UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Scienze Statistiche
SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE STATISTICHE
CICLO XXVI

COPULA-BASED MEASURES OF TAIL DEPENDENCE WITH APPLICATIONS

Direttore della Scuola: Ch.ma Prof.ssa Monica Chiogna

Supervisore: Ch.mo Prof. Nicola Torelli

Co-supervisore: Dott. Fabrizio Durante

Dottoranda: Roberta Pappadà

31 Gennaio 2014

To my dear Parents and my twin Brother

Acknowledgements

First of all, I would like to express my deep and sincere gratitude to my supervisors, Prof. Nicola Torelli and Dr. Fabrizio Durante, for their guidance and constant support. Throughout my Ph.D. they provided me with encouragements and invaluable advice. I feel fortunate to have had the opportunity to benefit from their knowledge and incredible enthusiasm for scientific research.

I am forever grateful to Dr. Gianfausto Salvadori, advisor of my Master Thesis at the Department of Mathematics (University of Salento, Italy), without whom I would never have undertaken this Ph.D. program. He introduced me to the world of Mathematical Statistics and encouraged me in pursuing further academic studies.

I wish to express my appreciation to the faculty, staff and colleagues of the Department of Statistical Sciences in Padova, for giving me the opportunity to spend three years in such a stimulating environment. I am thankful to Prof. Pace from the University of Udine, for valuable and critical comments which helped improve this work significantly.

I specially would like to thank my colleagues of the XXVI cycle of the Ph.D. School, Shireen, Ivan, Luca, Erlis, Lorenzo, Akram and Md Abud Darda. Their support and friendship accompanied me through joyful and difficult moments. They have been a major driving force during these years, providing me with good company and constructive discussions. I extend my thanks to the colleagues of the XXV cycle, for providing me with much encouragements and support.

I also would like to acknowledge the research team at the Center for Mathematics, Technische Universität München (Germany), and in particular Prof. Czado and Dr. Eike Brechmann, for the hospitality extended to me during my Ph.D. visiting period and precious collaboration. I am also grateful to my friend Silvia, who has made my stay in Munich a very enjoyable experience.

I owe my sincere gratitude to Elisa and Marta, that I met in Padova when everything was new for me. Their friendship which still lasts and unconditional support over the years will never be forgotten.

My heartfelt thanks go to my partner Lorenzo, for his unconditional love and understanding. He supported me and my work in such a way that I have never been alone to face my challenges. Thank you for the joy you bring to my life.

Finally, I wish to thank all members of my family for their love and strong support. My deepest gratitude goes to my parents, Sergio and Maria Rita, my greatest role models. I am very lucky to have my twin brother Alessandro, who has always been and remains to be my pillar. Thanks to them I became who I am today. To them I dedicate my thesis.

Roberta Pappadà

Padova
January 31, 2014

Abstract

With the advent of globalization and the recent financial turmoil, the interest for the analysis of dependencies between financial time series has significantly increased. Risk measures such as value-at-risk are heavily affected by the joint extreme comovements of associated risk factors. This thesis suggests some copula-based statistical tools which can be useful in order to have more insights into the nature of the association between random variables in the tail of their distributions.

Preliminarily, an overview of important definitions and properties in copula theory is given, and some known measures of tail dependence based on the notion of tail dependence coefficients and rank correlations are introduced. A first proposal consists of a graphical tool based on the so-called tail concentration function, in order to distinguish different families of copulas in a 2D configuration. This can be used as a copula selection tool in practical fitting problems, when one wants to choose one or more copulas to model the dependence structure in the data, highlighting the information contained in the tail.

The thesis mainly deals with financial time series applications, where copula functions and the related concepts of tail copula and tail dependence coefficients are used to characterize the dependence structure of asset returns.

Classical cluster analysis tools are revisited by introducing suitable copula-based tail dependence measures, which are exploited in the identification of similarities or dissimilarities between the variables of interest and, in particular, between financial time series. Such an approach is designed to investigate the joint behaviour of pairs of time series when they are taking on extremely low values. Either the asymptotic and the finite behaviour are assessed. The proposed methodology is based on a suitable copula-based time series model (GARCH-copula model), in order to model the marginal behaviour of each time series separately from the dependence pattern. Moreover, non-parametric estimation procedures are adopted for describing the pairwise dependencies, thus avoiding any model assumption. Simulation studies are conducted in order to check the performances of the proposed procedures and applications to financial data are presented showing their practical implementation.

The information coming from the output of the introduced clustering techniques can be exploited for automatic portfolio selection procedures in order to

hedge the risk of a portfolio, by taking into account the occurrence of joint losses. A two-stage portfolio diversification strategy is proposed and empirical analysis are provided.

Results show how the suggested approach to the clustering of financial time series can be used by an investor to have more insights into the relationships among different assets in crisis periods. Moreover, the application to portfolio selection framework suggests a cautious usage of standard procedures that may not work when the markets are expected to experience periods of high volatility.

Sommario

Con l'avvento della globalizzazione e la recente crisi finanziaria, l'interesse verso l'analisi delle relazioni tra serie storiche finanziarie è notevolmente aumentato. Misure di rischio come il value-at-risk sono fortemente influenzate dai movimenti estremi congiunti dei fattori di rischio associati.

Nella presente tesi si suggeriscono alcuni strumenti statistici basati sulla nozione di copula, che possono essere utili al fine di ottenere informazioni sulla natura dell'associazione tra variabili casuali nella coda delle loro distribuzioni.

Preliminarmente, vengono introdotte definizioni e proprietà fondamentali della teoria delle copule, e discusse alcune note misure di dipendenza basate sul concetto di coefficienti di dipendenza nella coda e correlazioni fra i ranghi. Una prima proposta consiste in uno strumento grafico basato sulla cosiddetta funzione di concentrazione di coda per distinguere tra diverse famiglie di copule in una configurazione bidimensionale. Questo strumento può essere impiegato in problemi pratici, quando si vuole scegliere tra una o più copule per modellizzare la struttura di dipendenza nei dati, evidenziando le informazioni contenute nella coda.

La tesi prende in considerazione diverse applicazioni nell'analisi di serie storiche finanziarie, in cui le funzioni copula e i relativi concetti di copule di coda e coefficienti di dipendenza nelle code vengono impiegati per caratterizzare la struttura di dipendenza dei rendimenti finanziari.

Gli strumenti standard per l'Analisi dei Gruppi (Cluster Analysis) vengono rivisitati attraverso l'introduzione di opportune misure di dipendenza, che permettano di identificare similarità o dissimilarità tra le quantità di interesse, nello specifico rappresentate da serie finanziarie. Tale approccio ha lo scopo di studiare il comportamento congiunto di coppie di serie finanziarie nel momento in cui esse assumono valori estremamente bassi. Vengono valutate sia la dipendenza asintotica che il comportamento finito. La metodologia proposta utilizza un modello per serie storiche basato sulle copule (GARCH-copula model), che consente di modellizzare il comportamento marginale di ogni serie temporale separatamente dalla struttura di dipendenza. Inoltre, vengono adottate procedure di stima non parametriche in relazione alla struttura di dipendenza, evitando così qualunque assunzione sul modello. Vengono condotti degli studi di simulazione per testare le procedure

proposte e diverse applicazioni a dati finanziari mostrano la loro implementazione pratica.

Il risultato delle tecniche introdotte precedentemente può essere utilizzato in procedure di selezione automatica di portafoglio al fine di coprire il rischio dovuto al verificarsi di perdite congiunte. Viene proposta una strategia di diversificazione di portafoglio in due fasi e illustrate le analisi empiriche.

L'approccio suggerito per il raggruppamento di serie finanziarie può essere utile ad un investitore per avere una visione più approfondita delle correlazioni tra mercati finanziari in periodi di crisi. Inoltre, l'applicazione nell'ambito della selezione di portafogli suggerisce un uso prudente delle procedure standard che potrebbero non essere appropriate quando si prevede che i mercati possano attraversare periodi di alta volatilità .

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

At the end of the nineties copulas became increasingly popular as a new powerful tool in the construction and estimation of multivariate stochastic models (Nelsen, 2006; Joe, 1997). First introduced by A. Sklar (Sklar, 1959), copulas have inspired a fast growth of interest and papers published in the subject due to many successful applications in different fields (*e.g.*, finance and economics, insurance, hydrology). Copulas have proved to be useful especially when the major issue is to understand/quantify a risk coming from different sources. See, for instance, Cherubini et al. (2011), Jaworski et al. (2013), Mai & Scherer (2012), Genest & Favre (2007) and the references therein. In these contexts, copula functions allow to aggregate individual risk factors (usually, expressed in terms of random variables) into one global risk output. Generally, such global risk is measured by the multivariate probability distribution function coupling the individual (one–dimensional) risks by means of a copula. In many practical applications, the global risk does not depend on the whole expression of the copula, but only on the behaviour of the copula in specific regions of its domain, the *tails* of a distribution. The concept of tail dependence describes the degree of dependence in the corner of the lower-left quadrant or upper-right quadrant of a bivariate distribution (Frahm et al., 2005). Tail-dependent distributions are of interest in many contexts and especially in financial applications, where it is clear that the assumption of normality can no longer be preserved. The concept of tail dependence can be embedded within the copula theory, such as the so–called tail-dependence coefficient. The flexibility of the copula-based approach in modelling dependencies avoiding constraints on the marginals represents one of the main advantages of copula models. As stressed, for instance, by the Basel Committee on Banking Supervision[1]: "The copula approach allows the practitioner to precisely specify the dependencies in the areas of the loss distributions that are crucial in determining the level of risk". Several investigations have been carried out during the years from different perspectives, ranging from tools from extreme-value analysis (Gudendorf & Segers, 2010) to the concept of threshold copulas (Jaworski, 2013).

