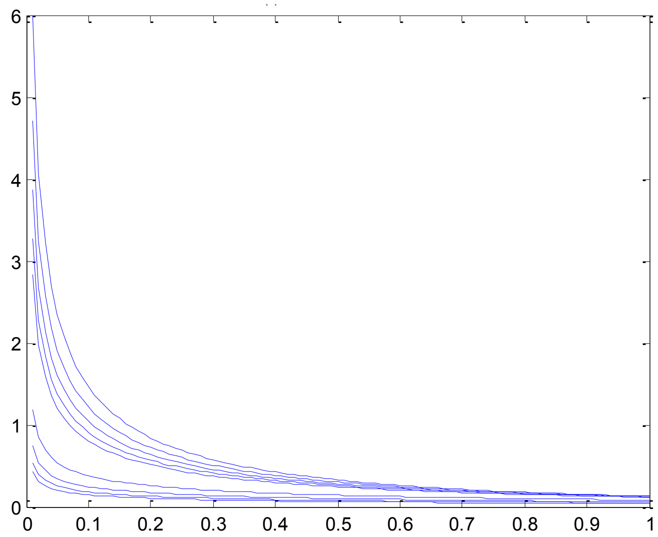


Figure 3.2: $g(t)$ plot of CscII for different value of $\theta \geq 0$

3.3.5 CscIII Copula

Properties 3: CscIII copula is Stochastically Increasing (SI).

To investigate whether CscIII copula has the strongest property of dependence, SI/SD, we check the function $g(t) = \ln(f(t))$ to be convex/concave:

$$\varphi(t) = \csc^{\theta}\left(\frac{\pi}{2}t\right) - 1,$$

$$\varphi^{-1}(t)=\frac{2}{\pi}arc\csc(t+1)^{\frac{1}{\theta}},$$

$$F(t)=1-\frac{2}{\pi}arc\csc(t+1)^{\frac{1}{\theta}}, \qquad \forall t\geq 0$$

Now $g(t)$ is defined as $\ln f(t)$ where $f(t)=\left(\frac{2}{\theta\pi|t+1|\sqrt{(t+1)^{\frac{2}{\theta}}-1}}\right)$:

$$g(t)=\ln\left(\frac{2}{\theta\pi|t+1|\sqrt{(t+1)^{\frac{2}{\theta}}-1}}\right). \tag{3.11}$$

To ease the calculation we check the convexity of $f(t)$. Then since $f(t)=\exp\left(g(t)\right)$ convexity of $f(t)$ implies convexity of $g(t)$ in equation (3.11).

$$f'(t)=\left(\frac{-2}{\pi\theta}\right)\left[\frac{\theta\left((t+1)^{\frac{2}{\theta}}-1\right)+(t+1)^{\frac{2}{\theta}-1}}{\theta\,(t+1)\left((t+1)^{\frac{2}{\theta}}-1\right)}\right]$$

$$=\left(\frac{-2}{\pi\theta}\right)\left[\frac{1}{(t+1)}+\frac{(t+1)^{\frac{2}{\theta}-2}}{\theta\left((t+1)^{\frac{2}{\theta}}-1\right)}\right],$$

and

$$f''(t)=\left(\frac{2}{\pi\theta}\right)\left(\frac{1}{(t+1)^2}+(t+1)^{\frac{2}{\theta}-3}\left(\frac{2\theta\left((t+1)^{\frac{2}{\theta}}-1\right)+2}{\theta^2\left((t+1)^{\frac{2}{\theta}}-1\right)^2}\right)\right) \tag{3.12}$$

which is always positive for $\forall t\geq 0$, $\theta\geq 0$. Therefore CscIII is SI function

$f''(t)$ in equation (3.12) is always positive which implies the $g(t)$ is convex function. It means CscIII is SI function. Therefore, considering the chain properties of dependence it is LTD, PKD and PQD.

3.3.6 Clayton Copula

Since trigonometric family of copula are Archimedean copula, we also recall the dependence properties of Gumbel, Clayton and A12th of this family. We start with the property of SI/SD for Clayton copula. According to proposition 1 Clayton, copula is SI/SD if and only if $\ln f_\varphi$ is convex/concave function. Where f_φ is derivative of $F_\varphi(t) = 1 - \varphi^{-1}(t)$ and for Clayton copula is define as

$$f_\varphi = \frac{1}{\theta}(1+t)^{-\frac{1}{\theta}-1} \ .$$

Then we rename logarithmic function f_φ as *g(t)*

$$g(t) = \ln\left[\frac{1}{\theta}(1+t)^{\frac{-1}{\theta}-1}\right].$$

To investigate where $g(t)$ is a convex function we find the second derivatives as follow:

$g''(t) = \left(\frac{1}{\theta}+1\right)(1+t)^{-2}$, which is greater than zero for all value of $0 \le t \, , \theta \ge 0$. Therefore, the Clayton copula is a Stochastically Increasing (SI) function. Consequently, it follows that the Clayton copula also got the properties of LTD, PKD and PQD.

3.3.7 Gumbel Copula

To consider the dependence properties of Gumbel copula, we start with the (SI/SD) property.

$$f_{\varphi} = \frac{1}{\theta} t^{\frac{1}{\theta}-1} \exp\left(-t^{\frac{1}{\theta}}\right) \qquad \forall \theta \geq 0, t \geq 0,$$

Similar to the previous cases $g(t)$ is define by $g(t) = \ln\left[\frac{1}{\theta} t^{\frac{1}{\theta}-1} \exp\left(-t^{\frac{1}{\theta}}\right)\right]$ and second derivatives is

$$g''(t) = \frac{1}{\theta} \exp\left(-t^{\frac{1}{\theta}}\right) \times \left[\frac{3}{\theta}\left(1-\frac{1}{\theta}\right) t^{\frac{2}{\theta}-3} + \frac{1}{\theta^2} t^{\frac{3}{\theta}-3} + \left(\frac{1}{\theta}-1\right)\left(\frac{1}{\theta}-2\right) t^{\frac{1}{\theta}-3}\right],$$

which is always positive for $0 \leq t, \theta \geq 1$. Therefore, Gumbel copula is SI and consequently it is LTD, PKD and PQD.

3.3.8 12th Family of Archimedean Copula

Starting with the (SI/SD) property for A12 we have

$$F(t) = 1 - \left(1 + t^{\frac{1}{\theta}}\right)^{-1} \qquad \forall t \geq 0$$

$$f(t) = \frac{1}{\theta} t^{\frac{1}{\theta}-1} \left(1 + t^{\frac{1}{\theta}}\right)^{-2} \qquad \forall t \geq 0, \theta \geq 1$$

Assuming that $g(t) = \ln\left[\frac{1}{\theta} t^{\frac{1}{\theta}-1} \left(1 + t^{\frac{1}{\theta}}\right)^{-2}\right]$ we have the followings:

$$g'(t) = t^{-1}\left[\left(\frac{1}{\theta}-1\right)-\frac{1}{\theta}t^{\frac{1}{\theta}}\right] \qquad \forall t \ge 0\,, \theta \ge 1$$

$$g''(t) = \left(1-\frac{1}{\theta}\right)\left(t^{-2}+\frac{1}{\theta}t^{\frac{1}{\theta}-2}\right) \qquad \forall t \ge 0\,, \theta \ge 1$$

$g''(t)$ is always positive, and therefore, the 12th family of Archimedean copula is Stochastically Increasing (SI) which implies the properties of LTD, PKD and PQD.

The summarized results are listed in Table 3.1. From Table 3.1, all copulas are positive dependence. Although the entire proposed copula do not have stronger positive dependence concept SI, they are similar to the properties of positive dependence since they are at least PKD. It implies that these copulas eventually enjoy similar properties. For example, positive dependence implies positive covariance between random variables.

Table 3.1: Dependence properties of trigonometric copulas

Dependence Concept	SI SD	RTD/LTI RTI/LTD	PKD NKD	PQD NQD
Cot	-	LTD	PKD	PQD
CSC	SI	LTD	PKD	PQD
CSC II	SI	LTD	PKD	PQD
CSC III	SI	LTD	PKD	PQD
Gumbel	SI	LTD	PKD	PQD
Clayton	SI	LTD	PKD	PQD
A12	SI	LTD	PKD	PQD

3.4 Dependence Measure

Fundamental researches which has been done on the concept of dependence measures show the importance of this topic. Obtaining a measure which can capture the dependence relationship between random variables is the final objective of those researches. Since dependence is an extensive concept and variables can be dependent in different aspects, one special measure can only capture dependence from one perspective. For example, the correlation coefficient measure the linear dependence between random variables and it is not equivalent to dependence. It means two independent random variables are surely uncorrelated while two uncorrelated random variables are not necessarily independent.

In this section, we review the global concept of dependence which every measure should have. Then some important measure of dependence such as correlation coefficient, rank correlation and tail dependence measure will be discussed.

3.4.1 Global Measure of Dependence

If X and Y are not totally dependent, then it may be helpful to find some quantities that can measure the strength or degree of dependence between them. If such a measure is a scalar, then we can refer to it as *index*. Let $\delta(X, Y)$ denotes an index of dependence between X and Y. The following conditions are the global dependence properties (Balakrishnan & Lai, 2009):

1) $\delta(X, Y)$ is defined for any pair of random variables, neither of random variables being constant, with probability 1.
2) $\delta(X, Y) = \delta(Y, X)$. While independence is a symmetric property, total dependence is not, (as one variable may be determined by the other, but not vice versa).
3) $0 \le \delta(X, Y) \le 1$.

4) $\delta(X,Y)=0$ if and only if X and Y are mutually dependent (Notice how strong this condition is made by the "only if" part).

5) $\delta(a(X),b(Y))=\delta(X,Y)$. The condition means that the index remains invariant under one-to-one transformation of the marginal random variables (Functions a and b map the spaces of X and Y, respectively, onto themselves, in a one-to-one manner).

6) $\delta(X,Y)=1$ if and only if X and Y are mutually completely dependent.

7) If X and Y are jointly normal, with correlation coefficient ρ, then

$$\delta(X,Y)=|\rho| .$$

8) $\delta(X,Y)=k(\theta)$, where k is a function of θ for any family of distributions defined by a vector parameter θ.

9) If (X,Y) and (X_n,Y_n), $n=1,2,...$ are pairs of random variables with joint distributions H and H_n, respectively, and if converges to H, as $n\to\infty$, then $\lim\delta(X_n,Y_n)=\delta(X,Y)$ (Hahn et al., 2002) .

Three most prominent global measures of dependence are correlation coefficient, Kendall's τ, and Spearman's correlation coefficient which are under rank correlation.

3.4.2 Correlation Coefficient

Correlation coefficient, also called Pearson product correlation coefficient, measures the strength and the direction of a linear relationship between two variables (Mari et al., 2001).

The correlation coefficient is given by,

$$r = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y},$$

where σ_X and σ_Y are the standard deviation of random variables, X and Y respectively.

Since r lies between [-1,1], a strong positive linear relationship represent with $r=+1$ and negative relationship with $r=-1$ and $r=0$ indicate none or weak linear correlation between these random variables.

3.4.3 Rank Correlation

The circumstance of moving random variables, X and Y toward each other without considering the exact mathematical relation can be defined as a measure of dependence, concordance measure or rank correlation, which is a fundamental tool for financial risk management. Here, we are interested to know that "the prices of two (or more) assets that tend to rise or fall together". Informally, a pair of random variables is concordant if large values of one tend to be associated to large value of the other and analogously for small values. Let X and Y are continuous random variables, then (x_i, y_i) $(x_i, y_i)$and (x_j, y_j) $(x_j, y_j)$are concordant if $x_i < x_j$ $x_i < x_j$and $y_i < y_j$ $y_i < y_j$or if $x_i > x_j$ $x_i < x_j$and $y_i > y_j$ $x_i > x_j$. Similarly, (x_i, y_i) $(x_i, y_i)$and (x_j, y_j) are discordant if $x_i < x_j$ and $y_i > y_j$ $y_i > y_j$, or if $x_i > x_j$ and $y_i < y_j$ $y_i < y_j$. For concordance, we note that $(x_i - x_j)(y_i - y_j) > 0$ $(x_i - x_j)(y_i - y_j) > 0$ and discordance where $(x_i - x_j)(y_i - y_j) < 0$ $(x_i - x_j)(y_i - y_j) < 0$. Any concordance measure of dependence ρ should satisfy the following properties (Kimeldorf et al., 1980) :

1) It defines for any pair of continuous random variables X and Y.

2) $\rho(X,Y)=\rho(Y,X)$, which implies symmetry.

3) $-1 \le \rho(X,Y) \le 1$, and reach these bounds in countermonotonic and comonotonic respectively.

4) $\rho(X,Y)=0$ if X and Y are independent random variables.

5) if the pair of random variables (X_1, X_2) is more dependent than the pair (Y_1, Y_2) in the following sense:

$C_X(u,v) \le C_Y(u,v) \quad \forall u,v \in [0,1]$, then the same ranking holds for any concordance measures ρ, it means that $\rho(X_1,X_2) \ge \rho(Y_1,Y_2)$.

This section introduces two well-known concordance measures, Kendall's τ and spearman's ρ.

3.4.3.1 Kendall's τ (τ)

By definition Kendall's τ is the difference between the probability of concordance and the probability of discordance. To define the sample version of Kendall's τ, let $\{(x_1, y_1),\dots,(x_n, y_n)\}$ denote a random sample of n observations from a vector (X,Y) of a continuous random variables. There are $\binom{n}{2}$ distinct pairs (x_i, y_i) and (x_j, y_j), let c denote the number of concordant pairs and d the number of discordant pairs. Then Kendall's τ for the sample is defined as

$$\tau = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}.$$

Mathematically Kendall's τ is defined as

$$\tau = \tau_{X_1,Y_1} = P\big[\,(X_1 - Y_1)(X_2 - Y)_2 \;\geq 0\;\big] - P\big[(X_1 - Y_1)\,(X_2 - Y_2) < 0\big]$$

where (X_1, Y_1) and (X_2, Y_2) be two independent and identically distributed random vectors. This quantity is invariant under increasing transformation of the marginal distributions. Accordingly, Kendall's τ depends only on the copula of (X, Y). Following theorems show the relation between Kendall's τ and copula (Nelsen, 2006).

Theorem (Nelsen 2006): *Let (X_1,Y_1) and (X_2,Y_2) be independent vectors of continuous random variables with joint distribution functions H_1 and H_2, respectively, with common margins F (of X_1 and X_2) and G (of $Y_1$and Y_2). Let C_1 and C_2 denote the copulas of (X_1,Y_1) and (X_2,Y_2), respectively, so that* $H_1(x,y) = \mathrm{C}_1(F(x),G(y))$ *and* $H_2(x,y) = C_2(F(x),G(y))$. *Let Q denote the difference between the probabilities of concordance and discordance of (X_1,Y_1) and (X_2,Y_2), i.e., let*

$$\tau = \tau(X_1,Y_1) = \tau_{X1,Y1} := P[(X_1 - Y_1)(X_2 - Y_2) > 0] - P[(X_1 - Y_1)(X_2 - Y_2) < 0]$$

Then

$$\tau_{X_1,X_2} = 4\iint_{I^2} C_1(u_1,u_2)\,dC_2(u_1,u_2) - 1.$$

Therefore:

Theorem (Nelsen 2006): *Let X and Y be continuous random variables with copula C. Then the population version of Kendall's τ for X and Y is given by*

$$\tau_{X,Y} = \tau_C = \begin{cases} 4\iint\limits_{[0,1]} C(u,v)dC(u,v)dudv - 1, \\ 1 - 4\iint\limits_{[0,1]} \frac{\partial C(u,v)}{\partial u}\frac{\partial C(u,v)}{\partial v}dudv. \end{cases}$$

We take note that the Kendall's τ depend only on the copula $C(u,v)$. Although the computation of this is difficult, there is a simple expression of Kendall's τ for the Archimedean copulas in terms of its generator φ,

$$\tau_{X,Y} = \tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 - 4\int_0^\infty u\left[\frac{d}{du}\varphi^{-1}(u)\right]^2 du.$$

*3.4.3.2 Spearman's ρ (*ρS*)*

As with Kendall's τ, the population version of the measure of association known as *Spearman's ρ* (denoted by *ρS*) is also based on concordance and discordance. Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent pairs of random variables with a common distribution function *H*. Then, *ρS* is defined to be proportional to the probability of concordance minus the probability of discordance for the two pairs (X_1, Y_1) and (X_2, Y_3),

$$\rho S = 3\left\{P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]\right\}.$$

In terms of the copula we have the following results:

$$\begin{aligned}\rho S &= 12\iint\limits_{[0,1]} C(u,v)dudv - 3 \\ &= 12\iint\limits_{[0,1]} uvdC(u,v) - 3\end{aligned}$$

By rewriting the above equation as $\rho S = \dfrac{E(UV) - 1/4}{1/12}$.

This simply means Spearman's rank correlation between X and Y is Pearson's product moment correlation coefficient between the uniform variates U and V (Nelsen, 2006).

The relation between Kendall's τ and Spearman's rank correlation for any copula C is as following formula:

$$\frac{3\tau - 1}{2} \le \rho S \le \frac{\tau^2 - 2\tau - 1}{2} \qquad \tau \ge 0,$$
$$\frac{\tau^2 + 2\tau - 1}{2} \le \rho S \le \frac{3\tau + 1}{2} \qquad \tau \le 0.$$

Figure 3.3, reflects the strong relation between Kendall's τ and spearman's ρ.

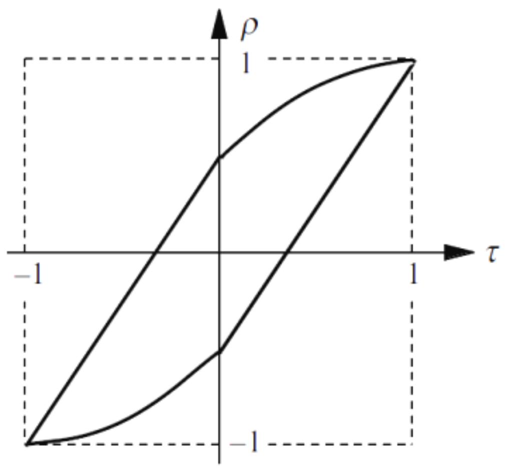


Figure 3.3: Relationship between Kendall's τ and spearman's ρ

3.4.4 Tail Dependence

In most financial applications, it is helpful to find the relation from the points of two random variables instead of focusing on the strong relation between every two points. The, tail dependence quantifies the properties of two random variables at the extreme movements; this provides a good measure of extreme risk.

Let X and Y be continuous random variables with distribution functions F and G, respectively. The upper tail dependence denoted as λ_u is defined as the probability of extremes in *Y* occurring, conditioned on the presence of extremes in *X*, that is,

$$\lambda_u = \lim_{x \to 1^-} P\left[Y > G^{-1}(t) \mid X > F^{-1}(t)\right] = \lim_{u \to 1^{-1}} \frac{\overline{C}(u,u)}{(1-u)}.$$

The joint distribution is said to be asymptotically dependent if $\lambda_u \in (0,1],$ and independent if $\lambda_u = 0.$ When C is Archimedean with generator φ, the upper tail dependence can be expressed as

$$\lambda_u = 2 - \lim_{t \to 0^+} \frac{1-\varphi^{[-1]}(2t)}{1-\varphi^{[-1]}(t)}.$$

Likewise, the lower tail dependence parameter λ_l is the limit (if it exists) of the conditional probability that *Y* is less than or equal to the $100t$-th percentile of G given that *F is* less than or equal to the $100t$-th percentile of *F* as t approaches 0, that is,

$$\lambda_l = \lim_{x \to 0^+} P[Y \le G^{-1}(t) \mid X \le F^{-1}(t)] = \lim_{u \to 0^+} \frac{C(u,u)}{u}.$$

Again, when C is Archimedean with generator φ, the lower tail dependence is expressible as (Nelsen, 2006),

$$\lambda_l = \lim_{t \to \infty} \frac{\varphi^{[-1]}(2t)}{\varphi^{[-1]}(t)}.$$

3.5 Trigonometric Dependence Measure

Using the definition of measure of dependences given in subsection 3.4, we next define the dependence measure of the proposed trigonometric functions.

3.5.1 Cot Copula

For the Cot generator $\varphi(t)=\cot^{\theta}\left(\frac{\pi}{2}t\right)$ define in chapter 2 the upper and lower tail dependence (λ_u and λ_l respectively) is defined as follows:

$$\lambda_u = 2-\lim_{t\to 0^+}\frac{1-\varphi^{[-1]}(2t)}{1-\varphi^{[-1]}(t)} = 2-2^{\frac{1}{\theta}},$$

$$\lambda_l = \lim_{t\to\infty}\frac{\varphi^{[-1]}(2t)}{\varphi^{[-1]}(t)} = 2^{-\frac{1}{\theta}}.$$

The Kendall's τ for the generator $\varphi(t)=\cot^{\theta}\left(\frac{\pi}{2}t\right)$, is then given by

$$\tau = 1+4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1+4\int_0^1 \frac{\cot\left(\frac{\pi}{2}t\right)}{\left(-\theta\frac{\pi}{2}\right)\left(1+\cot^2\left(\frac{\pi}{2}t\right)\right)}dt = 1-\frac{8}{\pi^2\theta}.$$

From this expression, the Cot-copula function has a range of dependency between $\left[1-\frac{8}{\pi^2\theta},1\right]$.

3.5.2 Cot II Copula

According to the information in previous section, for the Cot II generator $\varphi(t)=\cot\left(\frac{\pi}{2}t^{\theta}\right)$, the upper and lower tail dependence (λ_u and λ_l respectively) is defined as follows:

$$\lambda_u = 2-\lim_{t\to 0^+}\frac{1-\varphi^{[-1]}(2t)}{1-\varphi^{[-1]}(t)}=0,$$

$$\lambda_l = \lim_{t\to\infty}\frac{\varphi^{[-1]}(2t)}{\varphi^{[-1]}(t)}=1.$$

The Kendall's τ for the generator $\varphi(t)=\cot(\frac{\pi}{2}t^{\theta})$, is then given by:

$$\tau_{X_1,X_2} = 1+4\int_0^1\frac{\varphi(t)}{\varphi^{'}(t)}dt = 1+4\int_0^1\frac{\cot\left(\frac{\pi}{2}t^{\theta}\right)}{-\theta\frac{\pi}{2}t^{\theta-1}\csc^2\left(\frac{\pi}{2}t^{\theta}\right)}dt$$

$$=1-\frac{4}{\pi\theta}\int_0^1\frac{\sin\left(\pi t^{\theta}\right)}{t^{\theta-1}}dt \tag{3.13}$$

3.5.3 Csc Copula

In case of Csc function with generator $\varphi(t)=\left(\csc(\frac{\pi}{2}t)-1\right)^{\theta}$, its tail dependence is calculated by following formula:

$$\lambda_u = 2-\lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right)=2-2^{\frac{1}{2\theta}}$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 2^{\frac{-1}{\theta}}$$

According to the definition of Kendall's τ for two random variables X and Y we have following formula for $\varphi(t) = \left(\csc\left(\frac{\pi}{2}t\right) - 1\right)^{\theta}$:

$$\tau_{X,Y} = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 + 4\int_0^1 \frac{\left(\csc(\frac{\pi}{2}t) - 1\right)^{\theta}}{\left(-\frac{\pi}{2}\theta\right)\left(\csc\left(\frac{\pi}{2}t\right) - 1\right)^{\theta-1}\csc\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right)}dt$$

$$= 1 + 4\int_0^1 \frac{\left(\csc\left(\frac{\pi}{2}t\right) - 1\right)}{\left(-\frac{\pi}{2}\theta\right)\csc\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right)}dt$$

$$= 1 - \frac{8}{\pi\theta}\left[\int_0^1 \tan\left(\frac{\pi}{2}t\right) - \tan\left(\frac{\pi}{2}t\right)\sin\left(\frac{\pi}{2}t\right)dt\right]$$

$$= 1 + \frac{16}{\pi^2\theta}\left[\ln\left|\cos(\frac{\pi}{2}t)\right| + \ln\left|\sec\left(\frac{\pi}{2}t\right) + \tan\left(\frac{\pi}{2}t\right)\right| - \sin\left(\frac{\pi}{2}t\right)\right]_0^1$$

$$= 1 + \frac{16}{\pi^2\theta}\left(\ln(2) - 1\right)$$

According to the Kendall's τ the dependence coverage for the Csc copula is $[0.0051,\ 1]$ which provide a wider range of dependence compared to the Cot copula.