In recent years financial markets have been characterized by an increasing globalization and a complex set of relationships among asset returns. Moreover, several studies have demonstrated that the linkages among different assets vary across time and that their strength tends to increase especially during crisis periods. In particular, practitioners are often interested in minimizing the whole risk of a portfolio of assets by adopting some diversification techniques which are based on the selection of different assets from markets and/or regions that one believes to be weakly correlated. In this framework, clustering techniques for multivariate financial time series are adopted in order to find sub-groups of asset returns such that elements within a group have a similar stochastic dependence structure, while elements from distinct groups have different behaviour. These procedures typically involve the choice of a convenient dissimilarity measure (see, *e.g.*, Piccolo, 1990; Caiado & Crato, 2010; Bastos & Caiado, 2013). Several clustering methods have focused on the use of Pearson correlation in order to infer the hierarchical structure of a portfolio of financial assets (see, for instance, the book by Kaufman & Rousseeuw, 1990). However, in most real life situations, returns of financial time series exhibit clear evidence against the multivariate normal distribution (McNeil et al., 2005). Therefore clustering procedures based on linear measures of correlation are inadequate for capturing extremal joint behaviour which is common in financial time series. Mainly due to this considerations, copula-based time series models have become a standard tool in modelling dependencies among univariate time series (see Patton, 2012).

[1]Developments in Modelling Risk Aggregation, October 2010,
`http://www.bis.org/publ/joint25.htm`

1.2 Summary and main contributions of this Thesis

This Thesis concerns the issue of tail dependence, as a relevant pattern in multivariate financial data. The identification of similarities/dissimilarities between the variables of interest can provide a better understanding of the underlying dependence structure. A copula-based approach allows us to construct models which go beyond the standard ones at the level of dependence. The proposed methodology finds application in several fields and is particularly useful in the analysis of financial data. In view of the results, the present thesis may represent one step in the direction of novel developments of copula-based statistical tools.

After Chapter 2, in which we introduce the concept of copula in full generality, the third chapter gives an overview of tail dependence measures revisited into the general framework of copulas. In particular, firstly the so–called *tail dependence coefficient* as a simple measure of dependence of large loss events is discussed. Secondly, we focus on the conditional version of Spearman's rank correlation, as an alternative measure of the amount of dependence in a given (tail) region of a bivariate distribution. Thirdly, we introduce the so–called *tail concentration function* (or quantile dependence function, see Patton, 2012), that may serve to visualize the tail behaviour of a copula at some (finite) points near the corners of its domain. We also discuss the potentials of such a function as a graphical tool useful to provide information in the choice of the copula model adequately fitting the data.

Chapter 4 introduces the reader to the framework of financial time series of returns, emphasizing some *stylized facts* regarding returns on financial assets: market data returns tend to be uncorrelated, but dependent, they are heavy tailed, extremes appear in clusters and volatility is random. In this framework, copula-based time series models provide a tool to handle in a flexible way the link among different univariate time series and allow us to model more accurately joint negative returns, without any influence of marginal behaviour on the dependence structure.

Dealing with financial data, an important issue concerns the minimization of the whole risk by adopting some diversification techniques, that should take into account the comovements among the assets. Therefore, the diversification issue naturally poses the question of investigating the relationship between financial time series and of checking whether they can be grouped together in such a way that may be helpful to portfolio selection. Following this ideas, Chapter 5 and Chapter

6 propose two strategies for clustering financial time series in "extreme scenarios", that means, aiming at grouping financial time series according to a measure that accounts for a kind of extreme (tail) dependence. Such a different approach consists of finding groups that are similar in the sense that assets within the same group tend to comove when they are experiencing very large losses. A similar viewpoint can be found in a recent work by De Luca & Zuccolotto (2011). However our approach presents some fundamental differences, since we avoid the specification of a fully parametric model for describing the pairwise dependence between the markets under consideration and we assume that the multivariate time series process follows a copula-based semi-parametric model. The two methodologies are intended to be used by an investor to have deeper insights into the relationships among different assets in crisis periods. The main role in the clustering procedures is played by dissimilarity measures, that are suitably defined to reflect the strength of the (positive) dependence between the time series in a given tail region of their joint distribution. To this end, two measures of tail dependence are adopted, which express different ways of looking at tail dependence since they focus, respectively, on asymptotic and finite tail behaviour. The estimation of these quantities is discussed and simulation studies are conducted to check the performances of the two procedures in identifying the clusters.

Different applications of the proposed methodologies to the analysis of specific financial datasets are presented. In addition, a further application concerns the construction of a weighted portfolio from a group of assets in such a way that the assets are diversified in their tail behaviour. This procedure is expected to be useful to have an idea about possible portfolios to be built in bearish periods and warn against the automatic use of standard portfolio selection procedures.

Chapter 2

Introduction to Copulas

2.1 Introduction

Many real-life situations can be modelled by a large number of random variables which play a significant role, and such variates are generally not independent. Therefore, it is often of fundamental importance to be able to link the marginal distributions of different variables in order to give a flexible and accurate description of the joint law of the variables of interest. Copulas were introduced in 1959 in the context of probabilistic metric spaces and later exploited as a tool for understanding relationships among multivariate outcomes. A copula is a function that links univariate marginals to their joint multivariate distribution in such a way that it captures the entire dependence structure in the multivariate distribution. The main advantage provided by a copula-approach in dependence modelling is that the selection of an appropriate model for the dependence between variables X and Y, represented by the copula, can proceed independently from the choice of the marginal distributions. The seminal result in the history of copulas is due to Sklar that introduced in 1959 the notion, and the name, of *copula*, and proved the theorem that now bears his name (Sklar, 1959). The latter states that any multivariate distribution can be expressed as its copula function evaluated at its marginal distribution functions. Moreover, any copula function when evaluated at any marginal distributions is a multivariate distribution.

The literature on the statistical properties and applications of copulas has been developing rapidly in recent years. For an introduction to the theory of copulas and a large selection of related models, the reader may refer, *e.g.*, to the monographs by Joe (1997) and Nelsen (2006), or to reviews such as Durante & Sempi (2010) and Cherubini et al. (2004), in which actuarial and financial applications are considered. Several other surveys of copula theory and applications in many fields have appeared in the literature to date, ranging from finance and economics to hydrology and environmental sciences. However, such a wide diffusion of applications of copulas has recently raised several criticisms among some researchers (see, for instance, the paper by Mikosch, 2006). It is worth stressing that copulas should not be regarded as the solution to all problems related to stochastic dependence and multivariate distributions. Copula models are just beginning to make their way into the statistical literature and further research efforts in investigating the potentials and the limitations of copula functions will be needed. The present thesis represents one small step in the direction of novel developments of copula-based statistical tools.

The rest of the chapter is so organized. Section 2.2 introduces the definition of a copula and presents some basic properties. Section 2.3 is dedicated to Sklar's Theorem and its interpretation, which allows us to consider a copula as a 'dependence function' and states the importance of copulas for stochastic models. Copulas invariance property and related dependence concepts are discussed in Section 2.4. Section 2.5 briefly reviews two well-established measures of correlation known as Spearman's rank correlation and Kendall's rank correlation in terms of copulas. Finally, some examples of frequently employed families of copulas are illustrated (Section 2.6).

2.2 Basic Definitions and Properties

To begin with, we need to establish basic notation. Let $d \in \mathbb{N}$ and $\boldsymbol{x}$ denote a vector $(x_1, \ldots, x_d) \in \mathbb{R}^d$. The symbol $\mathbb{I}$ will denote the unit interval $[0, 1]$. Usually, the abbreviation r.v. will denote a random vector, and the term *joint d.f.* will be used for denoting the distribution function (d.f.) of a random vector having, at least, two components. We start with the definition of a copula.

Definition 2.1. For every $d \geq 2$, a d-dimensional copula (d-copula) is a d-variate d.f. on $\mathbb{I}^d$ whose univariate marginals are uniformly distributed on $\mathbb{I}$.

In other words, a d-dimensional copula is a d-dimensional d.f. with all d univariate margins being $U(0,1)$. Since copulas are multivariate d.f.'s, the following characterization theorem holds.

Theorem 2.2.1. *A function $C : \mathbb{I}^d \to \mathbb{I}$ is a copula if, and only if, it satisfies the following properties:*

1. $C(1, \ldots, 1, u_j, 1, \ldots, 1) = u_j$ *for every* $1 \leq j \leq d$ *and all* $u_j \in \mathbb{I}$*;*
2. C *is increasing in each variable,* i.e. $C(\boldsymbol{u}) \leq C(\boldsymbol{v})$ *for all* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{I}^d$*,* $\boldsymbol{u} \leq \boldsymbol{v}$*;*
3. C *is d-increasing,* i.e. *the C-volume of any d-dimensional interval is non-negative.*

As an easy consequence, the joint probability of all outcomes is zero if the marginal probability of any outcome is zero, that is, $C(u_1, \ldots, u_d) = 0$ if $u_j = 0$ for any $1 \leq j \leq d$. Now, consider a continuous d-dimensional distribution function $F(x_1, \ldots, x_d)$ of the r.v. $\mathbf{X}$ with univariate marginal distributions $F_1, \ldots F_d$ and inverse (quantile) functions $F_1^{-1}, \ldots F_d^{-1}$. The transforms of uniform variates are distributed as $F_i, i = 1, \ldots, d$. Hence

$$\begin{aligned} F(x_1, \ldots, x_d) &= F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)) \\ &= \Pr(U_1 \leq u_1, \ldots, U_d \leq u_d) \\ &= C(u_1, \ldots, u_d). \end{aligned}$$

This anticipates the content of Sklar's theorem, which states the link between distribution functions and copulas and will be discussed in the next section.
Since copulas are multivariate distribution functions, *Fréchet-Hoeffding bounds* give upper and lower bounds in the class of all d-copulas.

Theorem 2.2.2. *Every copula satisfies the following inequality:*

$$\max\left[\sum_{i=1}^{d} u_i - d + 1, 0\right] \leq C(\boldsymbol{u}) \leq \min(u_1, \ldots, u_d) \tag{2.1}$$

for every $\boldsymbol{u} \in \mathbb{I}^d$.

The upper bound still satisfies the definition of copula. The lower bound is a copula for $d = 2$, while it never satisfies the definition of copula for $d \geq 3$. However, it can be proved that the bound is the best possible: pointwise there always exists a copula that takes its value. Fréchet-Hoeffding bounds are important in selecting an appropriate copula. Often, a desirable feature of a family of copulas is that it spans all possible degrees of dependence between the lower and the upper bound. The upper and lower Fréchet bound, respectively, can be considered special cases: $M_d(\boldsymbol{u}) = \min(u_1, \ldots, u_d)$ is the *comonotonicity copula*, associated with a vector $\mathbf{U} = (U_1, \ldots, U_d)$, whose components are uniformly distributed on $\mathbb{I}$ and such that $U_1 = U_2 = \cdots = U_d$ almost surely; $W_2(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$ is the *countermonotonicity copula*, associated with a vector $\mathbf{U} = (U_1, U_2)$ of r.v.'s uniformly distributed on $\mathbb{I}$ and such that $U_1 = 1 - U_2$ almost surely. Moreover, the *product copula* has the form $\Pi_d(\boldsymbol{u}) = u_1 \cdots u_d$, which corresponds to independence.

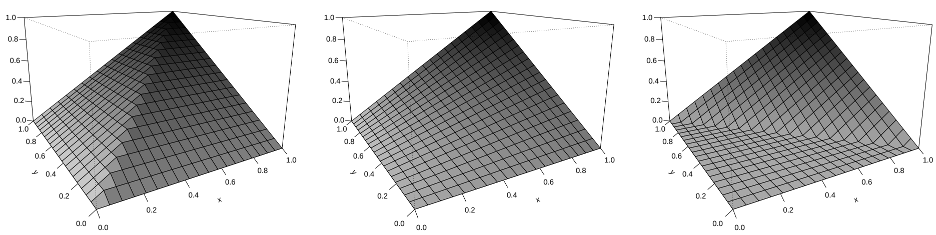


Figure 2.1: 3D plots of basic copulas. *Left* Comonotonicity copula (upper bound). *Middle* Product (or independence) copula. *Right* Countermonotonicity copula (lower bound).

2.3 Sklar's Theorem

Sklar's theorem is the building block of the theory of copulas. It guarantees that not only every joint continuous distribution function can be represented via a unique copula, but that the converse holds too. Sklar's theorem has been announced in Sklar (1959), however its first proof for the bivariate case appeared in Schweizer & Sklar (1974). Here, we enunciate the theorem in the d-dimensional case. A multivariate proof can be found in Schweizer & Sklar (1983) (compare also with Moore & Spruill, 1975; Deheuvels, 1978; Sklar, 1996).

Theorem 2.3.1. *Let $F_1, F_2, \dots, F_d$ be (given) marginal distribution functions and let A_j denote the range of F_j, $A_j := F_j(\overline{\mathbb{R}})$, for $j = 1, \dots, d$. Then, for every $(x_1, x_2, \dots, x_d) \in \overline{\mathbb{R}}^d$:*

(i) If F is a joint d.f. with univariate margins $F_1(x_1), F_2(x_2), \dots, F_d(x_d)$, there exists a unique copula C with domain $A_1 \times A_2 \cdots \times A_d$ such that

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)). \tag{2.2}$$

Hence, such a C is unique when $F_1, F_2, \dots, F_d$ are all continuous.

(ii) If C is any d-copula, then the function $F : \overline{\mathbb{R}}^d \to \mathbb{I}$ defined by (2.2) is a d-dimensional distribution function with margins $F_1, F_2, \dots, F_d$.

It follows as a corollary of Sklar's theorem that the copula that allows the representation (2.2) can be reconstructed from the margins and the joint distribution by inversion.

Corollary 2.3.2. *Under the hypothesis of part (i) of Sklar's theorem, if F_i is continuous for every $i \in \{1, \dots, d\}$, the copula $C : \overline{\mathbb{R}}^d \to \mathbb{I}$ is given by*

$$C(\boldsymbol{u}) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_d^{-1}(u_d)) \tag{2.3}$$

Corollary 2.3.2 states that the construction via Sklar's theorem exhausts the so–called *Fréchet class*, *i.e.* the class of joint distribution functions that have $F_1, F_2, \dots, F_d$ as margins.

By means of Sklar's result, while writing $F(\boldsymbol{x}) = C(F_1(x_1), \dots, F_d(x_d))$ we have the possibility to express the joint cumulative probability in terms of the marginal ones, referred to as the *basic probabilistic interpretation of copulas*; at the same time, we are able to separate marginal behaviour, as represented by the F_i, from the dependence between the marginals, only represented by the copula C of **X**. For this reason, a copula is often viewed as a dependence function.

Finally, consider a d-copula C. If C is absolutely continuous, then there exists *a.e.* in I^d a d-variate *density function* $c : \mathbb{I}^d \to [0, \infty)$ associated to C (see Durante et al., 2013a). The copula density reflects the strength of dependence of the margins. We have that

$$c(u_1, u_2, \dots, u_d) = \frac{\partial^d C(u_1, u_2, \dots, u_d)}{\partial u_1 \dots \partial u_d} \quad \text{a.e. on } \mathbb{I}^{\mathrm{d}}. \tag{2.4}$$

For absolutely continuous random variables with d.f. F, the copula density is related to the density of the distribution F, denoted as f, by the *canonical representation*

$$f(x_1, x_2, \dots, x_d) = c(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \cdot \prod_{j=1}^{d} f_j(x_j) \tag{2.5}$$

which implies that the copula density is equal to the ratio of the joint density f and the product of all marginal densities f_j.

2.4 Copulas as dependence functions

As pointed out in the previous section, copula functions allow researchers to consider marginal distributions and dependence as two separate but related issues. As a consequence, the copula function can be parametrized to include measures of dependence between the marginal distributions. Some properties of copulas deserve mention due to their important implications for empirical applications.

First, copulas have an attractive *invariance property* w.r.t. increasing transformations of the marginal distributions, that make copulas potentially very useful in applied work. The following theorem holds.

Theorem 2.4.1. (Schweizer & Wolff, 1976, 1981) *Let $X_1, \dots, X_d$ be continuous random variables with marginal distribution functions $F_1, \dots, F_d$ and copula C. Let $t_1, \dots, t_d$ be strictly increasing functions from $\mathbb{R}$ to $\mathbb{R}$. Then $t_1(X_1), \dots, t_d(X_d)$, which have marginal distribution functions $H_i = F_i \circ t_i^{-1}$, $i = 1, \dots, d$, and joint one $H(u_1, \dots, u_d) = \Pr(t_1(X_1) \leq u_1, \dots, t_d(X_d) \leq u_d)$, have copula C too:*

$$H(v_1, \dots, v_d) = C(H_1(v_1), \dots, H_d(v_d)). \tag{2.6}$$

The above result has important theoretical and applicative consequences. In fact, it implies that any 'property' of the joint distribution function of the random variables that is invariant under strictly increasing transformations of the random variables is a 'property' of their copula and independent of the individual distributions. In particular, all the concepts related to rank statistics (such as Kendall's τ) can be expressed in terms of copulas.

Some additional properties of copulas that can be inferred from Sklar's Theorem permit us to characterize independence, perfect positive and negative dependence, in terms of the basic copulas Π_d, M_d, W_2.

Proposition 2.4.2. *Let $(X_1, X_2, \ldots, X_d)$ be a r.v. with continuous joint d.f. F. Then*

(i) the copula of $(X_1, \ldots, X_d)$ is the product copula Π_d if, and only if, $X_1, \ldots, X_d$ are independent;

(ii) the copula of $(X_1, \ldots, X_d)$ is M_d if, and only if, there exists a r.v. Z and increasing functions $t_1, \ldots, t_d$ such that $\mathbf{X} = (t_1(Z), \ldots, t_d(Z))$ almost surely;

(iii) for $d = 2$, the copula of (X_1, X_2) is W_2 if, and only if, for some strictly decreasing function t, $X_2 = t\,(X_1)$ almost surely.

Random variates as in part (ii) of Proposition 2.4.2 are said to be comonotonic or *perfectly positively dependent*; while part (iii) corresponds to countermonotonicity or *perfectly negatively dependent* variables. That is, the association is positive if the copula attains the upper Fréchet-Hoeffding bound and negative if it attains the lower Fréchet–Hoeffding bound.

2.5 Measures of association

The general concept of association relates to random variates which are not independent according to the characterization in Section 2.4. That is, the random variables (X, Y) are said to be dependent or associated if they are not independent in the sense that $F(x, y) \neq F_1(x)F_2(y)$. In this section, we discuss the relationships between copula functions and some main association measures, in order to understand what is the nature of dependence that is captured by a copula. This issue is particularly relevant to the choice among different copula models. Well-known concepts of association include concordance, linear correlation, tail dependence, positive quadrant dependency. Some measures associated with them are rank correlations, the linear correlation coefficient, the indices of tail dependence. Two well-established measures of correlation known as Spearman's rank correlation (Spearman's *rho*) and Kendall's rank correlation (Kendall's *tau*) are reviewed in terms of copulas (see, *e.g.*, Joe, 1997; Embrechts et al., 2002). Both measures provide a valid alternative to the linear correlation coefficient which is often an inappropriate and misleading measure of dependence; they are invariant under monotonic transformations and do not depend on the functional forms of the marginal distributions. Tail dependence will be the topic of next chapter. We restrict the discussion to the bivariate case although generalization to higher dimensions is possible.

2.5.1 Kendall's *tau*

Consider two independent pairs of random variables $(X_1, Y_1), (X_2, Y_2)$ each with the same joint distribution function F. The vectors are said to be concordant if $X_1 > X_2$ whenever $Y_1 > Y_2$, and $X_1 < X_2$ whenever $Y_1 < Y_2$; and discordant in the opposite case. Thus, *concordance* refers to the property that large values of one random variable are associated with large values of another, whereas discordance refers to large values of one being associated with small values of the other.