3.5.4 CscII Copula

Tail dependence for CscII with generator function $\varphi(t) = \csc\left(\frac{\pi}{2}t^{\theta}\right) - 1$ is calculated as

$$\lambda_u = 2 - \lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right) = 2-\sqrt{2},$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 2^{-\frac{1}{\theta}}.$$

And in case of Kendall's τ for this function with generator function

$$\varphi(t) = \csc\left(\frac{\pi}{2}t^{\theta}\right) - 1:$$

$$\tau_{X_1,X_2} = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 + 4\int_0^1 \frac{\csc\left(\frac{\pi}{2}t^{\theta}\right)-1}{-\frac{\pi\theta}{2}t^{\theta-1}\csc\left(\frac{\pi}{2}t^{\theta}\right)\cot\left(\frac{\pi}{2}t^{\theta}\right)}dt$$

$$= 1 - \frac{8}{\pi\theta}\int_0^1 \frac{\csc\left(\frac{\pi}{2}t^{\theta}\right)-1}{t^{\theta-1}\csc\left(\frac{\pi}{2}t^{\theta}\right)\cot\left(\frac{\pi}{2}t^{\theta}\right)}dt \qquad (3.14)$$

$$= 1 - \frac{8}{\pi\theta}\left(\int_0^1 \frac{1}{t^{\theta-1}\cot\left(\frac{\pi}{2}t^{\theta}\right)}dt - \int_0^1 \frac{1}{t^{\theta-1}\csc\left(\frac{\pi}{2}t^{\theta}\right)\cot\left(\frac{\pi}{2}t^{\theta}\right)}dt\right).$$

3.5.5 CscIII Copula

The tail dependence for CscIII with generator function $\varphi(t) = \csc^{\theta}\left(\frac{\pi}{2}t\right) - 1$ is given by:

$$\lambda_u = 2 - \lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right) = 2-\sqrt{2}$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 2^{-\frac{1}{\theta}}$$

The Kendall's τ for this generator $\varphi(t) = \csc^{\theta}\left(\frac{\pi}{2}t\right) - 1$:

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt$$

$$= 1 + 4\int_0^1 \frac{\csc^{\theta}\left(\frac{\pi}{2}t\right) - 1}{\left(-\theta\frac{\pi}{2}\right)\csc^{\theta}\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right)}dt \tag{3.15}$$

$$= 1 - \frac{8}{\pi\theta}\int_0^1 \frac{\csc^{\theta}\left(\frac{\pi}{2}t\right) - 1}{\csc^{\theta}\left(\frac{\pi}{2}t\right)\cot\left(\frac{\pi}{2}t\right)}dt.$$

3.5.6 Gumbel II Copula

Tail dependence for Gumbel II with generator function $\varphi(t) = -\ln(t^{\theta})$ is given by:

$$\lambda_u = 2 - \lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right) = 2 - \frac{\theta+2}{\theta+1},$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 1.$$

Using the definition of Kendall's τ for Gumbel II, we have

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 + 4\int_0^1 \frac{t\ln\left(t^{\theta}\right)}{\theta}dt = 0$$

According to the Kendall's τ the dependence coverage for the Gumbel II- copula is null.

3.5.7 Gumbel Copula

The tail dependence for Gumbel with generator function $\varphi(t) = (-\ln t)^{\theta}$ and inverse $\varphi^{-1}(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$ is

$$\lambda_u = 2 - \lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right) = 2 - 2^{\frac{1}{\theta}},$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 0.$$

The Kendall's τ for the Gumbel is given by

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 + 4\int_0^1 \frac{t(-\ln t)^{\theta}}{-\theta(-\ln t)^{\theta-1}}dt = \frac{\theta-1}{\theta}.$$

Thus, the dependence coverage for Gumbel copula is [0, 1]. Though it has a wider coverage, the lower tail dependence is always constant, that is zero (independent of θ).

3.5.8 Clayton Copula

The tail dependence for Clayton with generator function $\varphi(t) = (t^{-\theta} - 1)$ is calculated as

$$\lambda_u = 2 - \lim_{t\to 0^+}\left(\frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)}\right) = 0,$$

$$\lambda_l = \lim_{t\to\infty}\left(\frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)}\right) = 2^{\frac{-1}{\theta}}.$$

The corresponding Kendall's τ the formula for Clayton is calculated as

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 - 4\int_0^1 \frac{\left(t^{-\theta} - 1\right)}{\theta\, t^{-\theta-1}}dt = \frac{\theta}{\theta+2}.$$

Accordingly the dependence coverage for Clayton is [0, 1]. Though it has a wider coverage, the upper tail dependence is always zero (independent of θ).

3.5.9 12th Family of Archimedean Copula (A12)

According to the Table 4.1 of Nelsen (Nelsen, 2006), the 12th families of Archimedean copula has both the upper and lower tail dependence. The tail dependence for A12 with generator function $\varphi(t) = \left(\frac{1}{t} - 1\right)^{\theta}$ is given by

$$\lambda_u = 2 - \lim_{t\to 0^+} \frac{1-\varphi^{[-1]}(2t)}{1-\varphi^{[-1]}(t)} = 2 - 2^{\frac{1}{\theta}}.$$

$$\lambda_l = \lim_{t\to\infty} \frac{\varphi^{[-1]}(2t)}{\varphi^{[-1]}(t)} = 2^{-\frac{1}{\theta}}.$$

The Kendall's τ for 12th family is:

$$\tau = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt = 1 + 4\int_0^1 \frac{\left(\frac{1}{t} - 1\right)^{\theta}}{(-\theta)\left(\frac{1}{t} - 1\right)^{\theta-1}\frac{1}{t^2}}\,dt = 1 - \frac{4}{6\theta}.$$

Therefore, the dependence coverage for 12th family is [0.34, 1].

Information regarding the lower and the upper tail dependences together with Kendall's τ to compare the dependence coverage of these copulas is summarized in Table 3.2.

Table 3.2: Both lower and upper tail dependence with dependence coverage for trigonometric copulas

Trigonometric copula	Lower tail	Upper Tail	Kendall's τ	Dependence Coverage
Cot	$2^{-\frac{1}{\theta}}$	$2-2^{\frac{1}{\theta}}$	$1-\frac{8}{\pi^2\theta}$	[0.19,1]
Cot II	1	0	Ref. (3.13)	[0.19,1]
CSC	$2^{\frac{-1}{\theta}}$	$2-2^{\frac{1}{2\theta}}$	$1+\frac{16}{\pi^2\theta}(\ln(2)-1)$	[0.005,1]
CSC II	$2^{-\frac{1}{\theta}}$	$2-\sqrt{2}$	Ref. (3.14)	[0.5, 1]
CSC III	$2^{-\frac{1}{\theta}}$	$2-\sqrt{2}$	Ref. (3.15)	[0.5, 1]
Gumbel	0	$2-2^{\frac{1}{\theta}}$	$\frac{\theta-1}{\theta}$	[0,1]
Gumbel II	1	$2-\frac{\theta+2}{\theta+1}$	0	0
Clayton	$2^{-\frac{1}{\theta}}$	0	$\frac{\theta}{\theta+2}$	[0,1]
A12	$2^{-\frac{1}{\theta}}$	$2-2^{\frac{1}{\theta}}$	$1-\frac{4}{6\theta}$	[0.34, 1]

Since we are interested in those copulas with both upper and lower tail dependences, we first choose a proper copula based on tail dependence. The Cot-copula and CSC-copula have flexible upper and lower tail dependences which is comparable with A12. Both the Cot-copula and CSC-copula have better dependence coverage than A12. Finally, we can conclude that CSC-copula is more superior in terms of tail dependence and also dependence coverage. An additional advantage of the CSC-copula is that it is a one-parameter copula with bivariate tail dependences. Figure 3.4 shows the differences between Clayton, Gumbel and Cot copula.

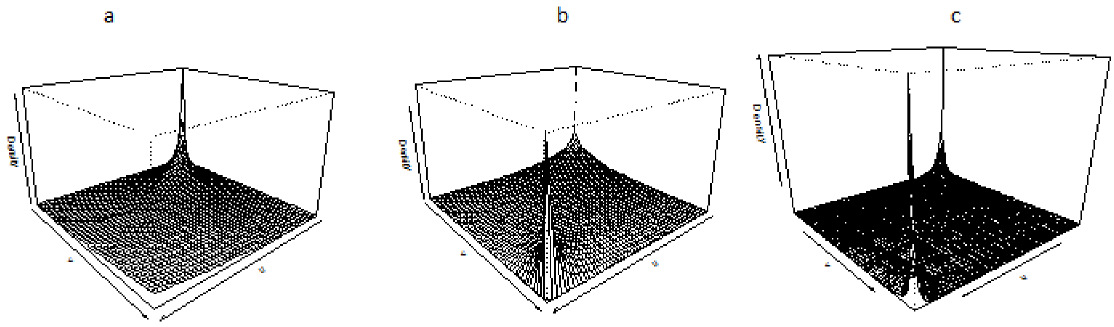


Figure 3.4: Tail dependence for (a) Clayton, (b) Gumbel and (c) Cot copula.

3.6 Illustrative Examples

In this section, we analyze a large bivariate data set to demonstrate the ability of the new bivariate copula to capture tail dependences in symmetric and asymmetric data sets.

Symmetric data are generated from t-Copula; specifically from t_2 , t-copula with 2 degree of freedem, and t_4, t-copula with 4 degree of frredem, with correlation coefficient in the range [0, 1). Our results showed that when the correlation coefficient is big, $\rho > 0.5$, with heavier tails, the Cot-copula and CSC-copula provide the best estimates among these copulas (in terms of GOF and tail dependences). Details of the lower and upper tail dependence calculation for t-Copula can be found in (Genest & MacKay, 1986b). Although the Gumbel and the Clayton capture a good range of tail dependence, they have the limitation of dealing with just one tail. The 12th family of Archimedean copula captures both upper and lower tail dependence but with a smaller range of dependence; examples given for *t*(0.5, 2) and *t*(0.5, 4). The problem of over-estimation of lower tail dependence for the Cot-copula under a small correlation coefficient $\rho \leq 0.5$ is attributed to the fact that the lower tail coverage is [0.5, 1] while in case of CSC-copula, the lower tail coverage is [0.25, 1]. Table 3.3 up to Table 3.9 show the results of estimation in different ranges of tail dependence. Graphical representations of goodness of fit based on Kendall's process are shown in Figure 3.5 up to Figure 3.11. The Cot-copula, in general, provides good tail estimates when the underlying distribution is heavy-tail with correlation coefficient ranging from moderate to large. This result is even better in the case of CSC-copula since the dependence coverage is wider.

Table 3.3: Estimated copulas with 1000 data from t-copula $t\,(0.9, 2)$

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.72	0.72		
Gumbel	0.0	0.7671	0.0002	3.3107
Clayton	0.8127	0	0.0013	3.3429
12th family	0.8007	0.7511	0.0006	3.1189
Cot	0.7746	0.7090	0.0001	2.7141
CSC	0.65815	0.7673	0.0001	1.6570
Csc2	0.61491	0.5857	0.0006	2.8508
Csc3	0.78255	0.5857	0.0006	2.8270

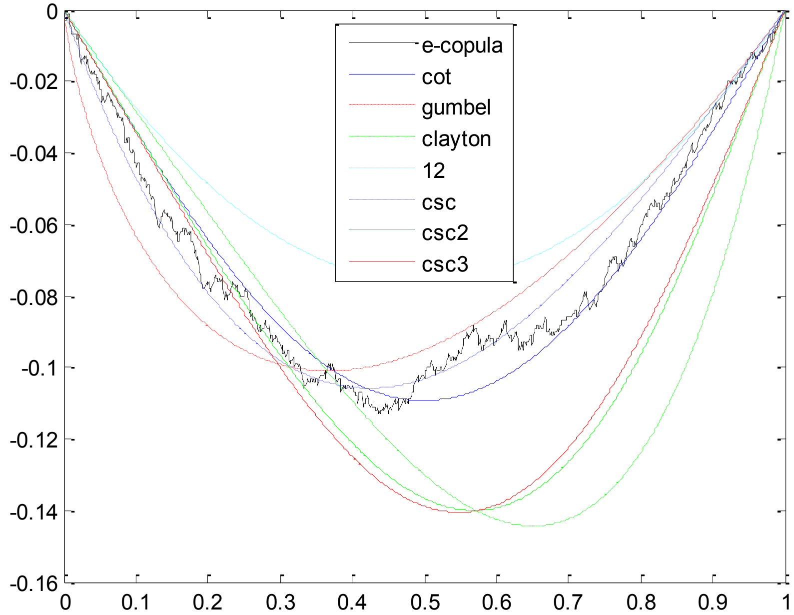


Figure 3.5: Goodness of fit representation of Table 3.3.

Table 3.4: Estimated copulas with 1000 data from t-copula *t* (0.5, 4)

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.25	0.25		
Gumbel	0	0.4058	0.0001	1.4862
Clayton	0.3520	0	0.0012	0.6638
12^{th} family	0.6876	0.5457	0.0068	1.8507
Cot	0.5651	0.2304	0.0012	1.2145
CSC	0.3882	0.3950	0.0003	0.7326
Csc2	0.0343	0.5858	0.0011	0.4112
Csc3	0.1136	0.5858	0.0004	0.3186

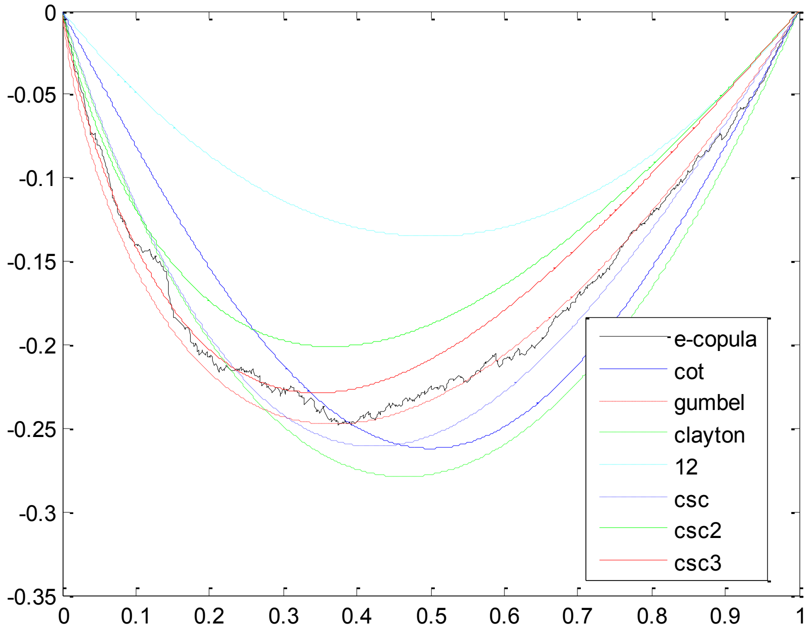


Figure 3.6: Goodness of fit representation of Table 3.4.

Table 3.5: Estimated copulas with 1000 data from t-copula, $t\,(0\,,\,2)$

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.18	0.18		
Gumbel	0	0.0913	0.0007	1.0723
Clayton	0.0019	0	0.0006	0.1108
12^{th} family	0.6502	0.4620	0.0268	1.6103
Cot	0.5114	0.0448	0.0066	1.0338
CSC	0.2853	0.1279	0.0016	0.5527
Csc2	0.0367	0.5858	0.0158	0.4196
Csc3	0.0000	0.5858	0.0085	0.0206

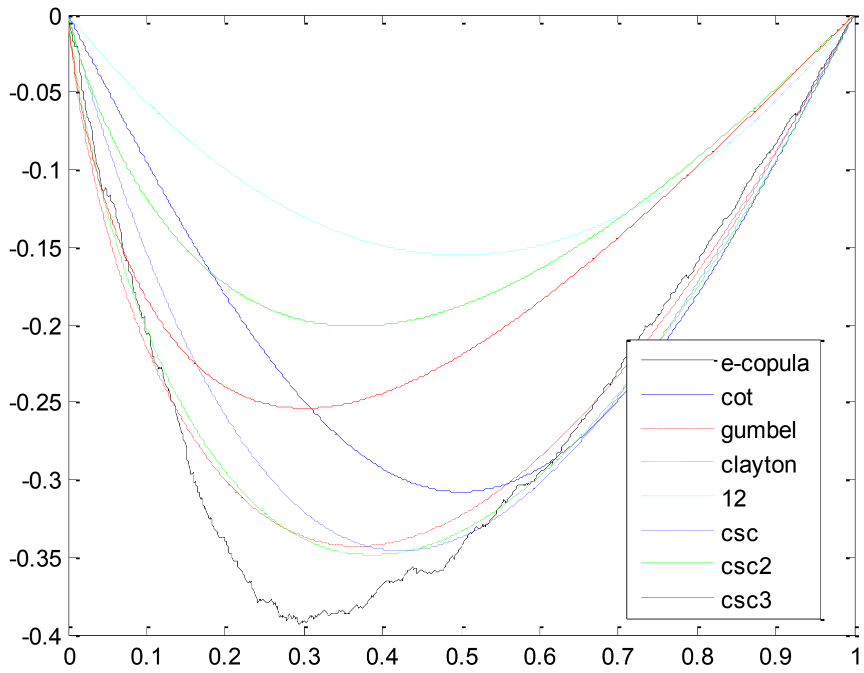


Figure 3.7: Goodness of fit representation of Table 3.5.

Table 3.6: Estimated copulas with 1000 data from t-copula, *t* (0.5, 4).

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.08	0.08		
Gumbel	0.0	0.0054	0.0001	1.0039
Clayton	0.0000	0.0	0.0001	0.0049
12^{th} family	0.6468	0.4539	0.0282	1.5909
Cot	0.5015	0.0060	0.0064	1.0044
CSC	0.2668	0.0640	0.0012	0.5246
Csc2	0.0338	0.5858	0.0173	0.4093
Csc3	0.0	0.5858	0.0093	0.0000

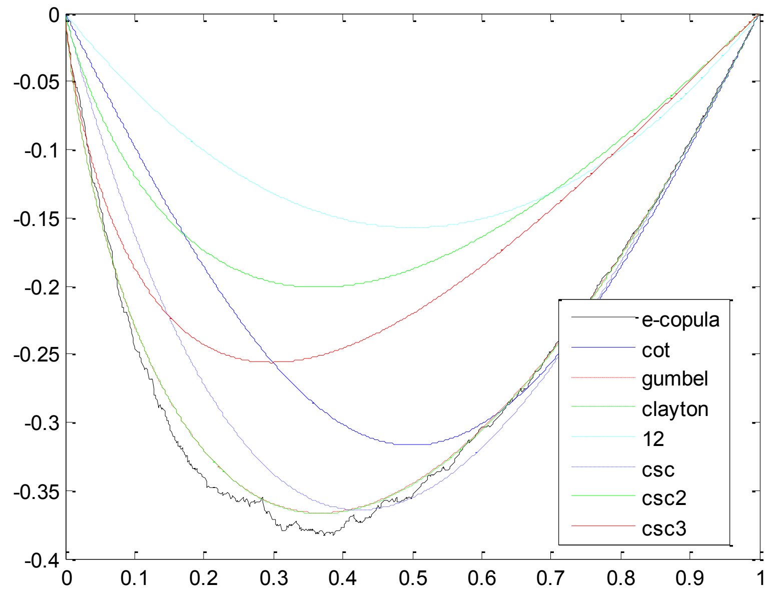


Figure 3.8: Goodness of fit representation of Table 3.6.

The asymmetric data has been generated from the Joe –Clayton copula. The results illustrates that for heavy asymmetric tail dependence, Cot-copula and CSC-copula serve a good one parameter distribution functions. Also from the Table 3.7 , we note that when the lower tail is heavier than the upper tail, the Cot-copula gives better coverage.

Table 3.7: Estimated copulas with 1000 data from BB7(2,2)

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.7071	0.5858		
Gumbel	0.000	0.6517	0.0007	2.3197
Clayton	0.7573	0.000	0.0005	2.4939
12^{th} family	0.7584	0.6814	0.0020	2.5061
Cot	0.7083	0.5881	0.0001	2.0094
CSC	0.5579	0.6611	0.0002	1.1876
Csc2	0.5232	0.5858	0.0001	2.1399
Csc3	0.7211	0.5858	0.0001	2.1202

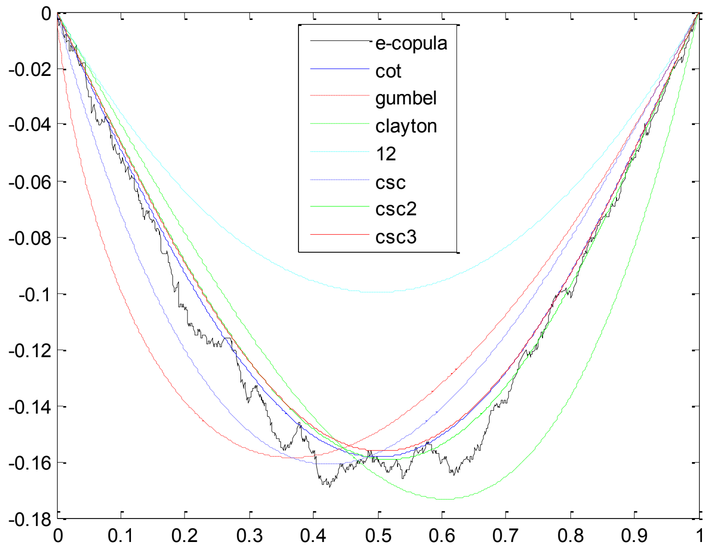


Figure 3.9: Goodness of fit representation of Table 3.7.

For the cases of heavy lower and upper tail dependence, the results are satisfactory for both the Cot and CSC copulas as shown in Table 3.8.

Table 3.8: Estimated copulas with 1000 data from BB7(3,2)

	Lower tail	Upper tail	Goodness of fit	parameter
Real data simulated	0.7071	0.7401		
Gumbel	0.000	0.7417	0.0002	3.0171
Clayton	0.7786	0.000	0.0016	2.7691
12^{th} family	0.7850	0.7262	0.0009	2.8638
Cot	0.7546	0.6748	0.0001	2.4618
CSC	0.6316	0.7417	0	1.5084
Csc2	0.5359	0.5858	0.0006	2.2226
Csc3	0.7331	0.5858	0.0005	2.2328

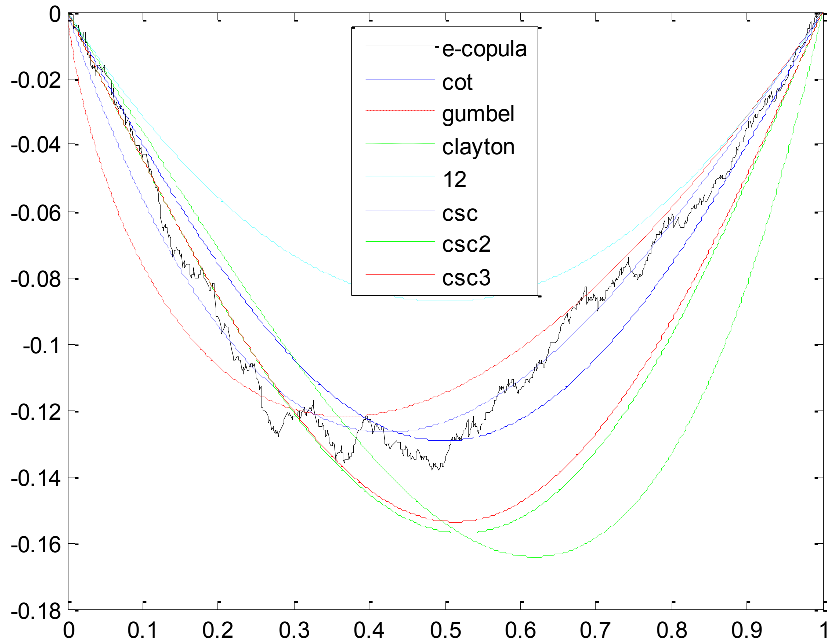


Figure 3.10: Goodness of fit representation of Table 3.8.