Definition 2.2. (Kendall, 1938) Kendall's *tau* for the random vector (X_1, Y_1) is defined as

$$\tau = \Pr((X_1 - X_2)(Y_1 - Y_2) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_2) < 0). \qquad (2.7)$$

Interpreting Definition 2.2 in terms of concordance, one can easily obtain

$$\tau = \Pr[\text{concordance}] - \Pr[\text{discordance}]$$

Given a r.v. (X, Y) with continuous marginals having copula C, Kendall's τ can be expressed in terms of the copula as follows:

$$\tau_C = 4 \int\int_{\mathbb{I}^2} C(v, z) dC(v, z) - 1 \qquad (2.8)$$

Hence Kendall's tau is a copula property. Notice that $-1 \leq \tau \leq 1$, it is symmetric and assume the value zero under independence. Further,

$$\begin{array}{lll} \tau = -1 & \text{iff } C = W_2 & \text{iff } Y = t(X) \text{ a.e., } t \text{ decreasing;} \\ \tau = 1 & \text{iff } C = M_2 & \text{iff } Y = t(X) \text{ a.e., } t \text{ increasing.} \end{array}$$

It is also worth to stress that Kendall's coefficient is a normalized expected value. In fact, equation (2.8) can be rewritten by replacing the double integral with the expected value of the function $C(U_1, U_2)$, with U_1 and U_2 standard uniform and joint distribution C. Hence,

$$\tau = 4\mathrm{E}[C(U_1, U_2)] - 1$$

and

$$-1 \leq 4\mathrm{E}[C(U_1, U_2)] - 1 \leq 1.$$

A result by Nelsen (1991) establishes that the following lemma holds.

Lemma 2.5.1. *If C is a copula*

$$\int\int_{\mathbb{I}^2} C(v,z)dC(v,z) + \int\int_{\mathbb{I}^2} \frac{\partial C(v,z)}{\partial v}\frac{\partial C(v,z)}{\partial z}dvdz = \frac{1}{2}$$

It follows that when C has both an absolutely continuous and a singular component, or is singular, Kendall's τ can be computed as:

$$\tau = 1 - 4\int\int_{\mathbb{I}^2} \frac{\partial C(v,z)}{\partial v}\frac{\partial C(v,z)}{\partial z}dvdz \tag{2.9}$$

Finally, consider a random sample of n pairs (X_i, Y_i), $i = 1, \ldots, n$ and define $A_{ij} := \mathrm{sgn}(X_i - X_j)(Y_i - Y_j)$. It is straightforward that

$$\mathrm{E}(A_{ij}) = (+1)\Pr((X_i - X_j)(Y_i - Y_j) > 0) + (-1)\Pr((X_i - X_j)(Y_i - Y_j) < 0)$$

It follows that an unbiased estimator of Kendall's tau is the so–called *Sample Kendall's* τ, which is consistent as well:

$$\tau_S = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j>i} A_{ij}. \tag{2.10}$$

2.5.2 Spearman's *rho*

Consider three pairs of independent and identically distributed random vectors, namely (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3). As for Kendall's τ, Spearman's *rho* can be defined in terms of the probability of concordance and discordance as follows.

Definition 2.3. Spearman's *rho* for the random vector (X_1, Y_1) is defined as

$$\rho_S = 3[\Pr((X_1 - X_2)(Y_1 - Y_3) > 0) - \Pr((X_1 - X_2)(Y_1 - Y_3) < 0)]. \tag{2.11}$$

Note that X_2, Y_3 are independent. Therefore, given a r.v. (X, Y) with continuous marginals having copula C, Spearman's ρ for (X, Y) can be expressed in terms of the copula as follows (Schweizer & Wolff, 1981):

$$\rho_S = 12\int\int_{\mathbb{I}^2} vz\, dC(v,z) - 3 = 12\int\int_{\mathbb{I}^2} C(v,z)\, dvdz - 3.$$

One can easily derive

$$\rho_S = 12 \int\int_{\mathbb{I}^2} \{C(v,z) - uz\} dvdz. \qquad (2.12)$$

Now, if $X \sim F_1$ and $Y \sim F_2$, and we let $U_1 = F_1(X)$, $U_2 = F_2(Y)$, then

$$\begin{aligned}\rho_S &= 12\mathrm{E}(U_1U_2) - 3 \\ &= \frac{\mathrm{E}(U_1U_2) - 1/4}{1/12} = \frac{\mathrm{Cov}(U_1, U_2)}{\sqrt{\mathrm{Var}(U_1)}\sqrt{\mathrm{Var}(U_2)}}.\end{aligned}$$

It turns out that Spearman's coefficient is the linear correlation between $F_1(X)$ and $F_2(Y)$. Spearman's ρ_S is therefore the rank correlation, in the sense of correlation of the integral transforms of X and Y. Also for Spearman's ρ_S one could prove that it reaches its bounds if, and only if, X and Y are respectively countermonotonic and comonotonic continuous random variates:

$$\begin{aligned}&\rho_S = -1 \quad \text{iff } C = W_2 \quad \text{iff } Y = t(X) \text{ a.e., } t \text{ decreasing;} \\ &\rho_S = 1 \qquad \text{iff } C = M_2 \quad \text{iff } Y = t(X) \text{ a.e., } t \text{ increasing.}\end{aligned}$$

Again, consider a random sample of n pairs (X_i, Y_i), $i = 1, \ldots, n$, and define $R_i := \mathrm{rank(X_i)}$ and $S_i := \mathrm{rank(Y_i)}$. The sample version of ρ_S is an unbiased estimator of the population one and is computed as:

$$1 - 6\frac{\sum_{i=1}^{n}(R_i - S_i)^2}{n(n^2-1)}.$$

From the definitions of Kendall's tau and Spearman's rho it follows that both are concordance measures and depend only on the copula under consideration. Moreover, for continuous random variables all values in the interval $[-1, 1]$ can be obtained for Kendall's tau or Spearman's rho by a suitable choice of the underlying copula. Further details and more discussion on the relationships between these two measures can be found in Nelsen (2006).

2.6 Families of Copulas

Several investigations have been carried out concerning the construction of different families of copulas and their properties. We will focus on a few of them, those that seem to be more popular in the literature and frequently employed in financial applications: Gaussian copulas, t-copulas, and the class of Archimedean copulas.

2.6.1 Gaussian (Normal) copulas

The Gaussian (Normal) copula belongs to the so–called class of Elliptical copulas, the latter being simply the copulas of elliptical distributions (see, for instance, Fang et al., 1987).

The copula of the n-variate normal distribution with correlation matrix R has the form

$$C_R^{Ga}(\boldsymbol{u}) = \Phi_R\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_d)\right), \tag{2.13}$$

where Φ_R is the standardized multivariate normal distribution with correlation matrix R and Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution. For $d = 2$, the bivariate Gaussian copula is given by

$$\begin{aligned} C_\theta^{Ga}(u_1, u_2) &= \Phi_\theta\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{1/2}} \left(-\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)}\right) dsdt \end{aligned}$$

where $\theta \in [-1, 1]$, and Φ^{-1} denotes the inverse of the univariate Gaussian distribution. It is easy to verify that the Gaussian copula generates the standard Gaussian joint distribution function - via Sklar's Theorem - whenever the margins are standard normal. The expression of the density of the normal copula c_R^{Ga} can be derived from (2.13), via the canonical representation:

$$c_R^{Ga}\left(\Phi(x_1), \ldots, \Phi(x_d)\right) = \frac{\dfrac{1}{(2\pi)^{\frac{d}{2}} |R|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\boldsymbol{x}^T R^{-1} \boldsymbol{x}\right)}{\prod_{j=1}^{d} \left(\dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{1}{2} x_j^2\right)\right)},$$

where $|R|$ is the determinant of R (see, *e.g.*, Cherubini et al., 2004). Finally, it is worth stressing that there are several bivariate random variables having Gaussian margins but do not possess a Gaussian dependence structure (a Gaussian copula).

2.6.2 *t*-copulas

Let $t_{R,\nu}$ denote the standardized multivariate Student's t distribution with correlation matrix R and ν degrees of freedom. The multivariate Student's t-copula is defined as follows:

$$C_{R,\nu}^t(\boldsymbol{u}) = t_{R,\nu}(t_\nu^{-1}(u_1,), t_\nu^{-1}(u_2,), \ldots, t_\nu^{-1}(u_d)), \tag{2.14}$$

where t_ν^{-1} is the inverse of the univariate c.d.f. of Student's t with ν degrees of freedom. In the bivariate case the copula expression can be written as

$$\begin{aligned} C^t_{\rho,\nu}(u_1,u_2) &= t_{\rho,\nu}\left(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)\right) \\ &= \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left(1+\frac{s^2-2\rho st+t^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} dsdt \end{aligned}$$

where ρ is simply the linear correlation coefficient of the corresponding bivariate t_ν-distribution if $\nu > 2$.

Student's t-copulas, as Gaussian copulas, are easily parametrized by the linear correlation matrix, but only the former yield dependence structures with tail dependence. Thus, t-copulas represent a valid alternative to Gaussian copulas, especially in financial applications where the need to get the extreme joint tail observations, clearly present in the real data, often arises. Finally, expressing the copula density as the ratio of the joint density and the product of all marginal densities, one has:

$$c_{R,\nu}(u_1,\dots,u_d) = |R|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+d}{2}\right)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^{n-1}}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^{n}} \frac{\left(1+\frac{1}{\nu}\zeta^T R^{-1}\zeta\right)^{-\frac{\nu+d}{2}}}{\prod_{j=1}^{d}\left(1+\frac{\zeta_j^2}{\nu}\right)^{-\frac{\nu+1}{2}}},$$

where $\zeta_j = t_\nu^{-1}(u_j)$ (see, *e.g.*, Cherubini et al., 2004).

2.6.3 Archimedean copulas

In this section we discuss an important class of copulas called Archimedean copulas. They have proved to be useful in several applications since they are capable of capturing wide ranges of dependence structures. Furthermore, in contrast to elliptical copulas, all commonly encountered Archimedean copulas have closed form expressions. There is a vast literature about Archimedean copulas. See, for instance, McNeil & Nešlehová (2009).

Archimedean copulas may be constructed using a function $\varphi : \mathbb{I} \to [0,\infty]$, continuous, decreasing, convex and such that $\varphi(1) = 0$. Such a function φ is called a *generator*. It is called a strict generator whenever $\varphi(0) = +\infty$. The pseudo-inverse of φ is defined as follows:

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

The pseudo-inverse is such that, by composition with the generator, it gives the identity, and it coincides with the usual inverse if φ is a strict generator.

Definition 2.4. Given a generator and its pseudo-inverse, an Archimedean 2-copula takes the form

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \tag{2.15}$$

If the generator is strict, the copula is said to be a strict Archimedean 2-copula.

Table 2.1 lists three bivariate one-parameter Archimedean copulas that appear regularly in statistical literature: Gumbel, Clayton, Frank (see Figure 2.2). They are constructed using a generator $\varphi_\alpha(t)$, indexed by the (real) parameter α. These three copulas accommodate different patterns of dependence and have relatively straightforward functional forms. Archimedean 2-copulas are easily verified to be symmetric (*i.e.* $C(u, v) = C(v, u)$, for every $(u, v) \in \mathbb{I}^2$), and associative (*i.e.* $C(C(u, v), z) = C(u, C(v, z))$, for every $(u, v, z) \in \mathbb{I}^3$). Moreover, they are easily related to measures of association and tail dependency (see Chapter 3).

Table 2.1: Selected Archimedean 2-copulas and their generators.

	$C(u,v)$	$\varphi_\alpha(t)$	range for α
Gumbel	$\exp\{-[(-\log u)^\alpha + (-\log v)^\alpha]\}$	$(-\log t)^\alpha$	$[1, +\infty)$
Clayton	$(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}$	$\alpha^{-1}(t^{-\alpha} - 1)$	$(0, +\infty)$
Clayton*	$\max[(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}, 0]$	$\alpha^{-1}(t^{-\alpha} - 1)$	$(-1, +\infty)\setminus\{0\}$
Frank	$-\frac{1}{\alpha}\log\left(1 + \frac{(\exp(-\alpha u)-1)(\exp(-\alpha v)-1)}{\exp(-\alpha)-1}\right)$	$-\log\frac{\exp(-\alpha t)-1}{\exp(-\alpha)-1}$	$(-\infty, \infty)\setminus\{0\}$

*For Clayton, the two cases correspond to strict and nonstrict generator, respectively.

Multivariate extensions can be obtained if restrictions are placed on the generator (see, *e.g.*, Durante & Sempi, 2010; Cherubini et al., 2004).

Definition 2.5. A function $h(t) : \mathbb{R} \to \mathbb{R}$ is said to be completely monotone on the interval J if it belongs to C^∞ and it has derivatives of all orders which alternate in sign, *i.e.* if it satisfies

$$(-1)^n \frac{d^n h(t)}{dt^n} \geq 0, \quad n = 0, 1, 2, ...$$

The following theorem from Kimberling (1974) is useful.

Theorem 2.6.1. *Let φ be a generator. The function $C : \mathbb{I}^d \to \mathbb{I}$ defined by*

$$C(u_1, u_2, \ldots, u_d) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_d))$$

is a copula for all $d \geq 2$ if, and only if, φ^{-1} is completely monotonic on $[0, \infty]$.

Finally, we give the definition of Archimedean d-copula.

Definition 2.6. Let φ be a strict generator, with φ^{-1} completely monotonic on $[0, \infty]$. Then a d-dimensional copula C is called Archimedean if it admits the representation

$$C(\boldsymbol{u}) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \cdots + \varphi(u_d)) \tag{2.16}$$

for all $\boldsymbol{u} \in \mathbb{I}$.

In the following, we briefly review the multivariate Archimedean copulas belonging to Gumbel, Clayton and Frank family.

Gumbel d-copula

The Gumbel family has been introduced by Gumbel (1960). Since it has been discussed in Hougaard (1986), it is also known as the Gumbel–Hougaard family. The standard expression for members of this family of d-copulas is

$$C(u_1, \ldots, u_d) = \exp\left(-\left(\sum_{i=1}^{d}(-\log(u_i))^{\alpha}\right)^{\frac{1}{\alpha}}\right), \ \alpha \geq 1. \tag{2.17}$$

The case $\alpha = 1$ gives the product copula as a special case, and the limit of (2.17) for $\alpha \to +\infty$ is the comonotonicity copula. It follows that the Gumbel family can represent independence and positive dependence only. The generator is given by $\varphi_\alpha(u) = (-\log u)^\alpha$, $\alpha \geq 1$.

Clayton d-copula

The Clayton family was first proposed by Clayton (1978), and studied by Oakes (1982). The standard expression for members of this family of d-copulas is

$$C(u_1, \ldots, u_d) = \left(\sum_{i=1}^{d} u_i^{-\alpha} - (d-1)\right)^{-\frac{1}{\alpha}} \quad \alpha > 0. \tag{2.18}$$

The limiting case $\alpha = 0$ corresponds to the independence copula. The generator has the form $\varphi_\alpha(u) = u^{-\alpha} - 1$, $\alpha > 0$.

Frank d-copula

Copulas of this family have been introduced by Frank (1979), and have the expression:

$$C(u_1, \ldots, u_d) = -\frac{1}{\alpha} \log \left\{ 1 + \frac{\prod_{i=1}^{d} (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{d-1}} \right\}, \; \alpha > 0. \tag{2.19}$$

It reduces to the product copula if $\alpha = 0$. For the case $d = 2$, the parameter α can be extended also to the case $\alpha < 0$. The generator is given by

$$\varphi_\alpha(u) = -\log \left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1} \right), \; \alpha > 0.$$

In Figure 2.2 we present an example of 1000 simulated draws from a Gumbel copula with $\alpha = 3$, a Clayton copula with $\alpha = 4$ and a Frank copula with $\alpha = 6$, respectively, and $N(0, 1)$ margins. We also show the corresponding contour plots of the copula density.

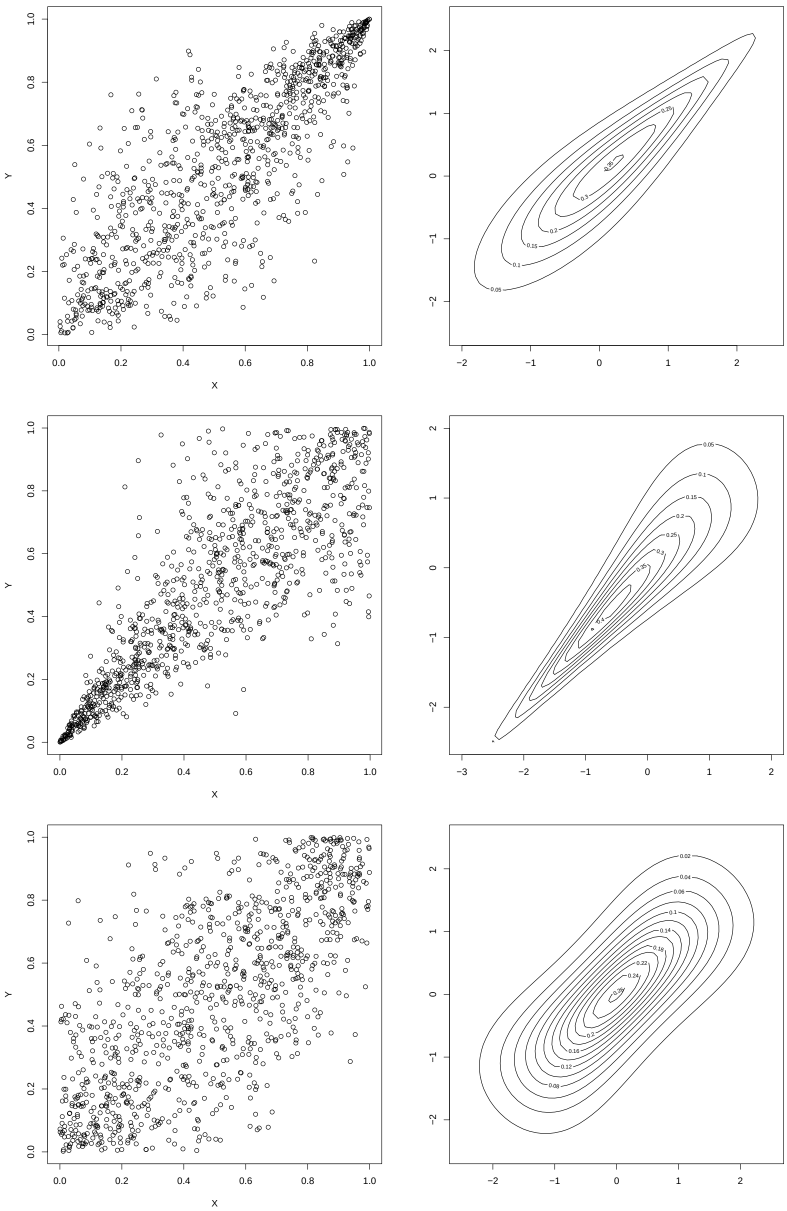


Figure 2.2: *Top* Bivariate sample of size $n = 1000$ from the Gumbel copula with parameter $\alpha = 3,0$ ($C^G_{3,0}$), and contour plot of the density of the d.f. $F = C(F_1, F_2)$, with $F_1, F_2 \sim N(0,1)$, $C = C^G_{3,0}$. *Middle* Bivariate sample of size $n = 1000$ from the Clayton copula with parameter $\alpha = 4,0$ ($C^{Cl}_{4,0}$), and contour plot of the density of the d.f. $F = C(F_1, F_2)$, with $F_1, F_2 \sim N(0,1)$, $C = C^{Cl}_{4,0}$. *Bottom* Bivariate sample of size $n = 1000$ from the Frank copula with parameter $\alpha = 6,0$ ($C^{Fr}_{6,0}$), and contour plot of the density of the d.f. $F = C(F_1, F_2)$, with $F_1, F_2 \sim N(0,1)$, $C = C^{Fr}_{6,0}$.