Our results support that the CSC-copula serves as an alternative proper distribution function when upper tail is heavier than the lower tail.

Table 3.9: Estimated copulas with 1000 data from BB7(4, 2)

	Lower tail	Upper tail	Goodness of fit	parameter
Simulated data	0.7071	0.8108		
Gumbel	0	0.7795	0.0002	3.4785
Clayton	0.7900	0	0.0018	2.9402
12^{th} family	0.7995	0.7493	0.0009	3.0981
Cot	0.7750	0.7096	0.0002	2.7187
CSC	0.6690	0.7774	0.0001	1.7242
Csc2	0.5660	0.5858	0.0008	2.4350
Csc3	0.7521	0.5858	0.0006	2.4334

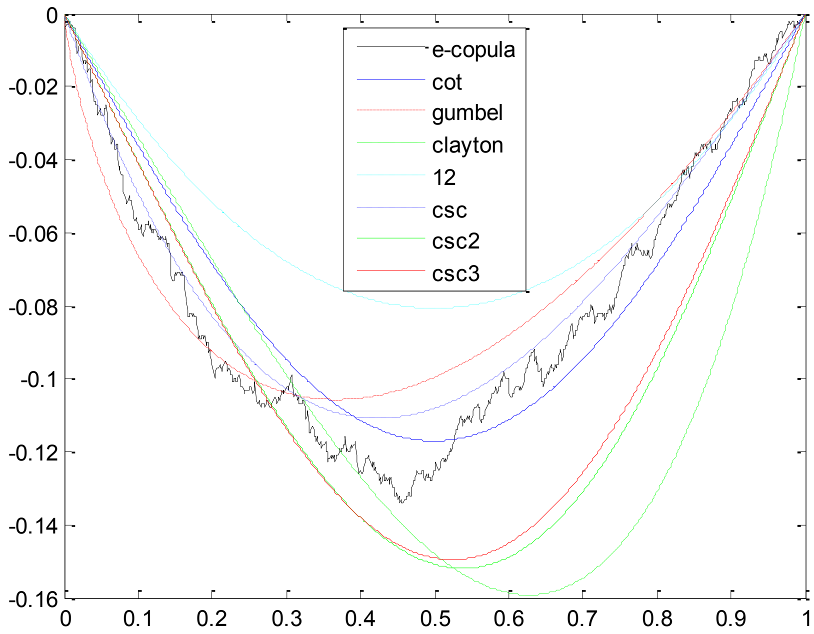


Figure 3.11: Goodness of fit representation of Table 3.9.

Chapter 4: Multivariate Copulas and Vine Structure

This chapter provides the concept of multivariate Archimedean copula together with the examination of the extension of Cot-copula and Csc-copula. Here, the problem of classical multivariate extension is considered as the starting point for introducing the vine structure and vine copula. We then extend the bivariate trigonometric copula into multivariate copula by using vine structure.

4.1 Multivariate Archimedean Copula

A bivariate Archimedean copula with a strict generator $\varphi:[0,1]\to[0,\infty]$ can be extended to $n-$dimensional copula $C:[0,1]^n \to [0,\infty]$ defined as

$$C(u)=\varphi^{-1}(\varphi(u_1),\varphi(u_2),\ldots,\varphi(u_n)).$$

This extension is possible if and only if φ^{-1} is entirely monotonic on R, i.e., if $\varphi^{-1}\in L_\infty$ with $L_m=\{\Phi:R_+\to[0,1]\,|\,\Phi(0)=1,\,\Phi(\infty)=0,\,(-1)^k\Phi^{(k)}(t)\geq 0,\,k\leq m\}$.

By considering this property, it is necessary then to check whether the Cot- and Csc-copula can be extended to multivariate dimension.

The inverse generator function of the Cot-copula is considered as $g(t)=\varphi^{-1}(t)=\dfrac{2}{\pi}arc\cot(t^{\frac{1}{\theta}})$. To examine whether *g(t)* is a complete monotonic function we proceed as follow:

$$g'(t)=\left(-\frac{2}{\pi\theta}\right)\left(t^{\frac{1}{\theta}-1}\right)\left(t^{\frac{2}{\theta}}+1\right)^{-1} \quad \to \quad -g'(t)\geq 0\ .$$

$$g''(t)=\left(-\frac{2}{\pi\theta}\right)\left(t^{\frac{1}{\theta}-2}\right)\left(t^{\frac{2}{\theta}}+1\right)^{-2}\left(\left(\frac{1}{\theta}-1\right)-\frac{2}{\theta}t^{\frac{2}{\theta}}\left(t^{\frac{2}{\theta}}+1\right)^{-1}\right)\quad\rightarrow\quad(-1)^2g''(t)\geq 0\ \ \theta\geq 1.$$

$$g'''(t)=\left(-\frac{2}{\pi\theta}\right)\left(A^{-2}\left(t^{\frac{1}{\theta}-2}\right)\left(\left(\frac{1}{\theta}-2\right)t-\frac{4}{\theta}t^{\frac{2}{\theta}}A^{-1}\right)B+\frac{4}{\theta^2}t^{\frac{3}{\theta}-3}A^{-3}\left(t^{\frac{2}{\theta}}A^{-1}-1\right)\right)\rightarrow\quad(-1)^3g'''(t)\leq 0.$$

where $B=\left(\frac{1}{\theta}-1\right)-\frac{2}{\theta}t^{\frac{2}{\theta}}\left(t^{\frac{2}{\theta}}+1\right)^{-1}$, $A=t^{\frac{2}{\theta}}+1$.

The same logic can be applied to the Csc-copula inverse function. The result would be similar as the Cot-copula, where the function is not extendable to higher dimensions.

Beside the fact that multivariate extension of trigonometric copulas is not possible, there are some more reasons for considering vine structure as a logic way of extension of multivariate copula.

As mentioned in (Joe, Li, & Nikoloulopoulos, 2010) a competent multivariate copula families has the following properties:

- Wide range of positive and negative dependence.
- Flexible range of upper and lower tail dependence.
- Computationally feasible density for (likelihood) estimation.
- Lower order margins belong to the same parametric family, means: Closure property under marginalization.

According to the literatures, none of the existing family of copulas satisfies all the conditions. Below some classified literatures is illustrated as proof.

The multivariate normal copula (Abdous et al., 2005; Fang et al., 2002) satisfied all properties except tail dependence. The multivariate t copula family (Demarta & McNeil, 2005) however does not have this problem. Multivariate t copula reflects the symmetric tail dependence. Thus for any bivariate margins, multivariate t copula resulted in the same value for the upper and lower tail dependence. Such advantage resulted in the extensive use of the t copula in the context of modeling multivariate financial return data (Breymann et al., 2003). However, the literatures reports that the financial data are asymmetric (Longin & Solnik, 2001), (Ang & Chen, 2002) and (Hong, Tu, & Zhou, 2007) and as such the t copula does not satisfy all the desired properties of multivariate copula family. There are some research done to improve the t copula for asymmetric data, which posses skewed tail dependence, for example (Demarta & McNeil, 2005), (Kotz & Nadarajah, 2004). However, this improvement resulted in expensive computations. The problem with Archimedean copula is related to the narrow range of negative dependence and exchangeable structure (Joe, 1997), (McNeil & Nešlehová, 2009). To overcome these problems, researchers extend the Archimedean copula to partially symmetric copula (Joe, 1993) and max-id copulas (Joe & Hu, 1996). Joe's (1993) proposal overcomes the problem of exchangeability but the problem with flexible tail dependence still persists. The max-id copula provides a flexible upper tail dependence but not the lower tail.

Arbitrary dimension problem of multivariate distribution can be overcome via vine structure. The vine copulas are a flexible graphical model for describing multivariate distributions built up using a cascade of bivariate copulas, so called pair-copulas. Because of this flexibility, the vine copulas do not face any of the previous multivariate copulas' problems by choosing appropriate bivariate copula (Brechmann & Schepsmeier).

By considering the advantages and the limitations in extending to multivariate copula, we now consider constructing the multivariate trigonometric copula via the vine structure.

The next part of this chapter covers the concept of vine structure and vine copula. In section 4.2 we provide an introduction of vine based on current available literatures. Section 4.3 provides precise mathematical definition of vine and vine copulas together with various types of vines. Section 4.4 presents copula vine or pair copula construction. The important dependence properties of copula will be given in section 4.5. The concept of copula estimation and model inference will be presented in section 4.6 and 4.7 respectively. Section 4.8 provides a structure of Archimedean vine copula with introducing trigonometric vines. Application of multivariate vine copula on US and Asia index will be presented in section 4.9.

4.2 Introduction to Vine

Vines are graphical structures that represent joint probability distributions. They were named for their close visual resemblance to grapes (see Figure 4.1) (Kurowicka & Joe, 2010).

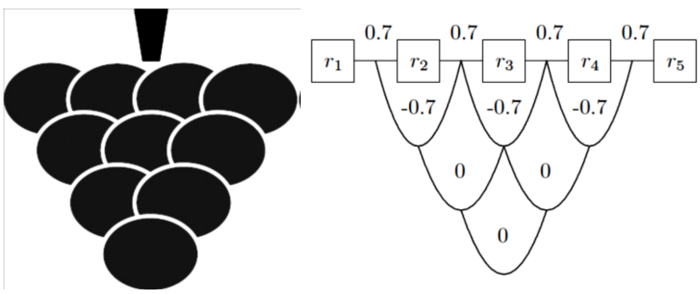


Figure 4.1: Comparison of vine structure with grapes.

An ordinary vine is a particular case for which all constraints are two-dimensional or conditional two-dimensional. Regular vines generalize trees, and are themselves specializations of something called Cantor trees (Kurowicka & Joe, 2010). The regular vine can be used successfully to model high-dimensional dependence together with copula.

Vine copula structure was first introduced by Joe (Joe, 1996) when he used the pairwise construction based on Sklar theorem (Sklar, 1959). Extension of this construction was done by Bedford and Cooks (Bedford & Cooke, 2001), (Bedford & Cooke, 2002). They used a graphical representation of the tree to make a multidirectional density by product of bivariate copulas, called pair-copula, instead of Sklar theorem. They called the structure as regular vine since it is based on graphical trees. Gaussian copula was used as bivariate copula in their structure. Using arbitrary pair-copula for the first time was conducted by Aas et.al,. (Aas, Czado, Frigessi, & Bakken, 2009). They developed standard Maximum Likelihood (ML) estimation for C- and D-vine copulas. The importance of C- and D-vine copulas are shown later by (Czado, 2010) where he shows that can be constructed in a simple recursive condition.

4.3 Definition and Concepts

A vine *v* on *N* variables is a nested group of trees *T*, where the edges of tree j, E_j are the nodes of tree $j+1$; $j=1,...,N-2$ and every tree contains the most number of edges (Kurowicka & Cooke, 2003). An ordinary vine on n variables is a vine in which two edges in tree j are combined by an edge in tree $j+1$ only if these edges share a common node, $j = 1, \ldots, N-2$. The formal definition follow according to (Kurowicka & Cooke, 2003; Kurowicka & Joe, 2010)

Definition (Regular Vine): v is a regular vine on n elements with edges $E(v) = E_1 \cup \ldots \cup E_{n-1}$ if

1) $v = \{T_1, \ldots, T_{n-1}\}$,

2) T_1 is a connected tree with nodes $N_1 = \{1, \ldots, n\}$ and edges E_1;and for $i = 2, \ldots, n-1$, T_i is a tree with nodes $N_i = E_{i-1}$.

3) (Proximity) for $i = 2, \ldots, n-1$, $\{a, b\} \in E_i, \#(a \Delta b) = 2$, where Δ denotes the symmetric differences operator . To put it another way, if a and b are nodes of T_i linked by an edge where $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$, then precisely one of the a_i are equivalent to one of the b_i and $\#$ denotes the cardinality of a set.

Definition (C-vine): A regular vine is labeled a canonical or C-vine if each tree T_j has a unique node of degree $n-1$, thus has the highest possible degree.

Definition (D-vine): A regular vine is labeled a D-vine if all nodes in T_1 has a degree of two or less.

The three important concepts known as conditioning, constraint and conditioned set of an edge is defined as follows:

Definition (Constraint, Conditioning and Conditioned set):

1) For an element, $e \in E_i, i \leq n-1$, the **Constraint set** associated with e is **the complete union** U_e^* of e, that is, the subset of $\{1,...,n\}$ reachable from e by the membership relation.

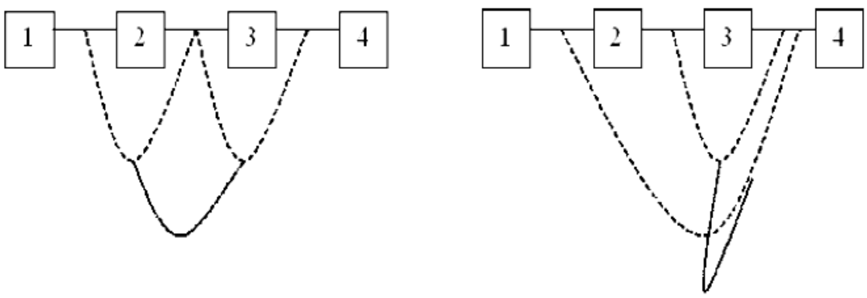
2) For $i = 1,...,n-1$, $e \in E_i$ if $e = \{j,k\}$ then the **conditioning set** associate with e is $U_e = U_j^* \cap U_k^*$

3) The **conditioned set** associated with e is $\{C_{e,j}, C_{e,k}\} = \{U_j^* \setminus D_e, U_k^* \setminus D_e\}$.

Definition (m-child, m-descendent): *If node e is an element of node f, we say that e is an* **m-child** *of f; similarly, if e is reachable from f via the membership relation: $e \in ... \in f$ we say that e is an* **m-descendent** *of f.*

To visualize the concept of the above definitions, we construct the following examples:

Example 4.1 (Non Regular vine): Figure 4.2 visualizes a regular and a non-regular vine. Figure 4.2(b) is not regular, because in T_2, with edges $a = \{1, 2\}$ and $b = \{3, 4\}$ there is not any common nodes in tree T_1. Accordingly Figure 4.2 (a) is a regular vine.



(a) (b)

Figure 4.2: Regular Vine (a) and Non Regular Vine (b)

Example 4.2 (C-vine): Figure 4.3 shows an example of C-vine with 5 nodes, n=5, with following trees:

a) $T_1:$ $N_1 = \{1,2,3,4,5\}$ and $E_1 = \{\{1,2\},\{1,3\},\{1,4\},\{1,5\}\} = \{1,2\,;1,3\,;1,4\,;1,5\}$,

b) $T_2:$ $N_2 = E_1$ and $E_2 = \{\{12,13\},\{12,14\},\{12,15\}\} = \{2,i\,|\,1, i = 3,4,5\}$,

c) $T_3:$ $N_3 = E_2$ and $E_3 = \{2,3\,|\,1\,;2,4\,|\,1\},\{2,3\,|\,1\,;2,5\,|\,1\}\} = \{3,i\,|\,1,2, i = 4,5\}$,

d) $T_4:$ $N_4 = E_3$ and $E_4 = \{3,4\,|\,1,2\,;3,5\,|\,1,2\} = \{4,5\,|\,1,2,3\}$.

For edges $\{3,2\,|\,1\}$. The Conditioning set is: $\{1,2\} \cap \{1,3\} = \{1\}$, and the Conditioned set is $\{1,2\} \setminus \{1\} \cap \{1,3\} \setminus \{1\} = \{2,3\}$,

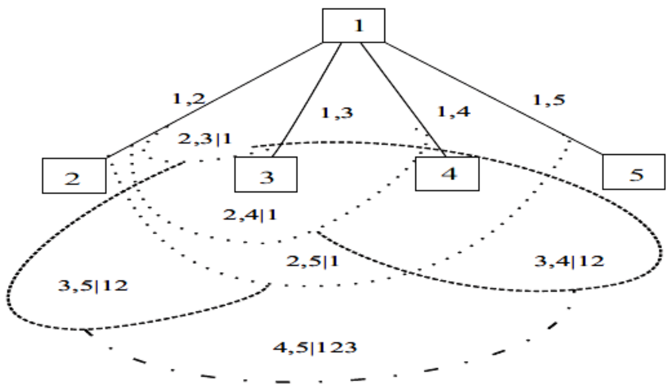


Figure 4.3: C-Vine with 5 Nodes

Example 4.3 (D-vine): Figure 4.4 shows a D-vine structure with five nodes, n=5 with following trees

a) $T_1:$ $N_1 = \{1,2,3,4,5\}$ and $E_1 = \{\{1,2\},\{2,3\},\{3,4\},\{4,5\}\} = \{1,2\,;2,3\,;3,4\,;4,5\}$,

b) $T_2: N_2 = E_1$ and $E_2 = \{\{12,23\},\{23,34\},\{34,45\}\} = \{i,i+2\,|\,i+1, i = 1,2,3,4\}$,

c) $T_3:$ $N_3 = E_2$ $E_3 = \{\{1,3\,|\,2\,;2,4\,|\,3\},\{2,4\,|\,3\,;3,5\,|\,4\}\} = \{i,i+3\,|\,i+1,i+2, i = 1,2\}$,

d) $T_4: N_4 = E_3$ and $E_4 = \{1,4 \mid 2,3 \,;\, 2,5 \mid 3,4\} = \{1,5 \mid 2,3,4\}$.

The Conditioning set is: $\{1,2\} \cap \{2,3\} = \{2\}$, and the Conditioned set is $\{1,2\} \setminus \{2\} \cap \{2,3\} \setminus \{2\} = \{1,3\}$,

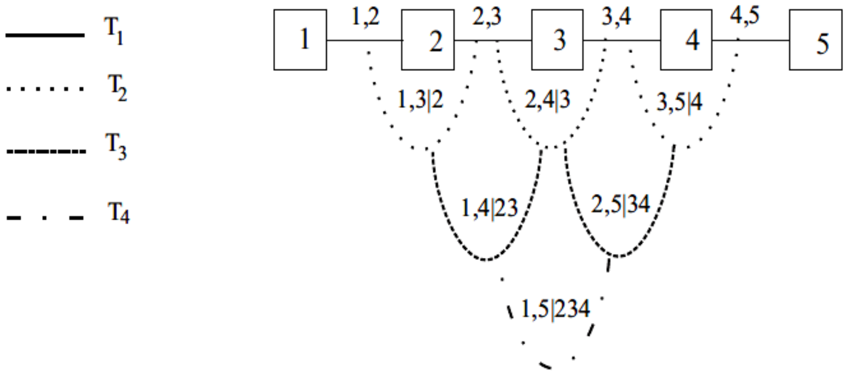


Figure 4.4: D-Vine with 5 nodes

Considering the vine definitions, we present some important properties of vines.

Properties 4.1: Let $v = \{T_1, \ldots, T_{n-1}\}$, be a regular vine then

1) The number of edges is $n(n-1)/2$,

2) Each pair of variables occurs exactly once as a conditioned set, this property is called doubleton,

3) If two edges have the same conditioning set, then they are the same edge.

Properties 4.2: For any node *K* of order *k > 0* in a regular vine, if variable i is a member of the conditioned set of *K*, then i is a member of the conditioned set of exactly one of the m-children of *K*, and the conditioning set of an m-child of *K* is a subset of the conditioning set of *K*.

Theorem 4.1(Bedford & Cooke, 2002) Bedford & Cooke, 2002)**:**

1) For any regular vine on $n-1$ elements, the number of regular $n-$dimensional vines which extend this vine is 2^{n-3}.

2) There are $\binom{n}{2} \times (n-2)! \times 2^{(n-2)(n-3)/2}$ labeled regular vines in total.

Interestingly all vines are in the same class when $n=3$. All regular vines for $n=4$ are C- or D-vines. But for $n=5$, there are many vines that are not either C or D-vines.

4.4 Copula Vine or Pair Copula Construction

Although the idea of copula vine started from pair copula decomposition from Sklar's theorem, the abstract breakthrough of constructing of multivariate copula is based on the vine structure. The graphical vine structure gives a bigger perspective of how multivariate copula can be constructed from a bivariate copula. In this section, we introduce the general idea and the theory of vine copula, and then we proceed to the concept of C- and D-vine copula. Finally, we touch the starting point of this idea by revising the concept of copula decomposition.

By Skalar's theorem, the world of statistical model has changed due to the breakthrough idea of decomposition of margins and dependence between random variables. The success of copula in bivariate rather than multivariate case is due to the shadow area in dependence. This shadow area is covered by concentrating mainly on the structure of dependence which is represented by vine as a graphical structure.

A vine copula is a specific type of regular vine which is constructed by assigning a bivariate copula to each edge in the $E(v) = E_1 \cup ... \cup E_{n-1}$. The set of $\binom{n}{2}$ copula is denoted by B which can be chosen independently from each other.

Theorem 4.2 gives the structure of the joint density of a regular vine copula with margins, $F_1,...,F_n$.

Theorem 4.2 (Bedford & Cooke, 2002) : Let $v = \{v_1, v_2, ..., v_{n-1}\}$ be a regular vine on n elements. For an edge $e \in E(v)$ with conditioned elements e_1, e_2 and conditioning set D_e, let the conditioning copula and its density be $C_{e_1,e_2|D_e}$ and $c_{e_1,e_2|D_e}$ respectively. For a given marginal distributions F_i with densities $f_i, i = 1,...,n$, the vine dependent distribution is uniquely determined with density given by:

$$f_{1,...,n} = f_1 ... f_n \prod c_{e_1,e_2|D_e}(F_{e_1|D_e}, F_{e_2|D_e}).$$

Vine copula have closed form densities when $F_1,...,F_n$ and the bivariate copula in B are differentiable.

Following examples are vine copula based on C- and D- vine structure which is defined in section 4.3.

Example 4.4: Consider a C-vine with 5 nodes as previously discussed in example 4.2. The set of bivariate copulas on every tree denoted separately as

a) $$T_1 : \begin{array}{l} E_1 = \{\{1,2\},\{1,3\},\{1,4\},\{1,5\}\} = \{12,13,14,15\} \\ B_1 = \{C_{12}, C_{13}, C_{14}, C_{15}\}, \end{array}$$

b) $T_2:$ $E_2 = \{\{12,13\},\{12,14\},\{12,15\}\} = \{2,i\,|\,1, i=3,4,5\}$
$B_2 = \{C_{23|1}, C_{24|1}, C_{25|1}\}$,

c) $T_3:$ $E_3 = \{2,3\,|\,1;2,4\,|\,1\},\{2,3\,|\,1;2,5\,|\,1\}\} = \{3,i\,|\,1,2, i=4,5\}$
$B_3 = \{C_{34|12}, C_{35|12}\}$,

d) $T_4:$ $E_4 = \{3,4\,|\,1,2;3,5\,|\,1,2\} = \{4,5\,|\,1,2,3\}$
$B_4 = \{C_{45|123}\}$.