Chapter 3

Tail Dependence measures

3.1 Introduction

A primary objective in modern risk management is to represent the comovement of markets as closely as possible, dealing with non-normality at the univariate and multivariate level. Corresponding to the heavy tail property in univariate distributions, *tail dependence* is used to model the co-occurrence of extreme events. Tail dependence refers to the degree of dependence in the corner of the lower-left quadrant or upper-right quadrant of a bivariate distribution. The most popular measure of tail behaviour is the so–called *tail dependence coefficient* (TDC), introduced by Sibuya (1959) and discussed by Joe (1993). In their approach, TDC corresponds to the probability that one margin exceeds a high/low threshold under the condition that the other margin exceeds a high/low threshold. In financial applications, the interest is usually concentrated on the probability that two stock indexes fall below given levels. When such a probability is invariant under strictly increasing transformations (for instance, it is based on ranks or quantiles), then it can be expressed only in terms of copulas. TDC's provide an asymptotic measure of tail dependence and represent one of many possible approaches. Conditional versions of common dependence measures have been considered in order to investigate the amount of dependence in a given region of a bivariate distribution. For example, a conditional version of Pearson's correlation coefficient ρ of a bivariate random vector $\mathbf{X} = (X_1, X_2)$ is defined by $\rho_A := \rho(X_1, X_2 | \mathbf{X} \in A)$ for some (measurable) set $A \subset \mathbb{R}^2$. However, Pearson's correlation pitfalls have been pointed out by a large number of authors (see, for instance, Embrechts et al., 2002). Possible alternatives are conditional versions of Spearman's rho and Kendall's tau, which only depend on the underlying copula (see Section 2.5).

This chapter defines and explores the concept of tail dependence from different perspectives. First, tail dependence is characterized by the TDC into the general framework of copulas, and computed for several families (Section 3.2). As stressed before, measures of finite tail dependence can be derived from conditional versions of rank correlations. In Section 3.3 we focus on conditional Spearman's rho, which will play a central role in the clustering procedure described in Chapter 5. Section 3.4 introduces an auxiliary function that may serve to visualize the tail behaviour of a copula C, the so–called *tail concentration function* (or quantile dependence function, see Patton, 2012). Finally, in Section 3.5 a variation of the graphical tool in Michiels & De Schepper (2013) is proposed, in order to detect which families of copulas are closer to the empirical copula in the tail dependence behaviour, as described by their tail concentration functions. This can be used as a copula selection tool in practical fitting problems, when one wants to choose one or more copulas to model the dependence structure in the data.

3.2 Tail dependence coefficients

We start with the most common definition of *upper* and *lower* tail dependence coefficients, for two continuous r.v.'s X_1 and X_2, with d.f.'s F_1 and F_2, respectively. Let F_1^{-1}, F_2^{-1} denote the quantile functions of X_1 and X_2, respectively.

Definition 3.1. The upper tail dependence coefficient λ_U (upper TDC) of (X_1, X_2) is defined by

$$\lambda_U := \lim_{t \to 1^-} \Pr\left(X_2 > F_2^{-1}(t) | X_1 > F_1^{-1}(t)\right) \tag{3.1}$$

provided that the limit $\lambda_U \in [0,1]$ exists. The lower tail dependence coefficient λ_L (lower TDC) of (X_1, X_2) is defined by

$$\lambda_L := \lim_{t \to 0^+} \Pr\left(X_2 \leq F_2^{-1}(t) | X_1 \leq F_1^{-1}(t)\right) \tag{3.2}$$

provided that the limit $\lambda_L \in [0,1]$ exists. If $\lambda_U \in (0,1]$, X_1 and X_2 are said to be *asymptotically dependent* in the upper tail; if $\lambda_U = 0$, X_1 and X_2 are said to be *asymptotically independent* in the upper tail. Analogous definitions hold for λ_L.

The following equivalent representation shows that tail dependence is a copula property (Joe, 1997).

Definition 3.2. If a 2-copula C is such that

$$\lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1-t} = \lambda_U \tag{3.3}$$

exists, then C has upper tail dependence if $\lambda_U \in (0, 1]$, and upper tail independence if $\lambda_U = 0$. Similarly, if the limit

$$\lim_{t \to 0^+} \frac{C(t,t)}{t} = \lambda_L \tag{3.4}$$

exists, then C has lower tail dependence if $\lambda_L \in (0, 1]$, and lower tail independence if $\lambda_L = 0$.

Since tail dependence is a copula property, it follows that many copula features transfer to the TDC's, such as the invariance under strictly increasing transformations of the margins. Given two random variables with copula C, the *survival copula* is so defined:

$$\widehat{C}(u,v) = u + v - 1 + C(1-u, 1-v), \tag{3.5}$$

and the joint survival function for two uniformly distributed random variables whose joint distribution function is C is given by

$$\overline{C}(u,v) = 1 - u - v + C(u,v) = \widehat{C}(1-u, 1-v).$$

Hence the following relationships can be easily derived:

$$\lim_{t \to 0^+} \frac{\widehat{C}(t,t)}{t} = \lim_{t \to 1^-} \frac{\widehat{C}(1-t, 1-t)}{1-t} = \lim_{t \to 1^-} \frac{\overline{C}(t,t)}{1-t}.$$

Thus, the lower TDC of $\widehat{C}$ is the upper TDC of C, *i.e.* $\lambda_L(\widehat{C}) = \lambda_U(C)$, where $\lambda_U(\cdot)$, $\lambda_L(\cdot)$ are defined as in (3.3) and (3.4), respectively. In a similar way, one can prove that $\lambda_U(\widehat{C}) = \lambda_L(C)$.

For copulas with simple analytical expressions, the computation of λ_U and λ_L can be straightforward. For copulas without a simple closed form an alternative formula for λ_U (λ_L) can be used. Consider a pair of random variables (U_1, U_2)

uniformly distributed on $\mathbb{I}$ with absolutely continuous copula C. Then

$$\begin{aligned}
\lambda_U &= \lim_{u \to 1^-} \frac{\overline{C}(u,u)}{1-u} \\
&= - \lim_{u \to 1^-} \frac{d\overline{C}(u,u)}{du} \\
&= \lim_{u \to 1^-} \left(2 - \frac{\partial}{\partial s} C(s,t) \Big|_{s=t=u} - \frac{\partial}{\partial t} C(s,t) \Big|_{s=t=u} \right) \\
&= \lim_{u \to 1^-} \left(\Pr(U_2 > u | U_1 = u) + \Pr(U_1 > u | U_2 = u) \right),
\end{aligned}$$

since $\Pr(V > v | U = u) = 1 - \partial C(u,v)/\partial u$. Moreover, if C is an *exchangeable copula*, *i.e.* $C(u,v) = C(v,u)$, then the latter expression simplifies to

$$\lambda_U = 2 \lim_{u \to 1^-} \Pr(U_2 > u | U_1 = u). \tag{3.6}$$

Analogously, the lower TDC can be expressed as

$$\lambda_L = \lim_{u \to 0^+} \left(\Pr(U_2 < u | U_1 = u) + \Pr(U_1 < u | U_2 = u) \right)$$

and simplifies to

$$\lambda_L = 2 \lim_{u \to 0^+} \Pr(U_2 < u | U_1 = u), \tag{3.7}$$

when C is exchangeable.

Example 3.2.1. Let C be a member of the Gaussian family with expression given in Section 2.6.1, $(X_1, X_2)^T \sim C(\Phi(x_1), \Phi(x_2))$, with linear correlation coefficient ρ. Then, formula (3.6) can be used to compute the upper TDC, which coincides with the lower TDC. We have

$$\begin{aligned}
\lim_{u \to 1^-} \Pr(U_2 > u | U_1 = u) &= \lim_{x \to \infty} \Pr(\Phi^{-1}(U_2) > x | \Phi^{-1}(U_1) = x) \\
&= \lim_{x \to \infty} \Pr(X_2 > x | X_1 = x).
\end{aligned}$$

We know that $X_2 | X_1 = x \sim N(\rho x, 1 - \rho^2)$. Thus,we obtain

$$\lambda_U = 2 \lim_{x \to \infty} \overline{\Phi}((x - \rho x)/\sqrt{1-\rho^2}) = 2 \lim_{x \to \infty} \overline{\Phi}(x\sqrt{1-\rho}/\sqrt{1+\rho}),$$

with Φ denoting the distribution function of the standard normal distribution. We showed that Gaussian copulas do have zero upper tail dependence coefficient. Since elliptical distributions are radially symmetric, the coefficient of upper and lower tail dependence coincide, that is, the Gaussian copula has either zero upper and lower tail dependence coefficient for $\rho < 1$.

Example 3.2.2. Let C be a member of the Student family with expression given in Section 2.6.2, $(X_1, X_2)^T \sim C(t_{\rho,\nu}(x_1), t_{\rho,\nu}(x_2))$, with ν degrees of freedom and linear correlation coefficient ρ. Then $X_2|X_1 = x \sim t_{\nu+1}\left(\rho x, \frac{\nu+x^2}{\nu+1}(1-\rho^2)\right)$. It follows that a t-copula's tail dependence can be evaluated by

$$
\begin{aligned}
\lambda_U &= 2 \lim_{x\to\infty} \Pr(X_2 > x | X_1 = x) \\
&= 2 \lim_{x\to\infty} \bar{t}_{\nu+1}\left(\left(\frac{\nu + x^2}{\nu+1} \right)^{-\frac{1}{2}} \frac{x(1-\rho)}{\sqrt{1-\rho^2}} \right) \\
&= 2 \lim_{x\to\infty} \bar{t}_{\nu+1}\left(\left(\frac{\nu/x^2 + 1}{\nu+1} \right)^{-\frac{1}{2}} \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right) \\
&= 2 - 2t_{\nu+1}\left(\sqrt{\nu+1}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right),
\end{aligned}
$$

provided that $\rho > -1$. Note that even for zero correlation this copula shows tail dependence. As shown, t-copulas have upper and lower tail dependence which tend to zero as the number of degrees of freedom tends to infinity for $-1 < \rho < 1$. One can see the difference of the tail dependence between Gaussian copulas and t-copulas from Figure 3.1.

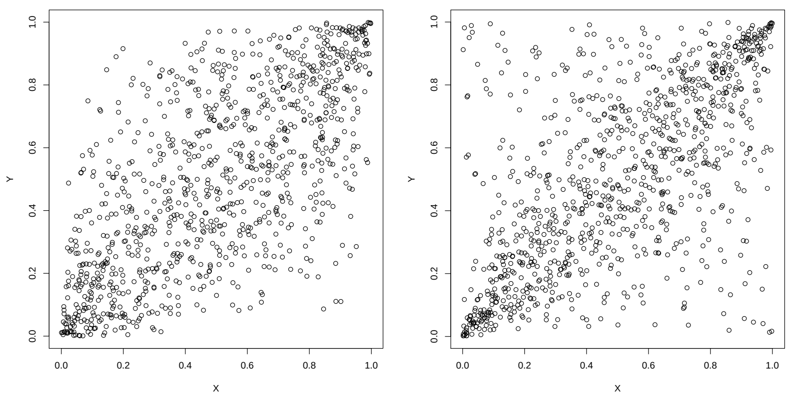


Figure 3.1: *Left* 1000 samples from a Gaussian copula with Kendall's $\tau = 0.5$. *Right* 1000 samples from a t-copula with Kendall's $\tau = 0.5$ and $\nu = 2$.

For Archimedean copulas, tail dependence can be expressed in terms of the generators. First, we recall the definition of Laplace transform.

Definition 3.3. Let X be a nonnegative random variable with distribution function

F. Consider the Laplace transform of X

$$\phi_X(s) = \int_0^\infty e^{-sx} dF(x),\ s > 0. \tag{3.8}$$

It's easy to prove that the inverse of Laplace transforms gives strict generators. The following result is demonstrated in Joe (1997).

Theorem 3.2.1. *Let φ be a strict generator such that φ^{-1} belongs to the class of Laplace transforms of strictly positive random variables. If ${\varphi^{-1}}'(0)$ is finite, then*

$$C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v))$$

does not have upper tail dependence. If C has upper tail dependence, then ${\varphi^{-1}}'(0) = -\infty$ and the upper tail dependence coefficient is given by

$$\lambda_U = 2 - 2 \lim_{s \to 0^+} \frac{{\varphi^{-1}}'(2s)}{{\varphi^{-1}}'(s)}. \tag{3.9}$$

The lower tail dependence coefficient is equal to

$$\lambda_L = 2 \lim_{s \to \infty} \frac{{\varphi^{-1}}'(2s)}{{\varphi^{-1}}'(s)}. \tag{3.10}$$

Applying Theorem 3.2.1 it is possible to compute tail dependency for the majority of the commonly encountered Archimedean copulas, as shown in the following examples.

Example 3.2.3. Gumbel copulas have strict generator $\varphi_\alpha(t) = (-\mathrm{log}t)^\alpha$, see Table 2.1. Thus $\varphi_\alpha^{-1}(s) = \exp(-s^{1/\alpha})$ and ${\varphi_\alpha^{-1}}'(s) = (-1/\alpha)s^{(1/\alpha)-1}\exp(-s^{1/\alpha})$. From Theorem 3.2.1 we obtain

$$\lambda_U = 2 - 2^{1/\alpha} \lim_{s \to 0^+} \frac{\exp(-(2s)^{1/\alpha})}{\exp(-s^{1/\alpha})} = 2 - 2^{1/\alpha}.$$

Example 3.2.4. The members of the strict Clayton family have generator $\varphi_\alpha(t) = \alpha^{-1}(t^{-\alpha} - 1)$ for $\alpha > 0$. Thus $\varphi_\alpha^{-1}(s) = (1 + \alpha s)^{-1/\alpha}$.
It follows that $\lambda_U = 0$ and

$$\lambda_L = 2 \lim_{s \to \infty} \frac{(1 + 2\alpha s)^{(-1/\alpha)-1}}{(1 + \alpha s)^{(-1/\alpha)-1}} = 2^{-1/\alpha}.$$

Example 3.2.5. Consider the Frank family in Table 2.1. The strict generator is $\varphi_\alpha(t) = -\log\frac{\exp(-\alpha t)-1}{\exp(-\alpha)-1}$, for $\alpha \in (-\infty,\infty)\backslash\{0\}$. Hence,

$$\varphi_\alpha^{-1}(s) = -\frac{1}{\alpha}\log(1-(1-e^{-\alpha})e^{-s}),$$
$$\varphi_\alpha^{-1\prime}(s) = -\frac{1}{\alpha}\frac{(1-e^{-\alpha})e^{-s}}{(1-(1-e^{-\alpha})e^{-s})}.$$

It follows that $\varphi_\alpha^{-1\prime}(0) = (-1/\alpha)(e^{\alpha}-1)$, which is finite. Therefore, the Frank family does not have upper tail dependence according to Theorem 3.2.1. Due to the radially symmetric property of this family, Frank copulas do not have lower tail dependence.

Table 3.1: Tail dependence coefficients for popular families of copulas.

Copula	λ_L	λ_U
Gaussian	0	0
Elliptical	≥ 0	≥ 0
Gumbel	0	$2-2^{1/\alpha}$
Clayton ($\alpha \geq 0$)	$2^{-1/\alpha}$	0
Frank	0	0
Plackett	0	0

3.3 Conditional Spearman's *rho*

According to the definition given in Chapter 2, the Spearman's rank-correlation coefficient of a r.v. (X,Y) with joint d.f F and continuous marginals F_X and F_Y, respectively, having copula C is given by

$$\rho_S = 12\int\int_{\mathbb{I}^2} vz\,dC(v,z) - 3 = 12\int\int_{\mathbb{I}^2} C(v,z)\,dvdz - 3. \tag{3.11}$$

For every fixed α with $0<\alpha<1$, define the set

$$A_T := \{\ (x,y) \mid x \leq F_X^{-1}(\alpha), y \leq F_Y^{-1}(\alpha)\}\ ,$$

where $F_X^{-1}(\alpha) = \inf\{x \mid F_X(x) \geq \alpha\}$ and $F_Y^{-1}(\alpha) = \inf\{y \mid F_Y(y) \geq \alpha\}$ denote the quantile functions with respect to F_X and F_Y, respectively, for $0<$

$\alpha < 1$. Thus $\Pr((X,Y) \in A_T) = C(\alpha, \alpha)$ and we assume that $C(\alpha, \alpha) > 0$. The *conditional joint c.d.f.* can be written as

$$\begin{aligned} F_T(x,y) =& \Pr(X \leq x, Y \leq y \mid (X,Y) \in A_T) = \frac{F\left(x \wedge F_X^{-1}(\alpha), y \wedge F_Y^{-1}(\alpha)\right)}{F\left(F_X^{-1}(\alpha), F_Y^{-1}(\alpha)\right)} \\ =& \frac{C\left(F_X\left(x \wedge F_X^{-1}(\alpha)\right), F_Y\left(y \wedge F_Y^{-1}(\alpha)\right)\right)}{C(\alpha,\alpha)}, \quad \forall x, y \in \mathbb{R}. \end{aligned}$$

The corresponding conditional marginal distribution functions are given by

$$\begin{aligned} F_{X_T}(x,y) =& \Pr(X \leq x \mid (X,Y) \in A_T) = F_T\left(x, F_Y^{-1}(\alpha)\right) \\ =& \frac{C\left(F_X\left(x \wedge F_X^{-1}(\alpha)\right), \alpha\right)}{C(\alpha,\alpha)}, \quad \forall x \in \mathbb{R}. \end{aligned}$$

and

$$\begin{aligned} F_{Y_T}(x,y) =& \Pr(Y \leq y \mid (X,Y) \in A_T) = F_T\left(F_X^{-1}(\alpha), y\right) \\ =& \frac{C\left(t, F_Y\left(y \wedge F_Y^{-1}(\alpha)\right)\right)}{C(\alpha,\alpha)}, \quad \forall y \in \mathbb{R}, \end{aligned}$$

with F_{X_T} and F_{Y_T} continuous. Due to Sklar's Theorem there exist a unique copula $C_T : [0,1]^2 \to [0,1]$ such that

$$F_T(x,y) = C_T(F_{X_T}(x), F_{Y_T}(y)), \ \forall x, y \in \mathbb{R}. \tag{3.12}$$

We refer to the copula $C_T(u,v) = F_T(F_{X_T}^{-1}(u), F_{Y_T}^{-1}(v))$, $u, v \in [0,1]$, as the *threshold (lower tail) copula* associated with the conditional joint c.d.f. (see Dobrić et al., 2013).

From equation (3.11) we can define the conditional Spearman's correlation coefficient by using the threshold tail copula, *i.e.*,

$$\rho_S(C_T) = 12 \int\int_{\mathbb{I}^2} vz\, dC_T(v,z) - 3 = 12 \int\int_{\mathbb{I}^2} C_T(v,z)\, dvdz - 3. \tag{3.13}$$

Now, consider the samples X_t, Y_t, $(t = 1, \ldots, n)$, with copula C. The procedure to calculate the conditional Spearman's coefficient $\widehat{\rho}_S$ associated to the sample observations runs as in Algorithm 3.1. For more details about its practical implementation, see Dobrić et al. (2013), Durante & Jaworski (2010). Notice that, by using similar arguments as in Schmid & Schmidt (2007), it can be proved that the conditional version of Spearman's rho described in Algorithm 3.1. are consistent and asymptotically normally distributed; *i.e.*,

$$\sqrt{n_T}(\widehat{\rho}_S(C_T) - \rho_S(C_T)) \xrightarrow{d} N(0, \sigma_T^2),$$

as n_T tends to ∞, provided that the threshold tail copula exists and satisfies some regularity assumptions. Here σ_T^2 depends on the threshold copula C_T. Furthermore, the calculation of Spearman's correlation depends on the number of points n_T in tail region. If such number is small, the estimated correlations would be affected by this small sample size (Dobrić et al., 2013). Typically, a convenient sample size may be reached by selecting an appropriate threshold α.

Algorithm 3.1 Calculation of $\rho_S(C_T)$.