Then the set of all copulas on regular C-vine is denoted by $B = B_1 \cup B_2 \cup B_3 \cup B_4$. Note that F_i and $C_{e_1,e_2|D_e}$ are assumed to be differentiable with density $f_i, i = 1,...,n$ and $c_{e_1,e_2|D_e}$. Then according to theorem 4.2 the density function is given by

$$\begin{aligned} f_{1,\dots,5} = {} & f_1 \cdots f_5 \\ & .c_{12}(F_1,F_2).c_{13}(F_1,F_3).c_{14}(F_1,F_4).c_{15}(F_1,F_5) \\ & .c_{23|1}(F_{2|1}, F_{3|1}).c_{24|1}(F_{2|1}, F_{3|1}).c_{25|1}(F_{2|1}, F_{3|1}) \\ & .c_{34|12}(F_{3|12}, F_{4|12}).c_{35|12}(F_{3|12}, F_{5|12}) \\ & .c_{45|123}(F_{4|123}, F_{5|123}). \end{aligned} \tag{4.1}$$

Equation 4.1 is a five-dimensional distribution function according to the C-vine structure, defined from copula that joints the bivariate random variables.

Example 4.5: Consider D-vine structure with a five nodes. For a graphical representation, consider the example 4.3 from section 4.3. We define the set of all bivariate copulas for D-vine structure as follows:

a) $T_1:$ $E_1 = \{\{1,2\},\{2,3\},\{3,4\},\{4,5\}\} = \{1,2\,;\,2,3\,;\,3,4\,;\,4,5\}$
$B_1 = \{C_{12}, C_{23}, C_{34}, C_{45}\}$,

b) $T_2:$ $E_2 = \{\{1,2\,;\,2,3\},\{2,3\,;\,3,4\},\{3,4\,;\,4,5\}\} = \{i, i+2\,|\,i+1, i=1,2,3,4\}$
$B_2 = \{C_{13|2}, C_{24|3}, C_{35|4}\}$,

c) $T_3:$ $E_3 = \{\{1,3\,|\,2\,;\,2,4\,|\,3\},\{2,4\,|\,3\,;\,3,5\,|\,4\}\} = \{i, i+3\,|\,i+1, i+2, i=1,2\}$
$B_3 = \{C_{14|23}, C_{25|34}\}$,

d) $$T_4: \begin{array}{l} E_4 = \{1,4 \mid 2,3\,;\, 2,5 \mid 3,4\} = \{1,5 \mid 2,3,4\} \\ B_4 = \{C_{15|234}\}. \end{array}$$

The set of all bivariate copulas on regular D-vine with five nodes is denoted by $B = B_1 \cup B_2 \cup B_3 \cup B_4$. By assuming differentiability as in the previous example, the five-dimensional density function is given by equation 4.2 as

$$\begin{aligned} f_{1,\ldots,5} = {} & f_1 \ldots f_5 \\ & c_{12}(F_1, F)_2\, c_{23}(F_2, F_3)\, c_{34}(F_3, F_4)\, c_{45}(F_4, F_5) \\ & c_{13|2}(F_{1|2}, F_{3|2})\, c_{24|3}(F_{2|3}, F_{4|3})\, c_{35|4}(F_{3|4}, F_{5|4}) \\ & c_{14|23}(F_{1|23}, F_{4|23})\, c_{25|34}(F_{2|34}, F_{5|34}) \\ & c_{15|234}(F_{1|234}, F_{5|234}). \end{aligned} \tag{4.2}$$

Equation 4.2 is a five dimensional distribution function according to the D-vine structure based on D-vines and theorem 4.2.

Extension of C- and D-vine copula from five-dimension to n-dimension is straightforward. It merely involves the definition of bivariate copula set. In case of C-vine a general definition of all bivariate copula sets is defined as $\{C_{i_1 i_2|1,\ldots,i_1-1} : 1 \le i_1 < i_2 \le n\}$. This gives the following decomposition of a multivariate density, the C-vine density with root nodes $1,\ldots,n$.

$$f_{1,\ldots n}(x) = \prod_{k=1}^{n} f_k \times \prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{i,i+j|1:(i-1)} \left[F(x_i \mid x_1,\ldots,x_{i-1}),\, F(x_{i+j} \mid x_1,\ldots,x_{i-1}) \mid \theta_{i|,i+j|1:(i-1)}\right], \tag{4.3}$$

In case of D-vine is defined as $\{C_{i_1 i_2|i_1+1,\ldots i_2-1} : 1 \le i_1 < i_2 \le n\}$ which gives equation 4.4 as multivariate density.

$$f_{1,..n}(x) = \prod_{k=1}^{n} f_k \times$$
$$\prod_{i=1}^{n-1} \prod_{j=1}^{n-i} c_{j,j+i|j+1:(j+i-1)} \left[F(x_j \mid x_{j+1},...,x_{j+i-1}),\ F(x_{j+i} \mid x_{j+1},...,x_{j+i-1}) \mid \theta_{j,j+i|j+1:(j+i-1)}\right], \tag{4.4}$$

where $f_k, k = 1,...,n,$ denote the marginal distributions and $c_{i,i+j|1:(i-1)}$ and $c_{j,j+i|j+1:(j+i-1)}$ are bivariate marginal distribution functions with parameters $\theta_{i|,i+j|1:(i-1)}$ and $\theta_{j,j+i|j+1:(j+i-1)}$.

Although the vine structure gives a wide perspective of the concept of vine copula, the idea started from pair- copula decomposition by Joe (Joe & Hu, 1996) who constructed the family of some multivariate distributions which were later called as D-vine. In the following sections, we try to explain the idea of vine copula according to Joe's point of view.

The joint density function of n random variables can be used as a starting point of definition of pair copula decomposition. We consider n random variables $X = (X_1,..., X_n)$ with a joint density function $f(x_1,...,x_n)$. The density can be factories as

$$f(x_1,...,x_n) = f(x_n)\, f(x_{n-1} \mid x_n)\, f(x_{n-2} \mid x_{n-1}, x_n)...f(x_1 \mid x_2,...,x_n), \tag{4.5}$$

According to Sklar's theorem (Sklar, 1959), every multivariate distribution F with margins $F_1, F_2,..., F_n$ can be written as

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

For an absolutely continuous F with strictly increasing, continuous marginal density $F_1, F_2,..., F_n$ we have

$$f(x_1,...,x_n) = c_{12...n}(F_1(x_1), F_2(x_2),..., F_n(x_n)) \ f_1(x_1)...f_n(x_n)$$

The components of equation (4.5) can be rewritten based on copula for example in case of two variables $f(x_{n-1} \mid x_n) = c_{n,n-1}(F_n(x_n), F_{n-1}(x_{n-1})) f_{n-1}(x_{n-1})$. It is because

$$f(x_{n-1}, x_n) = c_{n,n-1}(F_n(x_n), F_{n-1}(x_{n-1})) \ f_{n-1}(x_{n-1}) \ f_n(x_n)$$

To illustrate, as the number of variables increases, we have several alternatives to decompose the conditional density of X_1 given X_2 and X_3 according to equations (4.6) and (4.7).

$$f(x_1 \mid x_2, x_3) = c_{12|3}(F_{1|3}(x_1 \mid x_3), F_{2|3}(x_2 \mid x_3)) \ f(x_1 \mid x_3), \tag{4.6}$$

or

$$f(x_1 \mid x_2, x_3) = c_{13|2}(F_{1|2}(x_1 \mid x_2), F_{3|2}(x_3 \mid x_2)).f(x_1 \mid x_2), \tag{4.7}$$

Further decomposition of (4.7) leads to

$$f(x_1 \mid x_2, x_3) = c_{13|2}(F_{1|2}(x_1 \mid x_2), F_{3|2}(x_3 \mid x_2)).c_{12}(F_1(x_1), F_2(x_2)).f_1(x_1) \text{,}$$

Accordingly, it is obvious that each terms is able to be decomposed into the proper pair-copula times a conditional general density, through the general formula

$$f(x \mid v) = c_{xv_j|v_{-j}}(F(x \mid x_{v_{-j}}), F(v_j \mid v_{-j})).f(x \mid v_{-j}),$$

for a *d*-dimensional vector v. Here v_j is one arbitrarily chosen component of v and v_{-j} denotes the $v-$vector, excluding the jth component.

Marginal conditional distributions involving pair-copula construction can be calculated as (Joe, 1996)

$$F(x \mid v) = \frac{\partial C_{xv_j|v_{-j}}(F(x \mid v-j), F(v_j \mid v_{-j}))}{\partial F(v_j \mid v_{-j})},$$

where $C_{i,j|k}$ is a bivariate copula distribution function. The function $h(x,v,\theta)$ represent the conditional distribution function where X and V are uniform, i.e. $f(x) = f(v) = 1, F(x) = x$ and $F(v) = v,$ thus,

$$h(x,v,\theta) = F(x \mid v) = \frac{\partial C_{x,v}(x,v,\theta)}{\partial v},$$

To sum up, under suitable regularity conditions, a multivariate density is able to be expressed as a product of pair-copula, acting on several different conditional probability distributions. Needless to say that the construction is iterative in its nature, and that given a specific factorization, there are still many different re-parameterizations (Aas et al., 2009). As explained earlier, this re-parameterizations can be done by graphical regular vine structure.

4.5 Properties: Tail Dependence Properties of Vine Copula

The following dependence properties for vine copula can be found in (Joe, 1996; Joe et al., 2010):

a) Let the edge e be in F_l with $l > 1$ and let the conditioned set $e = \{e_1, e_2\}$. If C_e is more concordant than C'_e, then the margin $F_{e1,e2}$ is more concordance than $F'_{e1,e2}$.

b) If C_e has upper (lower) tail dependence for all $e \in E_1$, and the remaining copula have support on $[0,1]^2$, all bivariate margins of $F_{1...n}(x_1,...,x_n)$ have upper (lower) tail dependence.

c) For parametric vine copula with a parameter θ_e associated with C_e, a wide range of dependence is obtained if each $C_e(.,\theta_e)$ can vary from the bivariate Frèchet lower bound to the Frèchet upper bound. Consider the Kendall tau triple $(\tau_{12}, \tau_{13}, \tau_{23})$ for n = 3. It is shown in (Joe et al., 2010) for a 3-dimensional vine copula that if $C_{23|1}$ is the conditional Frèchet upper (lower) bound copula, and then τ_{23} achieves the maximum (minimum) possible bond, given τ_{12}, τ_{13} .

4.6 Copula Estimation

Estimation of copula vine or pair-copula construction is different from normal multivariate distribution. It is due to the fact that the assumption on the dependence between random variables and conditional random variables are based on graphical vine structure which is separated from the density function. From the theoretical point of view, one has to check the best possible vine structure, but it is impossible in application since the number of vines structure increases rapidly with dimensions. Therefore, to estimate the parameters, we first assume that the vine structure is fixed. We further assume that the conditional copula do not depend on conditioning variables.

With these assumptions, the estimation of copula parameters is achieved using maximum likelihood principle sequentially from the first tree. In this part, we present the maximum likelihood estimation method, developed by Aas et al. (Aas et al., 2009) for C- and D-vine. Although the maximum likelihood procedure can be extended for arbitrary regular vine, the corresponding algorithm is rather vague.

Let $X_i = (X_{i,1},...,X_{i,T}), i = 1,...,n$ denote the ith random variable observe at T time points. Initially, for the sake of simplicity, it is assumed that the T observations of every variable are independent over time. This assumption is not limiting, because when the temporal dependence is present, univariate time-series models can be fitted to the margins and the analysis could henceforward proceed with the residuals (Kurowicka & Joe, 2010). Moreover, since the method focuses on copula estimation, the marginal distributions can be estimated separately using a two-stage procedure or normalized rank of data. Next, we assume that the conditional bivariate copulas are constant over the values of the conditioning variables (Kurowicka & Joe, 2010).

Let $C_{i_1,i_2|m}(u_{i_1},u_{i_2})$ denote the copula with conditioned set $\{i_1,i_2\}$ and conditioning set m. For partial derivatives with respect to u_{i_1} and u_{i_2} we use following notation based on (Kurowicka & Joe, 2010):

$$F_{i_1|i_2:m}(x_{i_1} \mid x_{i_2} : m) = C_{i_1|i_2:m}(u_{i_1}, u_{i_2}) = \frac{\partial C_{i_1 i_2|m}}{\partial u_{i_2}},$$
$$F_{i_2|i_1:m}(x_{i_2} \mid x_{i_1} : m) = C_{i_2|i_1:m}(u_{i_2} \mid u_{i_1}) = \frac{\partial C_{i_1 i_2|m}}{\partial u_{i_1}}.$$

4.6.1 C-vine Model Estimation

Consider a C-vine copula with $n-$nodes and $n(n-1)/2$ pair-copulas are arranged on $(n-1)$ trees according to the C-vine structure. In the first C-vine tree, T_1, the dependence with respect to one particular variable, the first root node is modeled using bivariate copulas for each pair. Conditioned on this variable, pairwise dependencies with respect to a second pairwise variable are modeled the second root node. In general, a root node is chosen in each tree and all pairwise dependencies with respect to this node are modeled, conditioned on all previous root nodes. As mention in section 4.2, C-vine trees have a star structure with density function as in equation 4.3.

The log-likelihood for C-vine is given by

$$\ln f(x) = \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \sum_{t=1}^{T} \ln \left(c_{i,i+j|1:(i-1)} \left(F(x_{i|t} \mid x_{1,t},\dots,x_{i-1,t}) F\left(x_{i|+j,t} \mid x_{1,t},\dots,x_{i-1,t}\right) \mid \theta_{i|,i+j|1:(i-1)} \right) \right), \quad (4.8)$$

The number of parameters depends on the different type of copula used. The log-likelihood function must be maximized numerically all over the parameters. In case if marginal are estimated by maximum likelihood method, they can be added in the logarithmic function in equation (4.8):

$$L = \sum_{t=1}^{T} \ln f(x_{i,t}; \alpha_i).$$

Since the maximum likelihood estimation for C-vine copula is clear, the flowchart is straightforward:

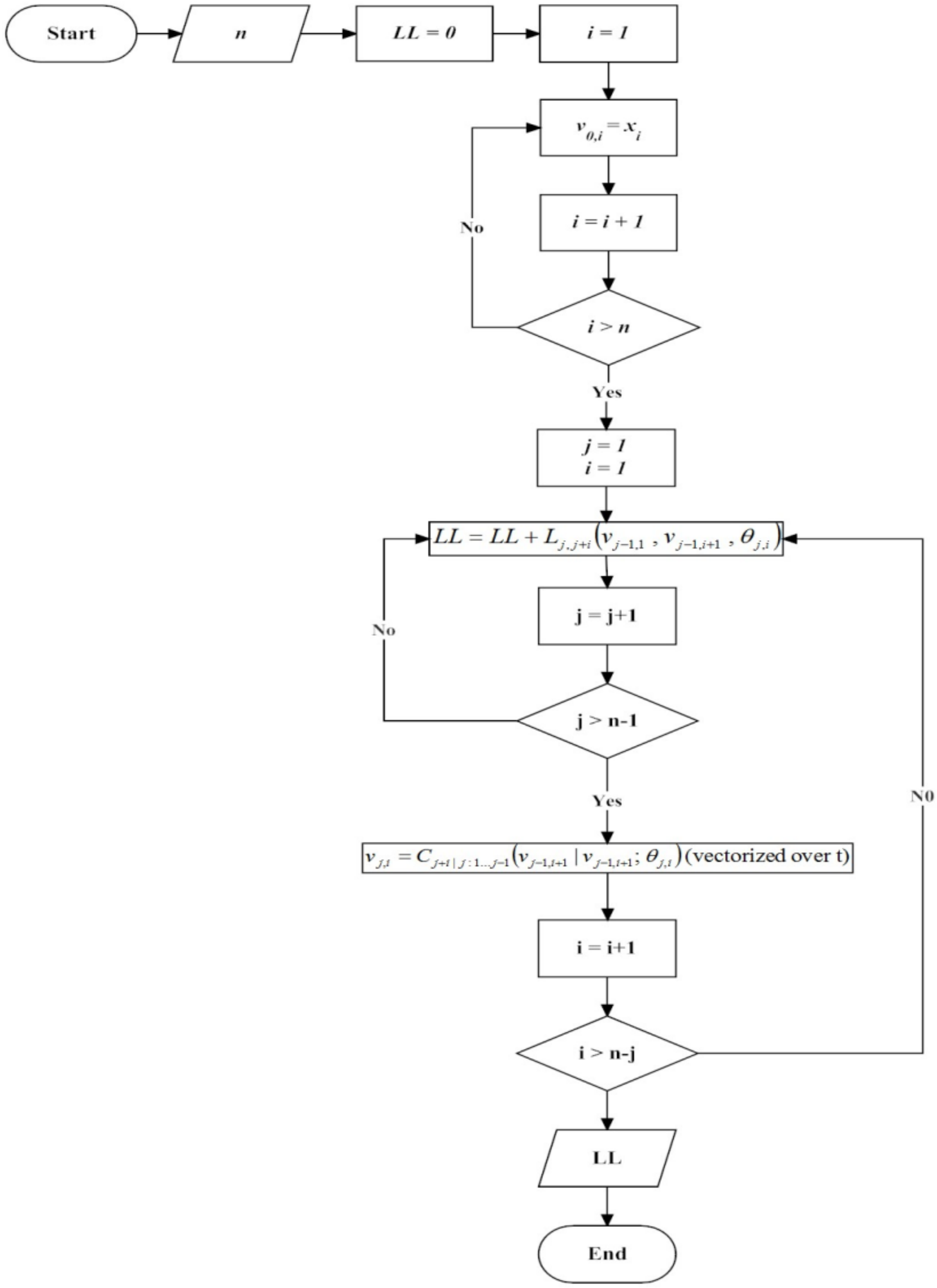


Figure 4.5: Flowchart C-vine copula inference

where $L_{i,j+i}(y,v,\theta)=\sum_{t=1}^{T}\log\{c_{i,i+j|1:(i-1)}(y_t,v_t,\theta)\}$.

Starting point for numerical maximization can be determined by

1) Estimate all parameters of copula in first tree from the original data.
2) Simulate the data in second tree using conditional distribution function from the first tree.
3) Estimate the copula parameter in second tree by data from (b).
4) Iterate until convergence.

Since the data set in each step is bivariate the computation is easy to perform.

4.6.2 D-vine Inference

The pattern is exactly the same as inference method in C-vine. The log-likelihood function is defined as logarithm of density function in equation 4.4. The maximum likelihood applies sequential from the first tree by considering the bivariate copula and D-vine structure. Interested reader refer to (Aas et al., 2009; Kurowicka & Joe, 2010; Nikoloulopoulos, Joe, & Li, 2012).

4.7 Model Inference

As mentioned in section 4.6, maximum likelihood estimation is based on the assumption of fixed vine copula, but full inference for pair-copula decomposition should in principle take into account (a) the selection of a regular vine, (b) the choice of (conditional) copula types, and (c) the estimation of the copula parameters. In this section, the problem of selecting a regular vine is considered.

Two approaches have been suggested for choosing the best regular vine; the first one is a sequential estimation and the second one is based on mutual information. The detail of sequential estimation will be reviewed in this section.

The first step in sequential estimation method relies on choosing either the C- or D-vine copula. When there is a canonical variable which the other variables depend on it, C-vine may be more appropriate than D-vine; otherwise use D-vine.

In model inference the parametric shape of every pair copula needs to be specified. It is obvious that the multivariate distribution is valid if the parametric copula best fit the data. It is possible to choose on predefining class of copula, but a more accurate strategy is to choose a copula for each pair of observation separately. To implement this method, we can apply the following algorithm (Kurowicka & Joe, 2010):

1) Determine the type of copulas to employ in first tree, T_1 , by plotting the original data, and checking for tail dependence or asymmetries (these are the patterns that make the multivariate normal copula inadequate).
2) Estimate the parameters of the selected copula using the original data.
3) Transform observations as required for the second tree, T_2, using the copula parameters from T_1 and the conditional functions in Section 3.7.
4) Identify the type of copula to use in T_2 (in the same way as in T_1).
5) Continue the steps until T_{n-1}.

Therefore, each copula selection depends on the selected copula in previous level. This selection does not guarantee a global optimal fit. By having the appropriate parametric shapes for each copula we can estimate the parameters according to maximum likelihood method, as in equation (4.8).

4.8 Archimedean C- and D-Vine Copula

In this section, we implement some examples of a three-dimensional copulas according to the C- and D-vine structure based on Gumbel and Clayton copula. Two important trigonometric copula, Csc and Cot copulas have been chosen to be a building block for C- and D- vine copula.

Since the role of different copula does not affect the structures, we construct the C- and D-vine distribution function in general, then specific copula density is calculated. According C-vine structure, we have following C-vine distribution function:

$$\begin{aligned} f(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3) & \\ & c_{12}(F(x_1), F(x_2)) c_{13}(F(x_1), F(x_3)) \\ & c_{23|1}(F(x_1 \mid x_2), F(x_1 \mid x_3)). \end{aligned}$$

Likewise, the density function with D-vine structure is given as

$$\begin{aligned} f(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3) & \\ & c_{12}(F(x_1), F(x_2)) c_{23}(F(x_2), F(x_3)) \\ & c_{13|2}(F(x_1 \mid x_2), F(x_3 \mid x_2)). \end{aligned}$$

4.8.1 Clayton Copula

To illustrate the construction of vine copula, we choose Clayton copula with the following distribution function: $C_{12}(u_1, u_2) = \left(u_1^{-\theta_{12}} + u_2^{-\theta_{12}}\right)^{-\frac{1}{\theta_{12}}}$.

The Clayton density is then:

$$
\begin{aligned}
h_{12}(u_1,u_2) &= F_{1|2}(u_1 \mid u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1,u_2) \\
&= \frac{\partial}{\partial u_2}\left(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1\right)^{-\frac{1}{\theta_{12}}} \\
&= u_2^{-\theta_{12}-1}\left[\left(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1\right)^{-1-\frac{1}{\theta_{12}}}\right].
\end{aligned}
$$

4.8.2 Gumble Copula

The Gumble copula distribution function is:

$$
C_{12}(u_1,u_2) = \exp\left(-\left[(-\ln u_1)^{\theta_{12}} + (-\ln u_2)^{\theta_{12}}\right]^{\frac{1}{\theta_{12}}}\right).
$$

For this copula we have

$$
\begin{aligned}
h_{12}(u_1,u_2) &= F_{1|2}(u_1 \mid u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1,u_2) \\
&= \frac{\partial}{\partial u_2} \exp\left(-\left[(-\ln u_1)^{\theta_{12}} + (-\ln u_2)^{\theta_{12}}\right]^{\frac{1}{\theta_{12}}}\right) \\
&= C_{12}(u_1,u_2)\frac{1}{u_2}(-\ln u_2)^{\theta_{12}-1}\left[(-\ln u_1)^{\theta_{12}} + (-\ln u_2)^{\theta_{12}}\right]^{\frac{1}{\theta_{12}}-1}.
\end{aligned}
$$

4.8.3 Cot Copula

The Cot copula distribution function is:

$$
C_{12}(u_1,u_2) = \frac{2}{\pi} arc\cot(\cot^{\theta_{12}}(\frac{\pi}{2}u_1) + \cot^{\theta_{12}}(\frac{\pi}{2}u_2))^{\frac{1}{\theta_{12}}}
$$

The Cot copula density is then

$$
\begin{aligned}
h_{12}(u_1,u_2) &= F_{1|2}(u_1 \mid u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1,u_2) \\
&= \frac{\partial}{\partial u_2}\left(\frac{2}{\pi} arc\cot\left(\cot^{\theta_{12}}\left(\frac{\pi}{2}u_1\right)+\cot^{\theta_{12}}\left(\frac{\pi}{2}u_2\right)\right)^{\frac{1}{\theta_{12}}}\right) \\
&= \frac{\csc^2\left(\frac{\pi}{2}u_2\right)\cot^{\theta_{12}-1}\left(\frac{\pi}{2}u_2\right)\left(\cot^{\theta_{12}}\left(\frac{\pi}{2}u_1\right)+\cot^{\theta_{12}}\left(\frac{\pi}{2}u_2\right)\right)^{\frac{1}{\theta_{12}}-1}}{\left(\cot^{\theta_{12}}\left(\frac{\pi}{2}u_1\right)+\cot^{\theta_{12}}\left(\frac{\pi}{2}u_2\right)\right)^{\frac{2}{\theta_{12}}}+1}.
\end{aligned}
$$

4.8.4 Csc Copula

The Csc copula distribution function is

$$
C_{12}(u_1,u_2) = \frac{2}{\pi} a\csc\left(\left(\csc\left(\frac{\pi}{2}u_1\right)-1\right)^{\theta_{12}}+\left(\csc\left(\frac{\pi}{2}u_2\right)-1\right)^{\theta_{12}}\right)^{\frac{1}{\theta_{12}}}+1
$$

The Csc density is define as

$$
\begin{aligned}
h_{12}(u_1,u_2) &= F_{1|2}(u_1 \mid u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1,u_2) \\
&= \frac{\partial}{\partial u_2}\left(\frac{2}{\pi} a\csc(A)^{\frac{1}{\theta_{12}}}+1\right) \\
&= -\frac{\csc\left(\frac{\pi}{2}u_2\right)\cot\left(\frac{\pi}{2}u_2\right)\left(\csc\left(\frac{\pi}{2}u_2\right)-1\right)^{\theta_{12}-1}}{A\sqrt{(A)^{\frac{2}{\theta_{12}}}-1}}.
\end{aligned}
$$

where $A = \left(\csc\left(\frac{\pi}{2}u_1\right)-1\right)^{\theta_{12}}+\left(\csc\left(\frac{\pi}{2}u_2\right)-1\right)^{\theta_{12}}$.

4.9 Application: US-Asia Index

The objective of this section is to model a multivariate distribution function with a copula-EGARCH model. We demonstrate the multivariate distribution function with C- and D-vine structure.

We examine daily data of three stock index returns: the Strait Times Index (STI) of Singapore, the Kuala Lumpur Composite Index (KLCI) of Malaysia and Standard and Poor index (SP500) of USA for the period January 01st, 1998 through December 31st, 2008. The data sets collected from DataStream consist of daily closing price with a total of n = 2780 observations. In the database, the daily return $R_{i,t}, i = 1,...,9$ consisted of daily closing price $P_{i,t}$, which is measured in local currency and computed as $R_{i,t} = \ln(P_{i,t} / P_{i,t-1})$.

Before proceeding to the estimation of the marginal and copula models, it is useful to assess their descriptive statistical properties. Table 4.1 reports the descriptive statistics of the daily financial market returns for the time series under consideration. Notably, in terms of daily returns, SP500 has the lowest mean returns with negative sign (-0.002%). The mean returns of KLCI and STI financial markets are positive with 0.014% and 0.004% respectively. It is clear that Malaysian financial market offer higher average returns than the most advanced financial markets, that is, US and Singapore financial markets but these high returns are also characterized by larger volatility, which is common for emerging financial markets and is consistent with previous studies (Abu et al. 2009; Miyakoshi T. 2003).

Table 4.1 displays the skewness, kurtosis, and related tests of the data collected. The Ljung-Box Q-statistics $Q(10)$ and $Q^2(10)$ which test for serial correlation in daily and squared returns, respectively, rejects the null hypotheses of non-serial correlation. These time series display typical features of stock returns such as fat tail, spiked peak, and persistence in variance. With evidence of ARCH effects as indicated by LM test, it is possible to proceed to the next step of the analysis which focuses on the bivariate EGARCH(1, 1) modeling of the dynamics of market's volatility in estimating the marginal distributions.

Table 4.2 presents the estimation results for the parameter and the use of asymmetric EGARCH model seems to be justified with all asymmetric coefficients significant at standard levels. The EGARCH model seems are reasonably good at describing the dynamics of the first two moments of the series as shown by the Ljung-Box statistics for the squared standardized residuals. LM test for presence of ARCH effects at lag 10, indicate that the conditional hetroskedasity that existed when the test was performed on the pure return series (see Table 4.1) are removed. The leverage effect term γ_i in the marginal EGARCH models are statistically significant, furthermore, with γ_i negative sign, as expected that negative shocks imply a higher next period conditional variance than positive shocks, indicating that the existence of leverage effect is observed in returns of the financial market series. Briefly, looking at the overall results, we can argue that EGARCH model adequately explains the data set under investigation. The marginal models seem to be able to capture the dynamics of the first and second moments of the returns of the financial time series. The time series plots of the returns are given in Figure 4.6.

The three return series behave similarly over time, exhibit periods of high and low volatility, and sometimes take on extremely large and small values, particularly STI and SP500 series for the more recent period.

Table 4.1: Summary statistics for daily equity market returns

	Mean	Std.	Skewness	Rob.Sk	Kurtosis	Rob.Kr	Q(10)	Q(10)	ARCH(5)
STI	0.0036	1.311	-0.1214	-0.021	8.5771	0.2701	17.895	912.2**	344.8**
KLCI	0.0135	1.506	0.5695	-0.004	60.2395	0.3758	85.29**	1365.9**	787.6**
SP500	-0.002	1.336	-0.1187	0.013	10.5733	0.3204	61.95**	2096.7**	572.1**

Rob.Sk and Rob.Kr are outlier-robust versions of skweness and kurtosis described as Sk2 and Kr2 in Kim and White (2004). **,* Significant at 1% and 5% respectively.

Table 4.2: Parameter estimates of marginal models.

	ψ_i	α_i	β_i	γ_i	Q(10)	$Q^2(10)$	LM(10)
STI	-0.166** [0.015]	0.237** [0.023]	0.981** [0.004]	-0.229** [0.053]	27.88**	10.788	10.035
KLCI	-0.141** [0.015]	0.192** [0.021]	0.981** [0.004]	-0.358** [0.067]	100.58**	2.786	2.709
SP500	-0.082** [0.011]	0.120** [0.016]	0.991** [0.002]	-0.508** [0.099]	11.74	12.781	12.777

Note: **,* Significant at 1% and 5% respectively. Standard errors are given in square brackets.

(a)

(b)

(c)

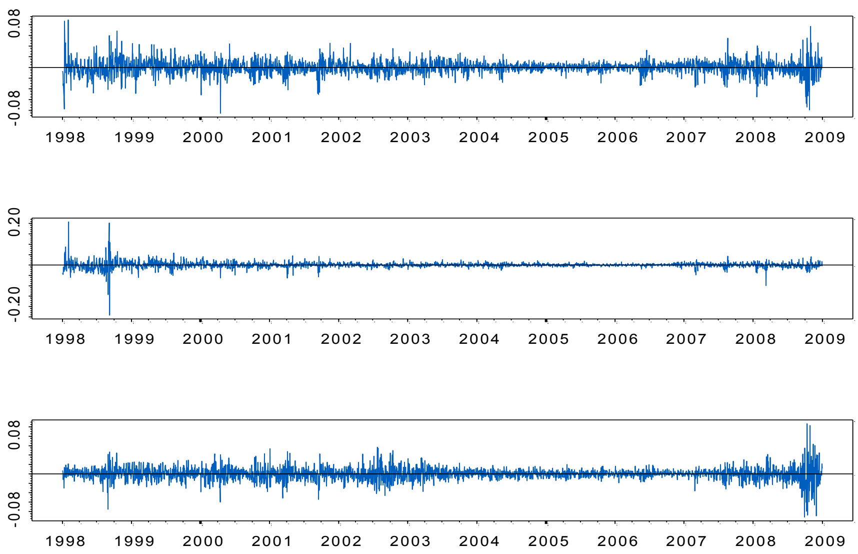


Figure 4.6: Daily log returns on (a) STI, (b) KLCI and (c) SP500 indices.

As mentioned earlier, the main aim of this section is to model a multivariate distribution function between international markets. In this framework, we have used the Inference for the Margins (IFM) method, to estimate the margins and the copula parameters. Firstly, the marginal distributions of each stock index are independently estimated via maximum likelihood through an EGARCH model. Then the standardized residuals are transformed into uniform margins with empirical probability integral transformation to copula data on (0, 1).

To investigate the best multivariate distribution function, we start with analyzing the residuals. Figure 4.7 shows the scatter plot of data. The scatter plot suggests week dependence among almost all indices. Therefore, the strongest positive dependence is between STI-SP500 which is also supported by evidence of chi- and k-plot of data in Figure 4.8. This result is in accordance with the result of Kendall's τ in Table 4.3. Interestingly, the empirical upper and lower tail dependence between the random variable given in Table 4.3 is not zero. By considering all dependence facts between random variables, Normal, Frank and t- copula has been selected for negative week dependence between KLCI-STI, KLCI –SP while a wider family of copula is selected for SP-STI. The estimation results of bivariate analysis are summarized in Table 4.4, Table 4.5, Table 4.6.

According to Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS)-goodness of fit method, the t-copula is the best to capture dependence properties between the pair of random variables.

Table 4.3: Kendall's τ calculation

	KLCI	SP500	STI
KLCI	1.0000	-0.0064	-0.0084
SP500	-0.0064	1.0000	0.0499
STI	-0.0084	0.0499	1.0000
AbsSum:	1.0148	1.0562	1.0583

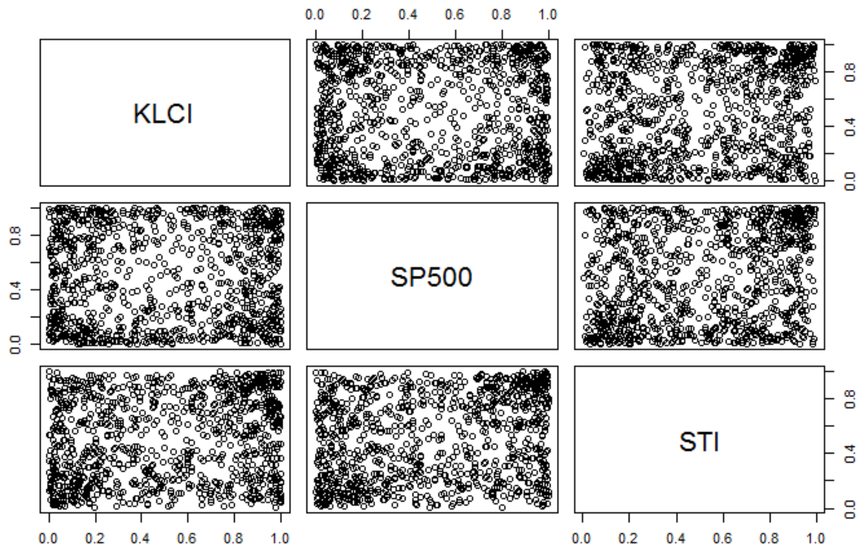


Figure 4.7: Scatter Plot of (KLCI-STI-SP).

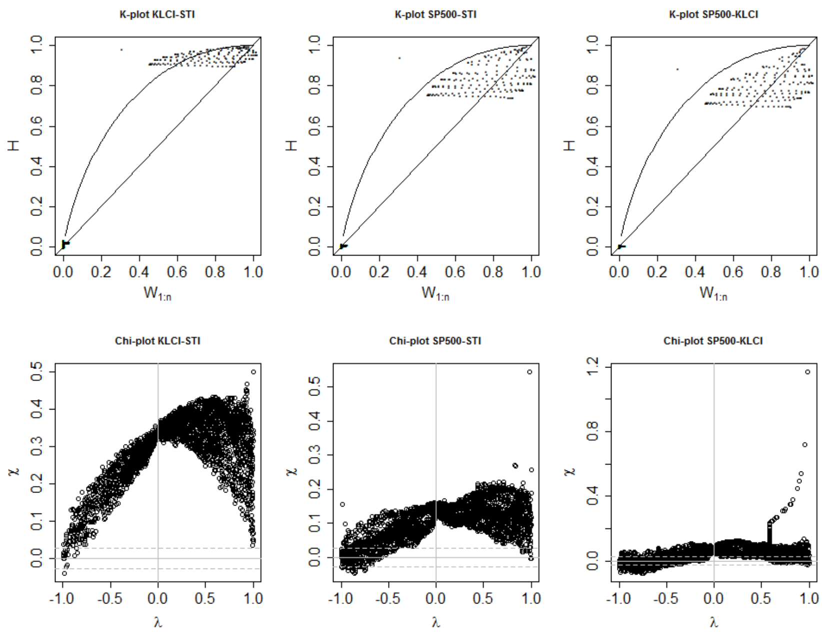


Figure 4.8: K-plot and Chi plot of stocks' pairs

Table 4.4: KLCI- SP500

	Param 1	Param 2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS
Normal	-0.0114 [0.0192]	0 [0]	0	0	0.5392	0.02	1.4448
t-Student	-0.0093 [0.0210]	10.6125 [0]	0.0051	0.0051	0.2065	0.15	1.0143
Frank	-0.0578 [0.1198]	0 [0]	0	0	0.3581	0	1.1920

Note: Significant at 5%

Table 4.5: KLCI- STI

	Param1	Param2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS
Normal	-0.0114 [0.0192]	0 [0]	0	0	0.5392	0.02	1.4448
t-Student	-0.0093 [0.0210]	10.6124 [2.6109]	0.0051	0.0051	0.2065	0.15	1.0143
Frank	-0.0578 [0.1198]	0 [0]	0	0	0.3581	0	1.1920

Note: Significant at 5%

Table 4.6: STI-SP500

	Param1	Param2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS	P-value KS
Normal	0.0746 [0.0191]	0.0000 [0.0000]	0.0000	0.0000	0.2685	0.0900	1.0183	0.2000
t-Student	0.0791 [0.0227]	3.4896 [0.3057]	0.1142	0.1142	0.1205	0.6600	0.8343	0.6400
Clayton	0.1155 [0.0238]	0.0000 [0.0000]	0.0025	0.0000	0.3332	0.0100	1.1957	0.0100
Gumbel	1.0657 [0.0127]	0.0000 [0.0000]	0.0000	0.0837	0.4212	0.0000	1.2029	0.0100
Frank	0.4718 [0.1228]	0.0000 [0.0000]	0.0000	0.0000	0.4526	0.0000	1.1981	0.0100
BB1	0.0659 [0.0249]	1.0459 [0.0132]	0.0000	0.0600	0.4976	0.0000	1.2996	0.0000
BB6	1.0010 [0.0561]	1.0649 [0.0442]	0.0000	0.0840	0.4210	0.0000	1.2061	0.0200
BB7	1.0621 [0.0171]	0.0865 [0.0237]	0.0003	0.0795	0.4394	0.0000	1.2337	0.0000
BB8	1.0956 [0.0191]	0.9988 [0.0004]	0.0000	0.0000	0.4822	0.0000	1.4901	0.0000

Note: Significant at 5%.
Trigonometric copula is not applicable in this simulation due to the weak dependence among the three stock indices

There are many alternative ways to build a multivariate distribution functions based on C- or D-vine structure by considering different order of variables. Here, we consider five different models; three based on C- and two based on D-vine model. The final multivariate distribution function will be chosen from those models.

The C-vine copulas are useful when one expects a variable to dominate the dependence with all other variables. This specific variable represents the canonical node in structure of multivariate copula. Considering the dependence properties of our data, STI index has dependence with SP500 from one side and the KLCI from the other side while the dependence between SP500 and KLCI is weak. Using this fact, the two models based on C-vine structure, MC1 and MC2, the STI would be the canonical node. Different ordering in second and third nodes between SP500 and KLCI will result in different models, MC1 and MC2. To compare the result, we also consider MC3 where KLCI with weak dependence, is canonical node in the C-vine structure.

We also implement the D-vine copula, since in our data set the difference in absolute sum of the empirical Kendall's τ dependence is not significant. Among several ordering of variables in D-vine structure, we consider two models MD1 and MD2.

The result of estimation for five methods, MC1, MC2, MC3, MD1 and MD2, based on sequential and MLE are listed in Table 4.7. The results indicate that the differences between the estimated parameters using both methods, sequential and MLE are insignificant. This property highlights the goodness of sequential estimation in application. Note that for each pair, t-copula has been selected as optimal bivariate joint distribution. This aligns with bivariate analysis seen above which suggests t-copula for every pair of random variables.

Table 4.7: Estimation of Parameters C- and D- vine models

Method	NO	Tree	Copula	Sequential		MLE	
				Para 1	Para 2	Para 1	Para 2
MC 1	1	STI-SP500	t	0.0275	2.8662	0.0239	2.8675
		STI-KLCI	t	0.0035	3.0844	0.0084	3.1378
	2	SP500-KLCI	t	0.0258	4.2294	0.0259	4.2121
MC 2	1	STI-KLCI	t	0.0275	2.8662	0.0239	2.8675
		STI-SP500	t	0.0035	3.0844	0.0084	3.1378
	2	KLCI- SP500	t	0.0258	4.2294	0.0259	4.2121
MC 3	1	KLCI- STI	t	0.0275	2.8662	0.0239	2.8675
		KLCI-SP500	t	0.0035	3.0844	0.0084	3.1378
	2	STI-SP500	t	0.0258	4.2294	0.0259	4.2121
MD 1	1	STI- SP500	t	0.0275	2.8662	0.0245	2.8621
		SP500- KLCI	t	0.0328	2.9459	0.0267	2.9506
	2	STI-KLCI	t	0.0087	4.6053	0.0087	4.6043
MD 2	1	SP-STI	t	0.0275	2.8662	0.0245	2.8621
		STI-KLCI	t	0.0328	2.9459	0.0267	2.9506
	2	SP-KLCI	t	0.0087	4.6053	0.0087	4.6043

To find the best fitting multivariate distribution function for our data set, we compare the goodness of fit for all models based on AIC and BIC criteria in Table 4.8.

The D-vine structure shows a better result compared to C-vine, which is attributed to weak dependence between the three stocks indices. However, the importance of canonical nodes in C-vine structure is obvious by comparing the result of MC1 and MC2 with MC3 where the KLCI is considered as canonical nodes instead of STI. Finally, the differences in MD1 and MD2 emphasize on the importance of ordering in D-vine structure. To conclude, we confirm the importance of both selection of vine structure and ordering of variables in construction a multivariate copula. The MD2 method is reported as the best among all models which suggests that there is no preference surmount variable due to weak dependence. In addition, it also suggests the importance of ordering in structure where STI depends on both KLCI and SP500.

Table 4.8: Goodness of Fit Test based on AIC and BIC criteria

	MC1	**MC2**	**MC3**	**MD1**	**MD2**
AIC MLE	-564.724	-564.7243	-563.2700	-568.1453	-569.5996
AIC seq.	-564.608	-564.6079	-564.6079	-568.0563	-569.5106
BIC MLE	-529.467	-529.4667	-529.4667	-532.8877	-533.2077
BIC seq.	-529.35	-529.3504	-529.3504	-532.7988	-533.6748
Log-likelihood MLE	288.3621	288.3621	286.9078	290.0726	291.4184
Log-likelihood Seq.	288.304	288.3040	286.6497	290.0282	291.0405
Parameters	6	6	6	6	6

Note: Significant at 5%

Chapter 5: Copula and Its Application

This chapter demonstrates the application of copula in modelling the finance data. We begin by analysing bivariate pairs of indices from three different continents, America, Europe and Asia, and examine the ability of trigonometric copula in capturing the dependence properties of indices. The second part of the finance application describes the design and characterization of multivariate joint distribution function which is built according to vine structure.

5.1 Bivariate Analysis of Index

This section critically examines the ability of bivariate trigonometric copula in seven indices from three continents: America, Europe and Asia.

Daily data of seven stock indices return were recruited for this study. Three index returns: German DAX, The French CAC40 and British FTSE 100 index from Europe, together with the Strait Times Index (STI) of Singapore, the Kuala Lumpur Composite Index (KLCI) of Malaysia and the Chinese composite index (SSE) from Asia and the Standard and Poor index (SP500) of USA for the period January 1st, 2000 to July 15th, 2012 were downloaded from Yahoo Finance. In the database, the daily return $R_{i,t}, i = i = 1,...,7$ consisted of daily closing price $P_{i,t}$ which is measured in local currency and computed as $R_{i,t} = \ln(P_{i,t} / P_{i,t-1})$.

As the first step, it is useful to describe the statistical properties of returns. The results obtained from the preliminary analysis of returns are presented in Table 5.1. It is apparent from this table that Asian indices have the highest mean returns with China index, SSE, as the biggest return value. It is followed by the European index with only

Germany DAX index with positive return. The minimum value of returns during this period belongs to America S&P500 with -0.00405. High returns in Germany and China is characterized by larger volatility. The Minimum volatility belongs to the Malaysia market, KLCI, which also can be observed in Figure 5.1. The skewness, kurtosis, and related tests of the data collected in Table 5.1. The Ljung-Box Q-statistics $Q^2(10)$ which test the serial correlation in squared returns rejects the null hypotheses of non-serial correlation.

Table 5.1: Summary statistics of Indices

	Mean	**Std.**	**Skewness**	**Kurtosis**	**Q2 (10)**
CAC40	-0.0192	1.5779	0.0700	4.4664	2.2e-16
DAX	0.0000	1.6437	0.0335	3.8323	2.2e-16
FTSE100	-0.0063	1.3120	-0.1128	5.5717	2.2e-16
SP500	-0.0041	1.3704	-0.1529	7.1188	2.2e-16
KLCI	0.0620	1.1611	-0.3203	84.3633	2.2e-16
STI	0.0227	1.2778	-0.4792	5.5651	2.2e-16
SSE	0.0396	1.6014	-0.1122	4.4683	2.2e-16

Note: Significat at 5%

With evidence of ARCH effects as indicated by Ljung-Box Q-statistics $Q^2(10)$ test of squared returns, it is possible to proceed to the next step of the analysis which focuses on the bivariate GARCH(1, 1) modelling of the dynamics of market's volatility in estimating the marginal distributions. Table 5.2 presents the estimated results for the parameter and the use of GARCH model seems to be justified with all coefficients significant at the standard levels. The GARCH model seems reasonably good at describing the dynamics of the first two moments of the series as shown by the Ljung-Box statistics for the standardized residuals with lag 10. LM test for presence of ARCH effects at lag 10, indicates that the conditional hetroskedasity that existed in the pure return series (Table 1) is removed. Briefly, looking at the overall results, it can be argued that a GARCH(1,1) model adequately explains the data set under investigation. The marginal models seem to be able to capture the dynamics of the first and second moments of the returns of the financial time series. Figure 5.1 shows the time series plot of indices.

Table 5.2: Time series GARCH model

	Mu	**Omega**	**Alpha1**	**Beta1**	**Shape**	**Q(10)**	**LM Arch**
CAC40	0.0462*	0.0167**	0.0889***	0.9075***	10.00***	16.0740 [0.0975]	23.2844 [0.0254]
DAX	0.0002**	0.0176***	0.0887***	0.9078***	10.00***	7.6124 [0.6666]	19.8766 [0.0695]
FTSE100	0.0441**	0.0127***	0.1033***	0.8935***	10.00***	11.7567 [0.3017]	26.9114 [0.08]
SP500	0.0404***	0.0102***	0.0868***	0.9099***	8.3562***	13.8987 [0.1777]	18.0392 [0.1145]
KLCI	0.0789***	0.0672***	0.2482***	0.7014***	5.0690***	17.8975 [0.0567]	0.1476 [1.0000]
STI	0.0706***	0.0145***	0.0819***	0.9104***	7.6272***	17.1819 [0.0704]	8.1499 [0.7733]
SSE	0.0667**	0.0296**	0.0718***	0.9227***	4.1883***	24.8060 [0.0518]	27735 [0.9969]

***,**,* Significant at 0.0 , 0.001 and 0.01 respectively. P-values are given in square brackets.

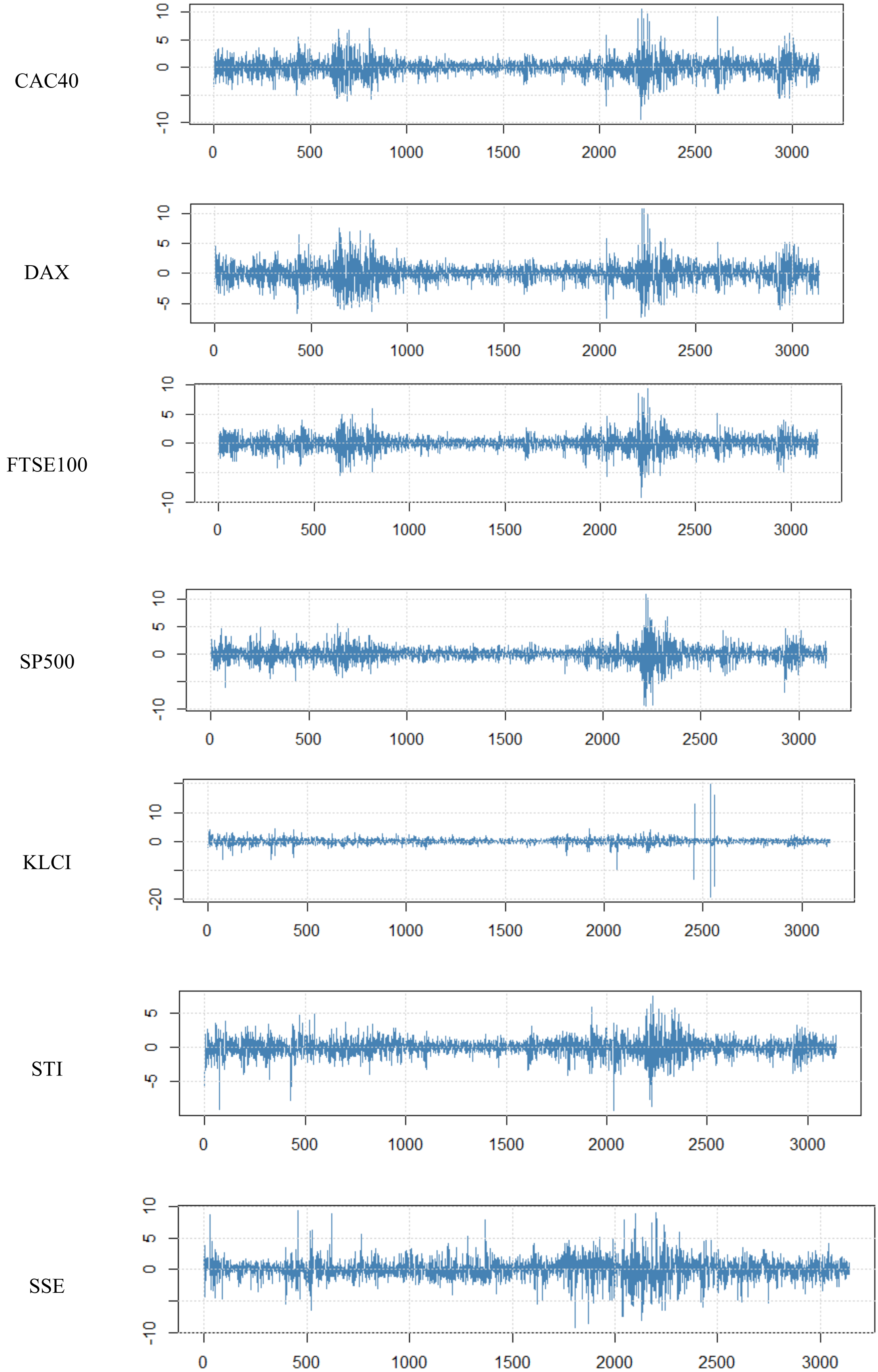


Figure 5.1: Time series of indices.

After removing the seasonality and trend by GARCH model from the data, the resulting standardized residuals of these models are transformed using the empirical probability integral transformation to copula data in [0,1]. To increase the reliability of analysis on data, we plot some graphical test for the purpose of presentation of the dependence among returns.

Figure 5.2 shows the scatter plot among seven returns. The figure is quite revealing in several ways. First, it highlights the strong dependence among European returns. Moreover, it shows relatively strong dependence between America and Europe markets while the dependence between America and Asia is weak. It is apparent from the figure that Asian data resulted in the lowest dependence compared with other markets.

By considering the K- and Chi- plot, the structure of dependence can be studied in more details. For convenience, the data can be classified into three main subgroups. Europe and American index was chosen as the first groups because of strong dependence between indices. Figure 5.3 compares the result of K- and Chi plot for this group. From this data set, a strong positive dependence among all pairs can be observed. This result is also confirmed, according to the empirical Kendall's τ and Spearman's ρ dependence measure.

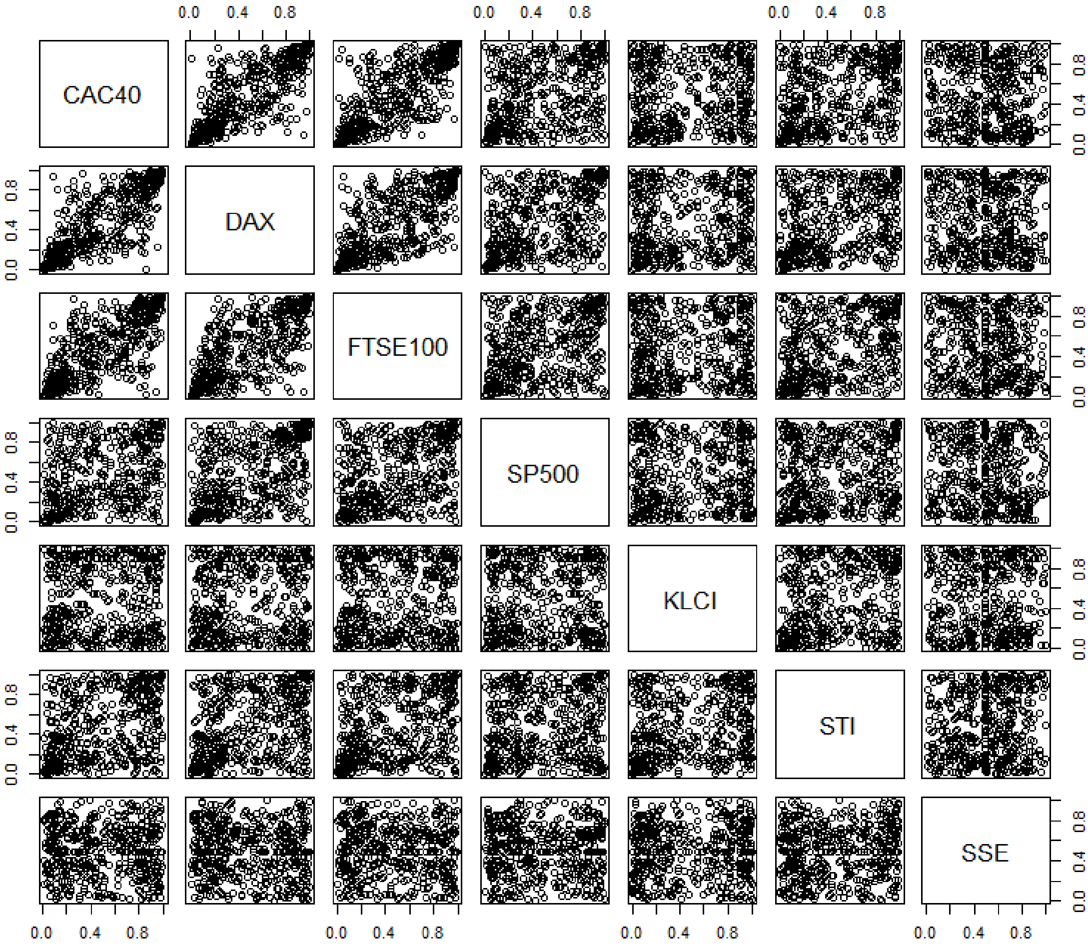


Figure 5.2: Scatter plot of returns.

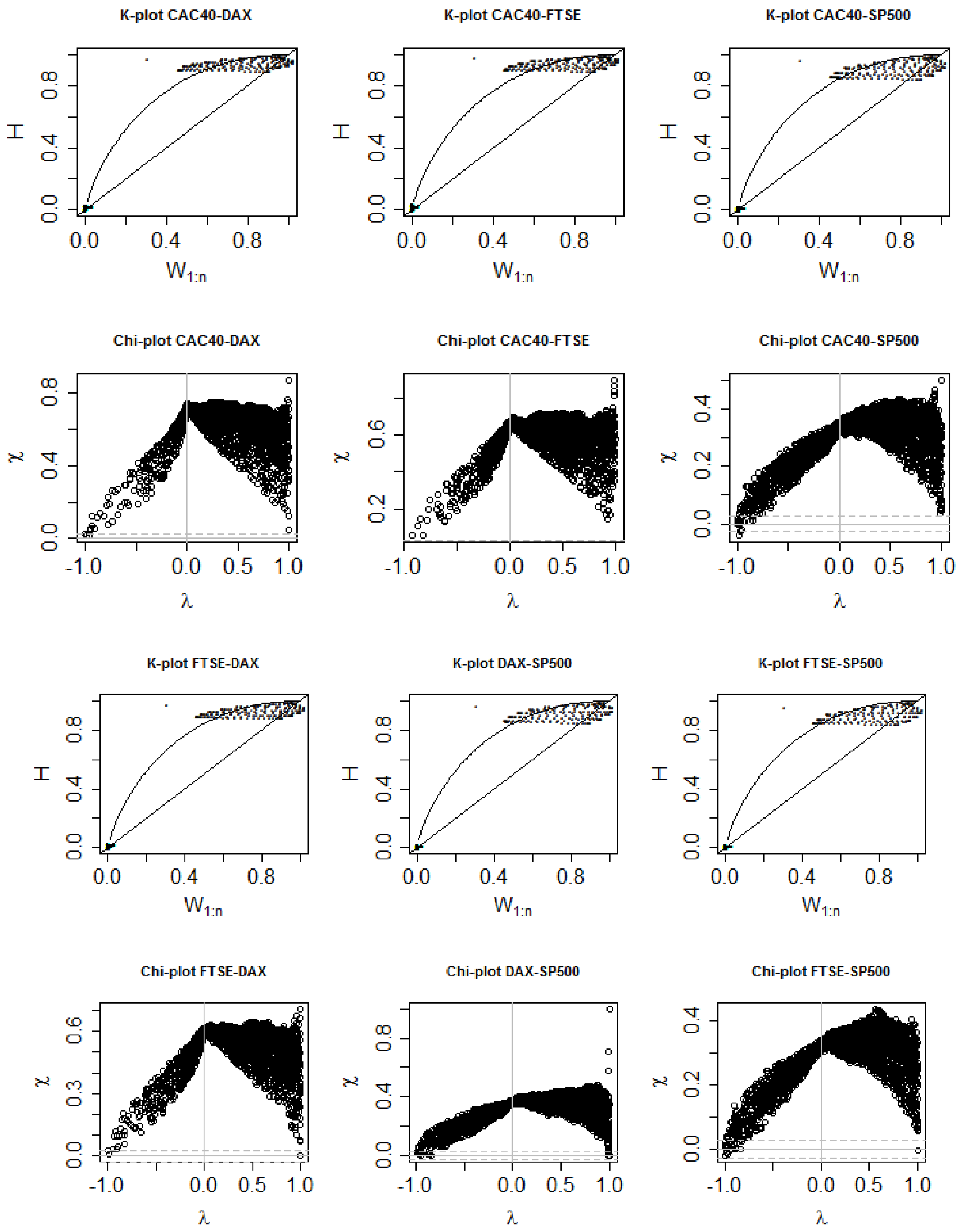


Figure 5.3: Chi- Plot and K-plot of European and America index

Table 5.3 illustrates some of the main characteristics of the dependence. The most striking result to emerge from the data is almost the same strong positive range of dependence between European indices by itself. This also appears among America and

European indices with weaker dependence. Interestingly, the trend is the same for tail dependence as well.

Table 5.3: P-value of independence test, empirical Kendall's τ , Spearman's ρ and both upper and lower tail dependence

INDEX	Independence	Empirical Kendall's τ	Empirical Spearman's ρ	Lower Tail	Upper Tail
CAC40-DAX	0.0000	0.7188	0.8779	0.6540	0.7729
CAC40-FTSE	0.0000	0.6800	0.8517	0.7241	0.7437
CAC40-SP500	0.0000	0.3695	0.5095	0.4123	0.4597
FTSE-DAX	0.0000	0.6142	0.7915	0.6576	0.6774
DAX-SP500	0.0000	0.3943	0.5418	0.4512	0.4858
FTSE-SP500	0.0000	0.3521	0.4873	0.4215	0.4413
SP500-KLCI	0.0000	0.0542	0.0793	0.0432	0.0749
SP500-STI	0.0000	0.1314	0.1893	0.0987	0.1878
SP500-SSE	0.0449	0.0239	0.0352	0.0124	0.0355
KLCI-STI	0.0000	0.2953	0.4182	0.2381	0.3388
KLCI-SSE	0.0000	0.1127	0.1661	0.05321	0.1201
STI –SSE	0.0000	0.1332	0.1955	0.1176	0.1529
DAX-KLCI	0.0000	0.1213	0.1773	0.1023	0.1409
DAX-STI	0.0000	0.2388	0.3407	0.2750	0.2921
DAX-SSE	0.0001	0.0468	0.0699	0.0283	0.0523

The evidence from the dependence results suggests using some strong positive copula which can capture both tail dependences. Considering this fact, trigonometric copulas are compared with other one-parameter Archimedean copula at the first stage in order to investigate the best one-parameter copula to fit on data.

Two pairs CAC40-DAX and DAX-SP500 were chosen. First pair can be a representative of the dependence structure between European indices, while the second one is an indicator of the dependence between America and Europe index. Therefore by choosing these two pairs, we can represent the dependence structure among European market which is slightly different with the dependence pattern among Europe and American indices. In order to assess the best fit among one-parameter copula family, several copula families are chosen. The result of parameter estimation and goodness of fit with tail dependence are listed in Table 5.4 and Table 5.5 CAC40-DAX and DAX-SP500 respectively.

Table 5.4: One – parameter estimation of Archimedean copula CAC40-DAX

	Lower tail	Upper tail	Goodness of fit	parameter
Empirical tail	0.6540	0.7729		
Gumbel	0	0.7703	0.0002	3.3529
Clayton	0.8203	0	0.0012	3.4993
12th family	0.8029	0.7545	0.0004	3.1573
Cot	0.7766	0.7123	0.0001	2.7413
CSC	0.6603	0.7694	0.0001	1.6702
Csc2	0.7939	0.5858	0.0006	3.0038
Csc3	0.7904	0.5858	0.0006	2.9474

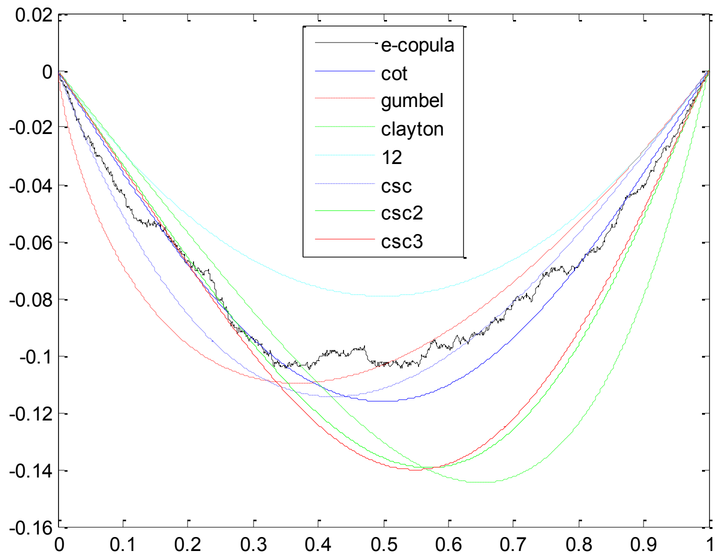


Figure 5.4: Goodness of fit representation of Table 5.4

Table 5.5: One – parameter estimation of Archimedean copula DAX-SP500

	Lower tail	Upper tail	Goodness of fit	parameter
Empirical tail	0.4512	0.4858		
Gumbel	0	0.482	0.0003	1.6606
Clayton	0.4945	0	0.0008	0.9844
12th family	0.7051	0.5817	0.0059	1.9837
Cot	0.6009	0.3358	0.0005	1.3609
CSC	0.4306	0.4761	0	0.8226
Csc2	0.3664	0.5858	0.0006	0.6904
Csc3	0.3272	0.5858	0.0004	0.6204

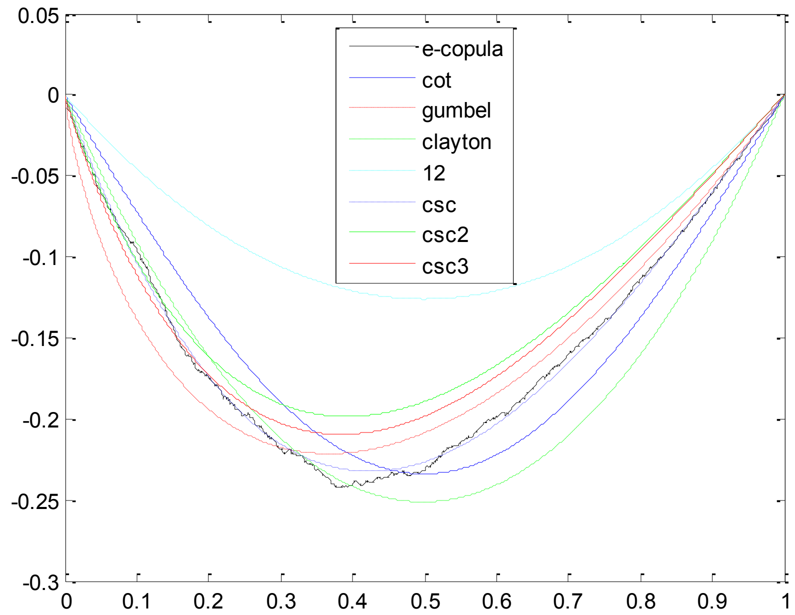


Figure 5.5: Goodness of fit representation of Table 5.5

It is apparent from Table 5.4 that Csc-copula and Cot-copula represent the best fit copula in terms of the goodness of fit measure as well as the closeness to the empirical tail measures. This result is confirmed by graphical representation of the goodness of fit plot. Further analysis on upper and lower tail dependence against the empirical measures also confirm this result. The Gumbel, Clayton, Csc2 and Csc3 copulas cannot represent both upper and lower tail dependence simultaneously; while 12^{th} family overestimate the value of tails. Between Cot and Csc, Csc copula yields similar upper tail dependence as the empirical tail dependence, this making Csc a superior choice. The result of Table 5.5 is similar to those in Table 5.4. The only difference is that it highlights the weak property of 12^{th} family of Archimedean copula which has small dependence coverage.

Following this, we compare the Cot- and Csc-copula with some two-parameter family of copulas. This analysis was chosen because the flexibility of capturing tail dependence is increased with the number of parameters. Table 5.6 and Table 5.7 present the result of this analysis for CAC40-DAX and CAC40-SP500 respectively. These tables show the value of parameters for each copula together with tail dependences and goodness of fit criteria. For this study, Cramer-von Mises (CvM) and Kolmogorov-Smirnov (KS) criteria was used to explore the goodness of fit for the given data. The most striking result is the ability of Csc copula compared with two- parameter family of copulas. It is not surprising that the performance of Frank and BB8 copula is the weakest among all since the dependence among pairs are strong. Although a clear benefit of using BB1 and BB7 rather than t-student could not be identified in the analysis of tail dependence, the goodness of fit tests show the priority of t-copula. In conclusion, we can emphasis on the ability of Cot- and Csc-copula in capturing strong dependence among the random variables while considering the asymmetric upper and lower tail dependences.

Table 5.6: Two – parameter estimation of Archimedean copula CAC40-DAX

	Param 1	Param 2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS	P-value KS
t-Student	0.9034	2.5721	0.6947	0.6947	0.2374	0.9100	1.1731	0.7000
Frank	12.4895	0.0000	0.0000	0.0000	0.9807	0.0000	1.6703	0.0000
BB1	0.8374	2.4899	0.7172	0.6790	0.0456	0.3300	0.5127	0.6900
BB7	2.9951	2.8323	0.7829	0.7396	0.8101	0.0000	1.9207	0.0000
BB8	6.0000	0.8943	0.0000	0.0000	4.0178	0.0000	3.8716	0.0000
Cot	2.7413	-	0.7766	0.7123	0.1302	0.9000	0.7558	0.9900
Csc	1.6702	-	0.6603	0.7694	0.2108	1.0000	1.0583	1.0000

Table 5.7: Two – parameter estimation of Archimedean copula DAX-SP500

	Param 1	Param 2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS	P-value KS
t-Student	0.5807	2.8164	0.3738	0.3738	0.4136	0.9300	1.6495	0.7400
Frank	4.2508	0.0000	0.0000	0.0000	2.0382	0.0000	2.5851	0.0000
BB1	0.3586	1.4452	0.2625	0.3845	0.3955	0.0000	1.5514	0.0000
BB7	1.5604	0.7190	0.3813	0.4407	0.1015	0.1700	0.8911	0.1200
BB8	3.1238	0.8324	0.0000	0.0000	3.2899	0.0000	3.6042	0.0000
Cot	1.3609	-	0.6009	0.3358	1.8525	0.23	2.2746	0.13
Csc	0.8226	-	0.4306	0.4761	0.0543	0.97	0.6704	0.96

The same structured approach which was conducted for the first group is employed for the second group of pairs with rather weak dependence properties. The data set include pairs of Asia market, Europe- Asia and America- Asia indices. Figure 5.2 demonstrates a weak dependence among almost all pairs. The result of Chi- and K-plot also confirms this week dependence. Regarding the first analysis on the data set, we also consider Table 5.3 for empirical Kendall's τ , Spearman's ρ and tail dependences. This table is quite revealing in several ways. First, it shows the weak dependence of Asia market among themselves and also with America and Europe. In Asia market, STI-KLCI has the strongest dependence. It is interesting that the dependence between China and the other two countries is almost the same and is unrelated to country; while the dependence in Malaysian market is defined according to the specific countries. Generally, Asia's market is more dependent to European rather than American market. What is interesting in data is that while Singapore, STI, has the highest dependence with America and European market, China acts almost independently from both markets.

Considering the different range of dependence in Table 5.3, we choose three representative pairs of indices: DAX-STI as almost strong dependence from Asia-Eroup market, KLCI-SSE as relatively weak dependence represented from Asia and SSP500-SSE as the lowest range of dependence in all pairs. We believe the result of analysing these pairs can be extended for the other pairs as well since these set of pairs are representing all range of dependences on data.

We start by analysing the data with the one-parameter family of Archimedean copula. Table 5.8, Table 5.9 and Table 5.10 represent the results for these three pairs.

The best fit copula with regard to the goodness of fit test and tail dependence measure, given in Table 5.8, is Csc-copula. The result also shows that Cot-copula cannot capture the lower tail dependence in this range. Although Gumbel and Clayton show a relatively proper goodness of fit, they fail to measure both tail dependences. The result also indicates that the 12^{th} family of Archimedean copula is not able to cover this range of dependence.

Interestingly, the best copula to fit the data given in Table 5.9 is Csc-copula although the problem of overestimation of lower tail dependence is conspicuous. Finally, Table 5.9 shows that Csc-copula is not working for almost zero dependence range.

Table 5.8: One – parameter estimation of Archimedean copula DAX-STI

	Lower tail	Upper tail	Goodness of fit	parameter
Empirical tail	0.2750	0.2921		
Gumbel	0	0.2864	0.0003	1.2870
Clayton	0.266	0	0.0002	0.5235
12th family	0.6735	0.5151	0.0111	1.7533
Cot	0.5354	0.1321	0.0009	1.1094
CSC	0.3404	0.286	0.0001	0.6432
Csc2	0.2456	0.5858	0.0040	0.4937
Csc3	0.0607	0.5858	0.0021	0.2475

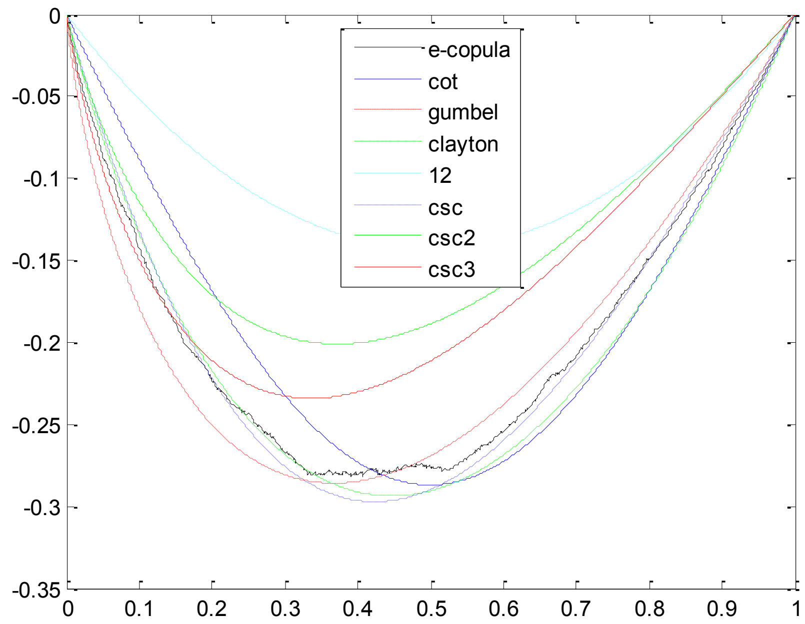


Figure 5.6: Goodness of fit representation of Table 5.8

Table 5.9: One – parameter estimation of Archimedean copula KLCI-SSE

	Lower tail	Upper tail	Goodness of fit	parameter
Empirical tail	0.05321	0.1201		
Gumbel	0	0.1187	0.0002	1.0968
Clayton	0.0493	0	0	0.2304
12th family	0.6531	0.4688	0.0179	1.6270
Cot	0.5062	0.0245	0.0022	1.0181
CSC	0.2892	0.1405	0.0002	0.5587
Csc2	0.1892	0.5858	0.0095	0.4163
Csc3	0	0.5858	0.0048	0.0374

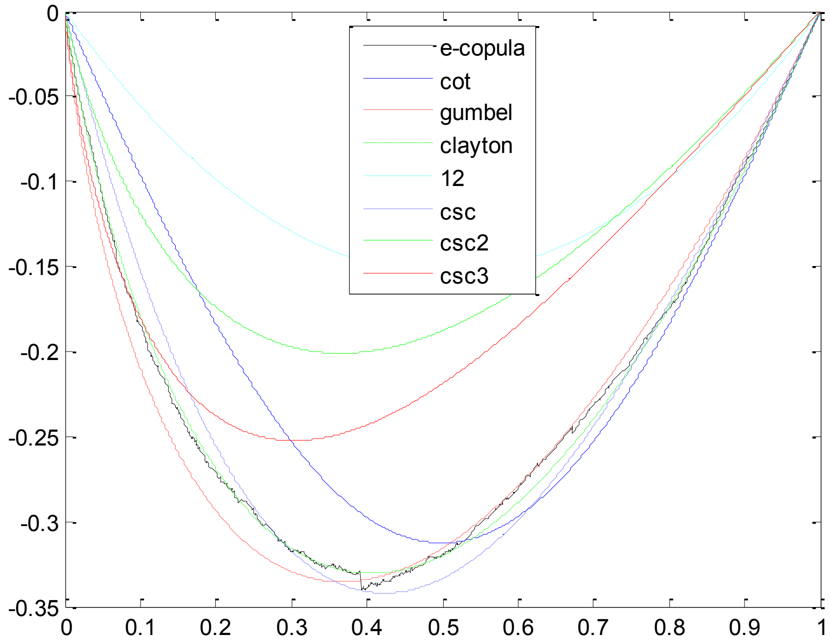


Figure 5.7: Goodness of fit representation of Table 5.9

Table 5.10: One – parameter estimation of Archimedean copula SP500-SSE

	Lower tail	Upper tail	Goodness of fit	parameter
Empirical tail	0.0124	0.0355		
Gumbel	0	0.0308	0	1.0229
Clayton	0	0	0	0.0525
12th family	0.6444	0.4482	0.0241	1.5775
Cot	0.5022	0.0088	0.0047	1.0064
CSC	0.2685	0.0703	0.0007	0.5272
Csc2	0.1599	0.5858	0.0145	0.3781
Csc3	0	0.5858	0.0075	0

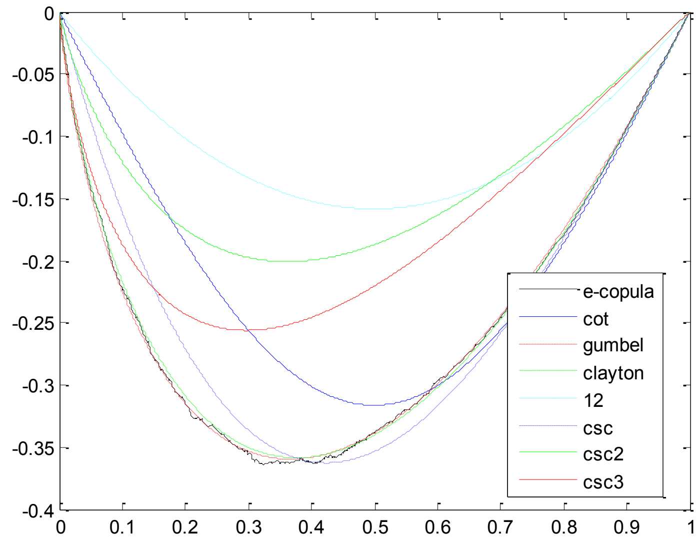


Figure 5.8: Goodness of fit representation of Table 5.10

We next compare the Cot and Ccs copula with the two–parameter family of copulas. The results are summarized in Table 5.11, Table 5.12 and Table 5.13. The results emphasizes that Csc-copula as the best fit for the first two pairs, DAX-STI and KLCI-SSE, of data compared with two-parameter family of copulas. Regarding the KLCI-SSE, t-copula is the best among all.

Table 5.11: Two – parameter estimation of Archimedean copula DAX-STI

	Param1	**Param2**	**Lower Tail**	**Upper Tail**	**Stat CVM**	**P-value CVM**	**Stat KS**	**P-value KS**
t-Student	0.3665	4.3842	0.1708	0.1708	0.1295	0.4600	0.8234	0.5800
Frank	2.3376	0.0000	0.0000	0.0000	0.9004	0.0000	1.7066	0.0000
BB1	0.3298	1.1341	0.1567	0.1573	0.0318	0.8900	0.5844	0.7000
BB7	1.1665	0.4338	0.2024	0.1884	0.0251	0.9700	0.4246	0.9900
BB8	6.0000	0.3476	0.0000	0.0000	1.1728	0.0000	2.0785	0.0000
Cot	1.1094	-	0.5354	0.1321	3.6744	0.3400	3.1552	0.4200
Csc	0.6432	-	0.3404	0.2860	0.2362	0.9900	0.9453	0.9800

Note: Significant at 5%

Table 5.12: Two – parameter estimation of Archimedean copula KLCI-SSE

	Param1	**Param2**	**Lower Tail**	**Upper Tail**	**Stat CVM**	**P-value CVM**	**Stat KS**	**P-value KS**
t-Student	0.1757	13.5842	0.0062	0.0062	0.1292	0.4400	0.9560	0.3300
Frank	1.0380	0.0000	0.0000	0.0000	0.4739	0.0000	1.4872	0.0000
BB1	0.1983	1.0240	0.0329	0.0322	0.0452	0.6700	0.7227	0.2400
BB7	1.0265	0.2155	0.0401	0.0355	0.0419	0.8000	0.6601	0.4700
BB8	6.0000	0.1765	0.0000	0.0000	0.5337	0.0000	1.5632	0.0000
Cot	1.0181	-	0.5062	0.0245	8.2010	0.0000	4.5625	0.0000
Csc	0.5587	-	0.2892	0.1405	1.1392	0.9900	2.0720	0.9600

Note: Significant at 5%

Table 5.13: Two – parameter estimation of Archimedean copula SP500-SSE

	Param1	Param2	Lower Tail	Upper Tail	Stat CVM	P-value CVM	Stat KS	P-value KS
t-Student	0.0359	15.3698	0.0012	0.0012	0.0965	0.6000	0.8311	0.4800
Frank	0.2173	0.0000	0.0000	0.0000	0.0591	0.4000	0.6342	0.4200
BB1	0.0411	1.0107	0.0000	0.0146	0.0638	0.3900	0.6484	0.4600
BB7	1.0135	0.0462	0.0000	0.0184	0.0587	0.4800	0.6264	0.5700
BB8	6.0000	0.0420	0.0000	0.0000	0.0592	0.5900	0.6354	0.6100
Cot	1.0064	-	0.5022	0.0088	15.4633	0.0000	5.9034	0.0000
Csc	0.5272	-	0.2685	0.0703	3.7655	0.0000	3.3406	0.0000

Note: Significant at 5%

This study set out to determine the ability of Cot- and Csc-copula in modelling of dependence between two random variables. One of the more significant findings to emerge from this study is that in high dependence range both Cot- and Csc-copula are capable even compare with two–parameter family of copulas. The results also emphasize on the capability of Csc copula in low range of dependence compared with even two parameter families of copulas.

5.2 Multivariate Vine Copula in Indices

Now we turn to the modelling multivariate joint distribution function using trigonometric copulas. The data sample consists of the first group of indices from the previous section; Europe and America indices. This sample is chosen because of strong dependence among random variables. Besides, we considered the Asian indices with weak dependence in Chapter 4 of this thesis.

The main issue addressed in this section is the importance of using trigonometric copula in modelling multivariate distribution functions. There are many alternative ways to building a multivariate distribution functions based on C- or D-vine structure by considering different order of variables. Therefore, we solve the problem of choosing vine structure before further analyses.

C-vine copulas are useful when one expects a variable to dominate the dependence with all other variables, while the guideline structure to use the D-vine copula is based on equivalent dependence among all variables. To compare the dependence between random variables we calculate the Kendall's τ for every variable. It is apparent from Table 5.14 that there is no significant difference in sum of the Kendall's τ for every index. According to this result, the D-vine structure is chosen to build multivariate distribution function. To reduce the effect of ordering of variables, we choose the same order for both models. The estimation of parameters has been done based on MLE and sequential estimation.

Table 5.14: Kendall's τ dependence between random variable

	CAC40	**DAX**	**FTSE**	**SP500**
CAC40	1.0000	0.7188	0.6800	0.3695
DAX	0.7188	1.0000	0.6142	0.3943
FTSE	0.6800	0.6142	1.0000	0.3521
SP500	0.3695	0.3695	0.3521	1.0000
SUM	2.7683	2.7025	2.6463	2.1159

The result of the estimation based on sequential and MLE for both models are presented in Table 5.15. Model MD1 choose the optimal bivariate copula from excising copula while the second one includes the trigonometric copula as well. Based on the results, in first model the best bivariate copula is t-copula. Interestingly, the Csc copula is the best option for all pairs in the second model.

Table 5.15: Estimation parameters based on Sequential and MLE

Method	No	Tree	Copula	Sequential		MLE	
				Par 1	Par 2	Par 1	Par 2
MD1	1	CAC40-DAX	t	-0.00599	6.033364	-0.01416	5.996535
		DAX-FTSE	t	-0.02931	4.275707	-0.01966	4.455765
		FTSE-SP500	t	-0.02499	4.811807	-0.03603	5.48559
	2	CAC40-FTSE | DAX	t	-0.03567	3.806965	-0.03575	3.82682
		DAX-SP500 | FTSE	t	-0.06262	3.708806	-0.07108	3.534773
	3	CAC-SP | DAX, FTSE	t	-0.05115	4.310417	-0.05127	4.250052
MD2	1	CAC40-DAX	Csc	1.016719	-	1.018019	-
		DAX-FTSE	Csc	1.014252	-	1.016095	-
		FTSE-SP500	Csc	1.012815	-	1.013139	-
	2	CAC40-FTSE| DAX	Csc	1.012334	-	1.012804	-
		DAX-SP500 | FTSE	Csc	1.013516	-	1.013719	-
	3	CAC-SP | DAX, FTSE	Csc	1.022395	-	1.02296	-

We compared these two models based on AIC and BIC goodness in Table 5.16. Results show a significant difference in AIC, BIC and Log-likelihood value between two models. Obviously, the second method, MD2, performs better than the first model. Further analysis on the number of parameters also show the advantage of the second model towards the first one where the number of parameters is reduced by a half, which is important in reduction of error estimation.

Table 5.16: AIC and BIC goodness of fit

	MC1	MC2
AIC MLE	-325.8528	-859.713
AIC seq.	-322.4941	-857.191
BIC MLE	397.77092	-787.089
BIC seq.	407.12965	-784.568
Log-likelihood MLE	289.42641	441.8566
Log-likelihood Seq.	287.24705	440.5957
Parameters	12	6

Note: Significant at 5%

This chapter has shown the importance of bivariate and multivariate trigonometric copula in finance data. One of the more significant finding to emerge from this study is the ability of Csc copula in modelling the multivariate probability joint distribution function with different range of dependence.

Chapter 6: Conclusion

This dissertation has investigated the importance of trigonometric copula as a one-parameter family of Archimedean copula in modelling joint probability distribution function of random variables. In this investigation, the main aim was to assess the dependence properties of copulas which were built on trigonometric generators.

This study has focused on five new Archimedean copulas based on trigonometric generators. The result of the dependence analysis on these functions suggests that the Cotangent (Cot) and Cosecant (Csc)-copulas emerged as reliable copula in modelling joint probability distribution function of random variables. The result of this study illustrates that both Cot and Csc-copula possess the flexible upper and lower tail dependences. The Csc-copula, however, covers a wider dependence than Cot-copula. The second major finding of this study is the construction of multivariate vine copula using Cot and Csc-copula as its basis. The advantage of having asymmetric upper and lower tail dependences with a single parameter copula contributes to constructing multivariate probability distribution function with the hope of less estimation error. Finally, the findings of this study suggest that the use of Cot and Csc-copula as a reliable tool for modelling joint probability distribution function of real data sets.

A number of caveats need to be noted with regards to the present study. Although Csc-copula cover all dependence between $[0,1]$, the dependence coverage of Cot-copula is in the range of $[0.18,1]$. The lower tail dependence in Csc-copula is $[0.25,1]$, whereas the Cot-copula suffer as it ranges in $[0.5,1]$. The limitation of the lower tail dependences has effect on modelling random variables with lower tail dependence that exceed the coverage ranges. This limitation was shown clearly for the Cot-copula in our simulation studies and application parts where the Csc-copula was chosen as the superior model.

For further research, the following work should be undertaken: with reference to the trigonometric copula that is, a study similar to one–parameter trigonometric copula should be carried out on two parameter families of trigonometric generators together with its dependence properties to compare by one-parameter family. As for the multivariate vine structure, future analysis on the choice of optimal structure of vine should be carried out while considering different dependences among the random variables that should be explored. The application of these finding can be implemented for hydrology and marketing data.

Appendix

Appendix A

Cot copula is Left Tail Decreasing (LTD).

The property $g''(t) \geq 0 \quad \forall\ 0 \leq t \leq 1$ when $\theta \geq 1.05$ implies that Cot copula is LTD. $g''(t)$ is define as

$$g''(t) = \frac{t^{\frac{1}{\theta}-2}\left[\left((\theta-1)+(\theta+1)\,t^{\frac{2}{\theta}}\right) arc\cot\left(t^{\frac{1}{\theta}}\right) - t^{\frac{1}{\theta}}\right]}{\left[\theta\left(1+t^{\frac{2}{\theta}}\right) arc\cot\left(t^{\frac{1}{\theta}}\right)\right]^2}.$$

Then the result of simulation for different value of $\theta \geq 1$ and for $\forall\ t \leq 0 \leq 1$ is given in Table A.1:

Table A.1: Simulation result of $g''(t)$ to identify Min and Max value.

θ	**1.05**	**1.1**	**1.15**	**1.20**	**1.25**	**1.30**	**2**	**3**	**4**	**5**	**6**	**7**
Min	0.0309	0.5406	0.9081	1.0516	1.1635	1.2728	2.6416	4.3791	6.0332	7.6540	9.2581	10.8527
Max	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Matlab Programming is provided as

```
function MXMI=cotplot

a=[1 1.05 1.1 1.15 1.20 1.25 1.30 2 3 4 5 6 7];
t= [0:0.01:1];
for i=1:length(a)

    term1= t.^((1/a(i))-2);
    term2= (a(i)-1);
```

```
    term3= (a(i)+1).*(t.^(2/a(i)));
    term4= acot(t.^(1/a(i)));
    term5=term2+term3;
    term6=term5.*term4;
    term7= term6-t.^1/a(i);
    g=term1.*term7;

        %plot(t, g, 'r'  )

MXMI(:,i) = [min(g), max(g)];
end
```

CotII copula is not SI/SD

Table A.2: Simulation result to identify Min and Max value for $g''(t)$.

θ	1	1.05	2	3	4	5	6	7
Min	-2.00	-1.98	-1.80	-1.73	-1.70	-1.68	-1.66	-1.65
Max	0.00	0.04	0.41	0.54	0.61	0.65	0.68	0.69

θ	8	9	10	11	40	70	100	1000	1000000
Min	-1.65	-1.64	-1.64	-1.63	-1.60	-1.60	-1.60	-1.60	-1.59
Max	0.71	0.72	0.73	0.74	0.79	0.80	0.80	0.81	0.81

As shown in table the value of $g''(t)$ is negative for min value while it is positive for maximum value. Therefore $g''(t)$ is not convex of concave function.

Matlab Programming is provided

```
function MXMI= cot2plot
a=[1 1.05 2 3 4 5 6 7 8 9 10 11 40 70 100 1000 1000000 ];
t= [0:0.01:1];
for i=1:length(a)
    term1=1-1./a(i);
    term2=1./(1+t.^2);
    term3=(acot(t)).^(-2);
```

```
    term4=(2.*t.^2-2)./(1+t.^2).^2;

    g = term1.*term2.*term3 + term4;

MXMI(:,i) = [min(g), max(g)];
end
```

CotII copula is not PKD/PQD

Table A.3: Simulation result to identify Min and Max value for $g'(t)$.

θ	1	1.05	2	3	4	5	6	7	8
Min	-0.06	-0.05	-0.03	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01
Max	0.20	0.19	0.10	0.07	0.05	0.04	0.03	0.03	0.03

θ	9	10	11	40	70	100	1000	1000000
Min	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00
Max	0.02	0.02	0.02	0.01	0.00	0.00	0.00	0.00

Matlab program is provided for this simulation.

```
function MXMI = cot2pkd
a=[1 1.05 2 3 4 5 6 7 8 9 10 11 40 70 100 1000 1000000 ];
t= [0:0.01:1];
for i=1:length(a)
    term1= a(i).*(1+t.^2).*acot(t);
    term2 = log(((2/pi).*acot(t)).^(1./a(i)));

    g=((t./term1)+term2)./(t.^2);

     plot(t, g)
MXMI(:,i) = [min(g), max(g)];
end
```

Csc copula is SI.

Matlab programming for simulation

```
function MXMI= cscSIplot
     a=[0.55 0.6 0.7 0.8 0.9 1 ];
t= [0:0.01:1];
for i=1:length(a)
    A= t.^(1./(2.*a(i))-1);

    B= t.^(1./a(i))+1;

    C= sqrt(t.^(1./a(i))+2);

    K= 2./(pi.*a(i));

    g = K.*(A./(B.*C));

MXMI(:,i) = [min(g), max(g)];
plot (t,g)
title('Plot of f(t) for CSC for different \Theta')
hold on
end
```

CscII copula is SI.

Matlab programming for simulation

```
function MXMI= csc2SIplot
a=[ 0.6 0.7 0.8 0.9 1 2 3 4 5];
t= [0:0.01:1];
for i=1:length(a)
    A= abs(t+1).*sqrt(t.^2+2.*t);
    B= acsc(t+1).^((1./a(i))-1);

    C= 4./(a(i).*(pi.^2));

    g = C.*(B./A);

MXMI(:,i) = [min(g), max(g)];

plot(t, g)
title('Plot of f(t) for CSCII for different \Theta')
hold on
end
```

Appendix B

MATLAB programming for Estimation and goodness of fit test of trigonometric copulas:

```
function [ y ,d , p] = tchek2(x)

[k,l ] = size(x);
for i=1:k
    for j=1:l
        if x(i,j)==1
             x(i,j)=0.999;
        elseif x(i,j)==0
            x(i,j)=0.0001;
        end
    end
end

pclayton = copulafit ('Clayton', x);
pgumbel = copulafit('gumbel', x);
tclayton  =  2^(- 1/  pclayton)  ;  %  lower  tail  dependnece  in  Clayton
(alpha >= 0 )
tgumbel = 2- 2^(1/pgumbel);
y1 = [0 , tgumbel; tclayton , 0 ];

pcot  = cotfit (x(:,1), x(:,2));
pcot2 = cot2fit (x(:,1), x(:,2));
p12 = fit12 (x(:,1), x(:,2));
pcsc  = cscfit (x(:,1), x(:,2));
pcsc2 = csc2fit (x(:,1), x(:,2));
pcsc3 = csc3fit(x(:,1), x(:,2));

pkcot=kendalcot(x);


tucot =  2- 2^(1/pcot);   %  upper tail similar to Gumbel.
tlcot =  2^(- 1/ pcot);   %  lower tail similar to Clayton.

tucot2 =  0 ;   %  upper tail similar to Gumbel.
tlcot2 =  2^(- 1/ pcot2);   %  lower tail similar to Clayton.

tu12 =  2- 2^(1/p12);   %  upper tail similar to Gumbel.
tl12 =  2^(- 1/ p12);   %  lower tail similar to Clayton.

tucsc =  2- 2^(1/(2*pcsc));   %  upper tail simila to Gumbel.
tlcsc =  2^(- 1/ pcsc);   %  lower tail similar to Clayton.

tucsc2 =  2 - sqrt(2);   %  upper tail similar to Gumbel.
tlcsc2 =  2^(- 1/ pcsc2);   %  lower tail similar to Clayton. % change
from 4 to 2

tucsc3 =  2 - sqrt(2) ;   %  upper tail similar to Gumbel.
tlcsc3 =  2^(- 1/ pcsc3);   %  lower tail similar to Clayton.

[dcot, w, q, z]= kfun ('cot', x, pcot);
plot(w,q,'k','LineWidth',1) %comparsion of u - K(u)
hold on
plot(w,z,'b')

[dgumbel, w, q, z] = kfun ('gumbel', x, pgumbel);
plot(w,z,'-.r')
```

```
[dclayton, w, q, z] = kfun ('clayton', x, pclayton);
plot(w,z, '--g')

[d12, w, q, z] = kfun ('12', x, p12);
plot(w,z,':c')

[dcsc, w, q, z] = kfun ('csc', x, pcsc);
plot(w,z,'b:')

[dcsc2, w, q, z] = kfun ('csc2', x, pcsc2);
plot(w,z,'g')

[dcsc3, w, q, z] = kfun ('csc3', x, pcsc3);
plot(w,z,'r')

%legend('e-copula','cot','clayton', '12', 'csc', 'csc2', 'csc3', 0)
legend('e-copula','cot','gumbel','clayton','12','csc','csc2',...
,'csc3', 0)
hold off
y = [y1 ; tl12 , tu12 ;tlcot ,  tucot ; tlcsc , tucsc ; tlcsc2 ,
tucsc2 ; tlcsc3 , tucsc3 ];
y = round(y*10000)/10000;
p = [ pgumbel ;  pclayton ; p12 ; pcot ; pcsc ; pcsc2 ; pcsc3];
p = round(p*10000)/10000;
d = [ dgumbel ;  dclayton ; d12 ; dcot ; dcsc ; dcsc2 ; dcsc3];
d = round(d*10000)/10000;
```

R program for Estimation and Goodness of fit test:

```
# Data

secdfKLCI=ecdfKLCI[1:500]
secdfSP=ecdfSP[1:500]
secdfSTI=ecdfSTI[1:500]

# PAirs Function
panel.cor <- function(x, y, digits=2, prefix="", cex.cor, ...)
{
  usr <- par("usr"); on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r <- abs(cor(x, y))
  txt <- format(c(r, 0.123456789), digits=digits)[1]
  txt <- paste(prefix, txt, sep="")
  if(missing(cex.cor)) cex.cor <- 0.4/strwidth(txt)
  text(0.5, 0.5, txt, cex = cex.cor * r)
}

x=matrix(c(secdfKLCI, secdfSP, secdfSTI),1000,3)
colnames(x, do.NULL = FALSE)
colnames(x) <- c("KLCI", "SP500", "STI")
pairs(x, lower.panel=panel.cor)

# Kendalls tau Correlation

ken1 = cor(ecdfKLCI,ecdfSP,method="kendall")
```

```
ken2 = cor(ecdfKLCI,ecdfSTI, method= "kendall")
ken3 = cor(ecdfSP,ecdfSTI, method= "kendall")

ken=c(ken1, ken2, ken3)

write.table(ken  , file  =  "Kendall.csv",  sep  =  ",",  col.names  =  NA,
      qmethod = "double")


# Estimation of Parametres

u1=ecdfCAC40
u2=ecdfDAX

c1= BiCopEst(u1,u2,family=1,method="mle",se=TRUE)
c2= BiCopEst(u1,u2,family=2,method="mle",se=TRUE)
c3= BiCopEst(u1,u2,family=3,method="mle",se=TRUE)
c4= BiCopEst(u1,u2,family=4,method="mle",se=TRUE)
c5= BiCopEst(u1,u2,family=5,method="mle",se=TRUE)
c7= BiCopEst(u1,u2,family=7,method="mle",se=TRUE)
c8= BiCopEst(u1,u2,family=8,method="mle",se=TRUE)
c9= BiCopEst(u1,u2,family=9,method="mle",se=TRUE)
c10= BiCopEst(u1,u2,family=10,method="mle",se=TRUE)


# Rewrite for parameters table

p1  = c(c1$par,  c1$par2)
p2  = c(c2$par,  c2$par2)
p3  = c(c3$par,  c3$par2)
p4  = c(c4$par,  c4$par2)
p5  = c(c5$par,  c5$par2)
p7  = c(c7$par,  c7$par2)
p8  = c(c8$par,  c8$par2)
p9  = c(c9$par,  c9$par2)
p10 = c(c10$par, c10$par2)


outpara = matrix( c(p1, p2, p3, p4, p5, p7, p8, p9, p10), 9, 2, byrow=
"T")
write.table(outpara  ,  file  =  "parameter-SP-STI.csv",  sep  =  ",",
col.names = NA, qmethod = "double")


# Rewrite for estandard error table

se1  = c(c1$se, c1$se2)
se2  = c(c2$se, c2$se2)
se3  = c(c3$se, c3$se2)
se4  = c(c4$se, c4$se2)
se5  = c(c5$se, c5$se2)
se7  = c(c7$se, c7$se2)
se8  = c(c8$se, c8$se2)
se9  = c(c9$se, c9$se2)
se10 = c(c10$se, c10$se2)


outpara = matrix( c(se1, se2, se3, se4, se5, se7, se8, se9, se10), 9,
2, byrow= "T")
```

```
write.table(outpara  ,  file  =  "Estandard-SP-STI.csv",  sep  =  ",", col.names = NA, qmethod = "double")
```

 # tail dependnece

```
tail1  = BiCopPar2TailDep(1,  c1$par,  c1$par2  )
tail2  = BiCopPar2TailDep(2,  c2$par,  c2$par2  )
tail3  = BiCopPar2TailDep(3,  c3$par,  c3$par2  )
tail4  = BiCopPar2TailDep(4,  c4$par,  c4$par2  )
tail5  = BiCopPar2TailDep(5,  c5$par,  c5$par2  )
tail7  = BiCopPar2TailDep(7,  c7$par,  c7$par2  )
tail8  = BiCopPar2TailDep(8,  c8$par,  c8$par2  )
tail9  = BiCopPar2TailDep(9,  c9$par,  c9$par2  )
tail10 = BiCopPar2TailDep(10, c10$par, c10$par2 )

tail =matrix(c(tail1, tail2, tail3, tail4, tail5, tail7, tail8, tail9, tail10 ),9,2, byrow="T")

write.table(tail,   file="taildependneceSP-STI.csv",   sep   =   ",", col.names = NA, qmethod = "double")
```

 # Goodness of fit calculation

```
gofc1  =  BiCopGofKendall(u1,u2, family= 1,  B=100, level=0.05)
gofc2  =  BiCopGofKendall(u1,u2, family= 2,  B=100, level=0.05)
gofc3  =  BiCopGofKendall(u1,u2, family= 3,  B=100, level=0.05)
gofc4  =  BiCopGofKendall(u1,u2, family= 4,  B=100, level=0.05)
gofc5  =  BiCopGofKendall(u1,u2, family= 5,  B=100, level=0.05)
gofc7  =  BiCopGofKendall(u1,u2, family= 7,  B=100, level=0.05)
gofc8  =  BiCopGofKendall(u1,u2, family= 8,  B=100, level=0.05)
gofc9  =  BiCopGofKendall(u1,u2, family= 9,  B=100, level=0.05)
gofc10 = BiCopGofKendall(u1,u2, family= 10, B=100, level=0.05)
```

 # rewrite the parametre for table:

```
gofcvm = c(gofc1$statistic.CvM, gofc2$statistic.CvM,
gofc3$statistic.CvM, gofc4$statistic.CvM, gofc5$statistic.CvM,
gofc7$statistic.CvM, gofc8$statistic.CvM, gofc9$statistic.CvM,
gofc10$statistic.CvM)

pvaluecvm = c(gofc1$p.value.CvM, gofc2$p.value.CvM, gofc3$p.value.CvM,
gofc4$p.value.CvM, gofc5$p.value.CvM, gofc7$p.value.CvM,
gofc8$p.value.CvM, gofc9$p.value.CvM, gofc10$p.value.CvM)

gofKS = c(gofc1$statistic.KS, gofc2$statistic.KS, gofc3$statistic.KS,
gofc4$statistic.KS, gofc5$statistic.KS,gofc7$statistic.KS,
gofc8$statistic.KS, gofc9$statistic.KS, gofc10$statistic.KS)


pvalueKS = c(gofc1$p.value.KS, gofc2$p.value.KS, gofc3$p.value.KS,
gofc4$p.value.KS, gofc5$p.value.KS, gofc7$p.value.KS,
gofc8$p.value.KS, gofc9$p.value.KS, gofc10$p.value.KS)



outgof = matrix(c(gofcvm, pvaluecvm, gofKS, pvalueKS),9,4, byrow="F")
write.table(outgof, file="gof-SP-STI.csv", sep=",", col.names=NA,
qmethod="double")
```

```
#********************** MULTIVARITE COPULA *********************

# data wit specific order

#dataCvine1 = matrix(c(ecdfSTI, ecdfSP, ecdfKLCI), 2634, 3, byrow="T")
dataCvine1  = matrix(c(ecdfKLCI, ecdfSTI, ecdfSP), 2634, 3, byrow="T")
# dataDvine = c(ecdfKLCI, ecdfSTI, ecdfSP)

#Selection of family

select1=
CDVineCopSelect(dataCvine1,type=1,familyset=c(1:10,13,14,23,24))

fam = select1$family

#Estimation of CDvine copula of data sequential method

est = CDVineSeqEst(dataCvine1, fam , 1, method="mle", se=TRUE,
progress=TRUE)

cdest = matrix( c(est$par, est$par2), 3,2, byrow="F")
cdse  = matrix( c(est$se,  est$se2) , 3,2, byrow="F")

write.table(cdest, file="MD2-Seq-ESTIMATION.csv", sep=",",
col.names=NA, qmethod="double")

write.table(cdse, file="MD2-Seq-serror.csv", sep=",", col.names=NA,
qmethod="double")

#Estimation of CDvine copula of data via MLE

estMLE = CDVineMLE(dataCvine1, fam , start=cdest[,1], start2=
cdest[,2], 1 )

cdestMLE  = matrix( c(estMLE$par, estMLE$par2), 3,2, byrow="F")
# cdseMLE = matrix( c(estMLE$se,  estMLE$se2), 3,2, byrow="F")

write.table(cdestMLE, file="MD2-MLE-ESTIMATION.csv", sep=",",
col.names = NA, qmethod="double")
# write.table(cdseMLE, file="MD2-MLE-Serror.csv", sep=",",
col.names=NA, qmethod="double")


# GOODNESS OF FIT TEST

MLEgofAIC= CDVineAIC(dataCvine1,fam, estMLE$par, estMLE$par2,1)
SeqgofAIC= CDVineAIC(dataCvine1,fam, est$par, est$par2,1)

MLEgofBIC= CDVineBIC(dataCvine1,fam, estMLE$par, estMLE$par2,1)
SeqgofBIC= CDVineBIC(dataCvine1,fam, est$par, est$par2,1)

# Log-liklihood function

MLEliklihood= CDVineLogLik(dataCvine1,fam, estMLE$par, estMLE$par2,1)
seqliklihood= CDVineLogLik(dataCvine1,fam, est$par, est$par2,1)
```

```
# GOF and Liklihood function
gof=c(MLEgofAIC$AIC, SeqgofAIC$AIC, MLEgofBIC$BIC, SeqgofBIC$BIC,
MLEliklihood$loglik , seqliklihood$loglik)

write.table(gof, file="MD2-gof.csv", sep=",", col.names=NA,
qmethod="double")
```

R programm for Multivaraite simulation in Chapter 5

```
# Kendalls tau Correlation

u1= ecdfCAC40
u2=ecdfDAX
u3=ecdfFTSE
u4=ecdfSP500

ken1 = cor(u1,u2,method="kendall")
ken2 = cor(u1,u3, method= "kendall")
ken3 = cor(u1,u4, method= "kendall")
ken4 = cor(u2,u3, method="kendall")
ken5 = cor(u2,u4, method= "kendall")
ken6 = cor(u3,u4, method= "kendall")


ken=c(ken1, ken2, ken3, ken4, ken5, ken6)

write.table(ken , file = "Kendall.csv", sep = ",", col.names = NA,
qmethod = "double")



#**************************** MULTIVARITE COPULA ***************

# data wit specific order

#dataCvine1 = matrix(c(ecdfSTI, ecdfSP, ecdfKLCI), 2634, 3, byrow="T")
dataCvine2  = matrix(c( u1, u2, u3, u4), 3140, 4, byrow="T")

# dataDvine=c(ecdfKLCI, ecdfSTI, ecdfSP)

#Selection of family

select1=
CDVineCopSelect(dataCvine2,type=2,familyset=c(1:10,13,14,23,24))

#fam = select1$family

fam = c(7,7,7,7,7,7)

#Estimation of CDvine copula of data sequential method

est = CDVineSeqEst(dataCvine2, fam , 2, method="mle", se=TRUE,
progress=TRUE)

cdest = matrix( c(est$par, est$par2), 6,2, byrow="F")
cdse = matrix( c(est$se,  est$se2), 6,2, byrow="F")
```

```
write.table(cdest, file="MD2-Seq-ESTIMATION.csv", sep=",",
col.names=NA, qmethod="double")
write.table(cdse, file="MD2-Seq-serror.csv", sep=",", col.names=NA,
qmethod="double")

#Estimation of CDvine copula of data via MLE

estMLE = CDVineMLE(dataCvine2, fam , start=cdest[,1], start2=
cdest[,2], 2 )

cdestMLE  = matrix( c(estMLE$par, estMLE$par2), 6,2, byrow="F")
# cdseMLE = matrix( c(estMLE$se,  estMLE$se2), 3,2, byrow="F")

write.table(cdestMLE, file="MD2-MLE-ESTIMATION.csv", sep=",",
col.names=NA, qmethod="double")
# write.table(cdseMLE, file="MD2-MLE-Serror.csv", sep=",",
col.names=NA, qmethod="double")


# GOODNESS OF FIT TEST

MLEgofAIC = CDVineAIC(dataCvine2,fam, estMLE$par, estMLE$par2,2)
SeqgofAIC = CDVineAIC(dataCvine2,fam, est$par, est$par2,2)

MLEgofBIC = CDVineBIC(dataCvine2,fam, estMLE$par, estMLE$par2,2)
SeqgofBIC = CDVineBIC(dataCvine2,fam, est$par, est$par2,2)

# Log-liklihood function

MLEliklihood = CDVineLogLik(dataCvine2,fam, estMLE$par, estMLE$par2,2)
seqliklihood = CDVineLogLik(dataCvine2,fam, est$par, est$par2,2)

# GOF and Liklihood function
gof = c(MLEgofAIC$AIC, SeqgofAIC$AIC, MLEgofBIC$BIC, SeqgofBIC$BIC,
MLEliklihood$loglik , seqliklihood$loglik)

write.table(gof, file="MD2-gof-t-copula.csv", sep=",", col.names=NA,
qmethod="double")
```

References

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics, 44*(2), 182-198. doi: 10.1016/j.insmatheco.2007.02.001

Abdous, B., Genest, C., & Rémillard, B. (2005). Dependence properties of meta-elliptical distributions *Statistical modeling and analysis for complex data problems* (pp. 1-15).

AghaKouchak, A., Bárdossy, A., & Habib, E. (2010). Conditional simulation of remotely sensed rainfall data using a non-Gaussian v-transformed copula. *Advances in Water Resources, 33*(6), 624-634.

Ane, T., & Kharoubi, C. (2003). Dependence structure and risk measure. *Journal of Business, 76*(3), 411-438.

Ang, A., & Bekaert, G. (2002). Short rate nonlinearities and regime switches. *Journal of Economic Dynamics and Control, 26*(7-8), 1243-1274.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios* 1. *Journal of Financial Economics, 63*(3), 443-494.

Avérous, J., & Dortet-Bernadet, J. L. (2004). Dependence for Archimedean copulas and aging properties of their generating functions. *Sankhyā: The Indian Journal of Statistics, 66*(4), 607-620.

Bae, K. H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *Review of Financial studies, 16*(3), 717-763.

Balakrishnan, N., & Lai, C. D. (2009). *Continuous bivariate distributions* (Vol. 171): Springer Verlag.

Barbe, P., Genest, C., Ghoudi, K., & Rémillard, B. (1996). On Kendall's process. *Journal of Multivariate Analysis, 58*(2), 197-229.

Barlow, R. E., & Proschan, F. (1975). Statistical theory of reliability and life testing: probability models: DTIC Document.

Bedford, T., & Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence, 32*(1), 245-268.

Bedford, T., & Cooke, R. M. (2002). Vines--a new graphical model for dependent random variables. *The Annals of Statistics, 30*(4), 1031-1068.

Berg, D. (2009). Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance, 15*(7-8), 675-701.

Berg, D., & Bakken, H. (2005). A goodness-of-fit test for copulae based on the probability integral transform. *Note, The Norwegian Computing Centre*

Boole, G. (1854). An investigation into the laws of thought. *Walton and Maberly, London.[Reprinted (1951) Dover, New York.]*.

Brechmann, E. C., & Schepsmeier, U. Modeling dependence with C-and D-vine copulas: The R-package CDVine.

Breymann, W., Dias, A., & Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *3*(1).

Charpentier, A., Fermanian, J. D., & Scaillet, O. (2007). The estimation of copulas: theory and practice *Copulas: From theory to Application in Finance. Risk Publications* (pp. 35-60).

Chatterjee, S. K. (2003). *Statistical thought: a perspective and history*: Oxford University Press, USA.

Choroś, B., Ibragimov, R., & Permiakova, E. (2010). Copula estimation *Workshop on Copula Theory and Its Applications* (pp. 77-91). Dortrecht (NL): Springer.

Czado, C. (2010). Pair-copula constructions of multivariate copulas *Workshop on Copula Theory and its Applications* (pp. 93-109). Dortrecht (NL): Springer.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.(5), 65*(6), 274–292.

Deheuvels, P. (1981). A Kolmogorov-Smirnov type test for independence and multivariate samples. *Rev. Roumaine Math. Pures Appl, 26*(2), 213-226.

Demarta, S., & McNeil, A. J. (2005). The t copula and related copulas. *International Statistical Review, 73*(1), 111-129.

Dorey, M., & Joubert, P. (2007). Modelling copulas: an overview. *The Staple Inn Actuarial Society*.

Durante, F., & Sempi, C. (2010). Copula theory: an introduction *Workshop on Copula Theory and Its Applications* (pp. 3-31). Dortrecht (NL): Springer.

Edgeworth, F. Y. (1893). Statistical correlation between social phenomena. *J. R. Statist. Soc, 58*, 670-675.

Elizalde, A. (2006). Credit risk models I: Default correlation in intensity models. *Documentos de Trabajo (CEMFI),* (5), 1. Retrieved from

Embrechts, P., McNeil, A., & Straumann, D. (1999). Correlation: pitfalls and alternatives. *RISK-LONDON-RISK MAGAZINE LIMITED-, 12*, 69-71.

Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls *Risk management: value at risk and beyond* (pp. 176-223): Cambridge University Press, Cambridge.

Fang, H. B., Fang, K. T., & Kotz, S. (2002). The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis, 82*(1), 1-16.

Fermanian, J. D., & Scaillet, O. (2003). Nonparametric estimation of copulas for time series. *Journal of Risk, 5*, 25-54.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon Sect. A, 9*, 53–77.

Frees, E. W., & Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal, 2*(1).

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland, 15*, 246-263.

Genest, C., & Favre, A. C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering, 12*, 347.

Genest, C., Gendron, M., & Bourdeau-Brien, M. (2009). The advent of copulas in finance. *The European Journal of Finance, 15*(7), 609-618.

Genest, C., Ghoudi, K., & Rivest, L. P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika, 82*(3), 543-552.

Genest, C., & MacKay, J. (1986a). Copules archimédiennes et families de lois bidimensionnelles dont les marges sont données. *Canadian Journal of Statistics, 14*(2), 145-159.

Genest, C., & MacKay, J. (1986b). The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician, 40*(4), 280-283.

Genest, C., QUESSY, J. F., & Rémillard, B. (2006). Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation. *Scandinavian Journal of Statistics, 33*(2), 337-366.

Genest, C., & Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de I'Institut Henri Poincare. Probabilit ´ es et Statistiques, 44*(6), 1096-1127.

Genest, C., Remillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics, 44*(2), 199-213.

Genest, C., & Rivest, L. P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association, 88*(423), 1034-1043.

Hahn, C., Wagner, F., & Pfingsten, A. (2002). An empirical investigation of the rank correlation between different risk measures. Westfalische Wilhelms-Universitat Munster: EFA 2002 Berlin Meetings.

Harris, R. (1970). A multivariate definition for increasing hazard rate distribution functions. *The Annals of Mathematical Statistics, 41*(2), 713-717.

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods* (2nd edn ed.): Wiley-Interscience.

Hong, Y., Tu, J., & Zhou, G. (2007). Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial studies, 20*(5), 1547-1581.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis, 46*(2), 262-282.

Joe, H. (1996). Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters. In I. L. R¨uschendorf, B. Schweizer & M. D. Taylor (Eds.), *Distributions with Fixed Marginals and Related Topics* (Vol. 28, pp. 120-141): JSTOR.

Joe, H. (1997). *Multivariate models and dependence concepts*: Chapman & Hall/CRC.

Joe, H., & Hu, T. (1996). Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of Multivariate Analysis, 57*(2), 240-265.

Joe, H., Li, H., & Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis, 101*(1), 252-270.

Jogdeo, K. (1982). Concepts of Dependence. In S. Kotz & L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 1, pp. 324-334). New York: Wiley Online Library.

Kimeldorf, G., May, J. H., & Sampson, A. R. (1980). Concordant and Discordant Monotone Correlations and Their Evaluation by Nonlinear Optimization. In S. H. Zanakis & J. S. Rustagi (Eds.), *TMS/Studies in the Management Sciences* (Vol. 19). North-Holland, Amsterdam.

Kimeldorf, G., & Sampson, A. R. (1978). Monotone dependence. *The Annals of Statistics, 6*(4), 895-903.

Kimeldorf, G., & Sampson, A. R. (1987). Positive dependence orderings. *Annals of the Institute of Statistical Mathematics, 39*(1), 113-128.

Klugman, S. A., & Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics, 24*(1), 139-148.

Kotz, S., & Nadarajah, S. (2004). *Multivariate t distributions and their applications*: Cambridge Univ Pr.

Kruskal, W. H. (1958). Ordinal measures of association. *Journal of the American Statistical Association, 53*(284), 814-861.

Kurowicka, D., & Cooke, R. (2003). A parameterization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications, 372*(10), 225-251.

Kurowicka, D., & Joe, H. (2010). *Dependence Modeling: Vine Copula Handbook*: World Scientific Publishing Co. Pte. Ltd.

Lancaster, H. (1982). Dependence, measures and indices of. In K. S & J. NL (Eds.), *Encyclopedia of Statistical Sciences* (Vol. 2, pp. 334-339): Wiley Online Library.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics, 37*(5), 1137-1153.

Lehmann, E. L., & D'Abrera, H. J. M. (1975). *Nonparametrics: statistical methods based on ranks* (Vol. 351): Holden-Day San Francisco.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance, 56*(2), 649-676.

Mari, D. D., Kotz, S., & ebrary, I. (2001). *Correlation and dependence*: Imperial College Press.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*: Princeton Univ Pr.

McNeil, A. J., & Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and 1-norm symmetric distributions. *The Annals of Statistics, 37*(5B), 3059-3097.

Moore, D. S., & Spruill, M. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics*, 599-616.

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression: a second course in statistics* (Vol. 1): Addison-Wesley, Reading

Muller, A., & Scarsini, M. (2005). Archimedean copulae and positive dependence. *Journal of Multivariate Analysis, 93*(2), 434-445.

Nelsen, R. B. (2006). *An introduction to copulas*: Springer Verlag.

Nikoloulopoulos, A. K., Joe, H., & Li, H. (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis, 56*(11), 3659–3673.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 414-422.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics, 21*(2), 147-173.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics, 23*(3), 470-472.

Salvadori, G., & De Michele, C. (2007). On the use of copulas in hydrology: theory and practice. *Journal of Hydrologic Engineering, 12*, 369.

Schweizer, B., & Sklar, A. (1983). *Probabilistic metric spaces*: North Holland.

Schweizer, B., & Wolff, E. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics, 9*(4), 879-885.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris, 8*, 229–231.

Sklar, A. (1973). Random variables: joint distribution functions and copulas (Vol. 9): Kybernetika.

Spearman, C. (1904). The proof and measurement of association between two things. *The American journal of psychology, 15*(1), 72-101.

Wang, W., & Wells, M. T. (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association, 95*(449), 62-72.

Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing*: Academic Press.

Wilkinson, T. (1872). Treatise on Probability. *Nature, 7*(164), 124.

Wolff, E. F. (1980). N-dimensional measures of dependence. *Stochastica, 4*(3), 175-188.