1. Set the threshold $\alpha \in (0, 0.5)$.
2. Calculate the empirical cumulative distribution functions $\widehat{F}_X$ and $\widehat{F}_Y$ associated with X_t and Y_t, $(t = 1, \ldots, n)$, respectively.
3. For any $t = 1, \ldots, n$, let $(R_t, S_t) = (\widehat{F}_X(X_t), \widehat{F}_Y(Y_t))$, which corresponds to pseudo copula observations.
4. Select all the observations in the sets
$$\widehat{T} = \{(R_t, S_t) \mid R_t \leq \alpha, S_t \leq \alpha\}$$
5. Denote by $\mathscr{I}_T$ the set of all indices t's such that $(R_t, S_t) \in \widehat{T}$.
6. Calculate the univariate empirical cumulative distribution functions $\widehat{F}_{T,X}$ and $\widehat{F}_{T,Y}$ associated with all the observations $(X_t, Y_t)_{t \in \mathscr{I}_T}$.
7. For any index $t \in \mathscr{I}_T$, let $(R'_t, S'_t) = (\widehat{F}_{T,X}(X_t), \widehat{F}_{T,Y}(Y_t))$.
8. Calculate Spearman's correlation $\widehat{\rho}(C_T)$ given by
$$\widehat{\rho}(C_T) = \frac{12}{n_T} \sum_{t \in \mathscr{I}_T} R'_t S'_t - 3$$
where n_T is the cardinality of $\widehat{T}$.

3.4 The tail concentration function

While tail dependence coefficients (see Definition 3.2) give an asymptotic approximation of the behaviour of the copula in the tail of the distribution, it is often interesting to look at the tail behavioural considered at some (finite) points near the corners of the copula domain.

The present section discusses an auxiliary function that may serve to visualize the tail dependence of a copula C, the so–called *tail concentration function* (shortly, TCF). In particular, the latter is regarded as a graphical tool to visualize tail behaviour and provide useful information in the choice of the copula model adequately fitting the data. The TCF has been defined, for instance, in Venter (2002), while its estimation from empirical data has been presented in Patton (2012, 2013).

Consider a random vector (U, V) such that $(U, V) \sim C$. For any $t \in (0, 1)$, define $q_L(t) = \Pr(U < t, V < t)/t$ and $q_U(t) = \Pr(U > t, V > t)/(1 - t)$. In terms of the copula C one has

$$q_L(t) = \frac{C(t,t)}{t} \quad \text{and} \quad q_U(t) = \frac{1 - 2t + C(t,t)}{(1-t)}. \tag{3.14}$$

Definition 3.4. Given functions q_L, q_U as in (3.14), the tail concentration function is defined as the function $q_C : \mathbb{I} \to \mathbb{I}$ given by

$$q_C(t) = q_L(t) \cdot \mathbf{1}_{[0,0.5]} + q_U(t) \cdot \mathbf{1}_{(0.5,1]}. \tag{3.15}$$

Notice that $q_C(0.5) = (1 + \beta_C)/2$, where β_C is the Blomqvist's measure of association given by $\beta_C = 4C(0.5, 0.5) - 1$ (see Nelsen, 2006).

For practical purposes, the tail concentration function can be more suited to assess the risks of joint extremes than its limits given by the upper and the lower TDC's. In fact, when the speed of convergence of tail concentration function to the boundary 0 (or 1) is slow, this implies that the dependence in the finite tail can be significantly stronger than in the limit (compare with Manner & Segers, 2011).

The practical effect of considering the tail concentration function can be summarized in Figure 3.2. Here, we can note the different tail behaviour of several copulas sharing the same Blomqvist's measure of association. The left figure displays the TCF plots for copulas with zero lower (upper) tail dependence coefficient, while the right figure considers copulas with non-zero lower (upper) tail dependence (see Table 3.1). In both figures the Blomqvist's Beta is set to $\beta_C = 0.5$.

As it can be noticed, the TCF of Gaussian copulas seems to converge to 0 (respectively, 1) slowly than Frank and Plackett copulas. As regard to the copulas with non-negative TDC's, it seems that the convergence of the TCF to 0 (respectively, 1) is slower in the case of Clayton copulas (respectively, survival Clayton). Thus, Clayton copulas represent a natural choice for a conservative (from a risk manager viewpoint) estimation of the tail of a joint distribution. Notice that the TCFs of Gumbel and Galambos copulas are very close each other.

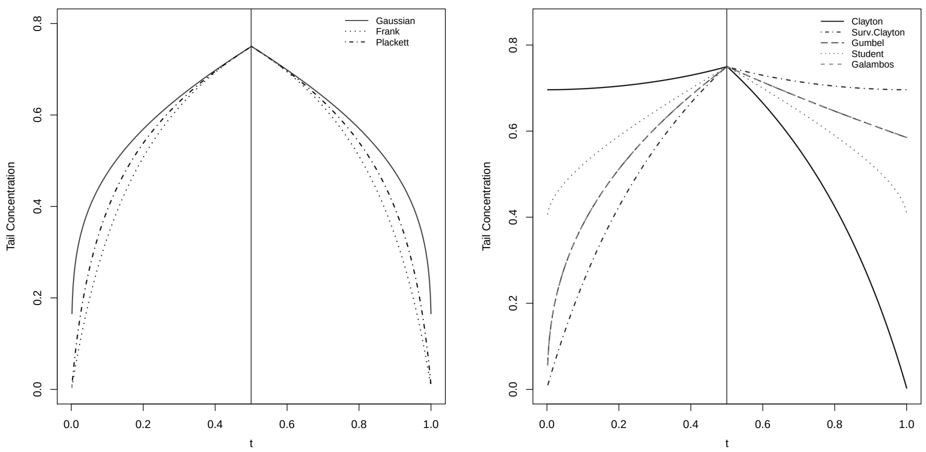


Figure 3.2: *Left* Tail concentration functions for a sample of size 1000 from a Gaussian, Frank and Plackett copula. *Right* Tail concentration functions for a sample of size 1000 from a Clayton, survival Clayton, Gumbel, Galambos copula and Student t-copula with $\nu = 4$ degrees of freedom.

In practice, given a random sample $\{(X_i, Y_i)\colon i = 1, \ldots, n\}$ from the random pair (X, Y) with copula C, the TCF can be estimated by the following procedure. First, the marginal distributions are estimated by their empirical versions,

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \le t), \qquad G_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(Y_i \le t)$$

Secondly, the copula C is estimated by the empirical copula C_n, given for all $(u, v) \in \mathbb{I}^2$ by

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_i \le u, V_i \le v),$$

where, for all $i = 1, \ldots, n$, $U_i = F_n(X_i)$ and $V_i = G_n(Y_i)$. Thus, for any $t > 0$, the (empirical) tail concentration function is given by

$$q_{\text{emp}}(t) = \frac{C_n(t,t)}{t} \cdot \mathbf{1}_{(0,0.5]}(t) + \frac{1 - 2t + C_n(t,t)}{1-t} \cdot \mathbf{1}_{(0.5,1)}(t) \tag{3.16}$$

Figure 3.3 presents the empirical estimation of the TCF for two time series from the MSCI World Index Data.

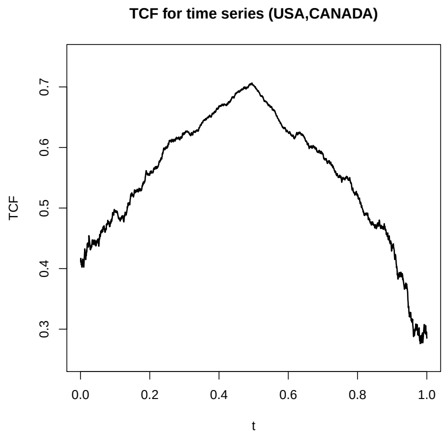


Figure 3.3: TCF related to two MSCI indices.

Notice that C_n depends on the sample size. In particular, $C_n(t,t)$ provides accurate estimation of the dependence structure when there is a sufficient number of points in $[0,t]^2$. Therefore, in order to allow more reliable estimate, the empirical tail concentration function is usually calculated on some interval $[\varepsilon, 1]$ for a suitable $\varepsilon > 0$ that is related to sample size.

3.5 Visualizing the tail-dependent copula-space

In Michiels & De Schepper (2013), a graphical tool has been provided in order to give some advice to the decision maker for the choice of the copula family to be used for fitting a given set of pairwise observations.

This method is based on two steps. First, one constructs the so–called *copula-test space*, *i.e.* the set of all possible families of copulas that are suitable for the data at hand (see Michiels & De Schepper, 2008). Such a space is, for instance, constructed by taking into account some empirical measures of association calculated from the observations. Then, a suitable distance between the empirical copula (as derived from data) and the (fitted) parametric families of copulas in the copula test space is introduced. The calculated distances are hence visualized in a Cartesian plane via principal coordinate analysis in such a way that one can assess visually which copula family is closer (in the given distance) to the empirical copula. In Michiels & De Schepper (2013) the considered distance is a L^2-type distance calculated on a finite grid of the copula domain. It measures the mean squared error between the graph of the empirical copula and the graph of a fitted parametric copula.

However, as stressed also in Michiels & De Schepper (2013), in some cases it would be also convenient the choice of a distance metric that provides different weights in the tails. In fact, while a global distance is surely convenient to have an idea about the overall goodness-of-fit of the proposed model, in most practical problems connected with risk management, one is also interested whether the model captures some important aspects of tail dependence behaviour. In fact, as stressed in Hua & Joe (2012), it may happen that the central part of the distribution influences the estimation more than the tail part, a fact that could be not conservative from the viewpoint of a risk manager that wants to estimates some risk quantities derived from the model (such as Value-at-Risk).

Therefore, here we introduce a variation of the graphical tool in Michiels & De Schepper (2013) in order to detect which families of copulas are closer to the empirical copula in the tail dependence behaviour. Specifically, the main ingredient of our tool will be the introduction of suitable dissimilarity measure based on the tail concentration function that can be used to visualize the distances among the copulas of the test space to the empirical copula.

To this end, let $(X_i, Y_i)_{i=1,\dots,n}$ be a bivariate sample from an unknown copula; or, better, use rank transformation to obtain from any sample the pseudo-observations that are useful to identify the copula structure (see, *e.g.*, Genest & Favre, 2007). Consider a set of k copulas $C_1, C_2, \dots, C_k$ belonging to different families that have been fitted to the available data. A dissimilarity between the empirical copula C_n and the copula C_i, $i = 1, \dots, k$ can be so defined:

$$\delta(C_n, C_i) = \int_0^1 \left(q_{\text{emp}}(t) - q_{C_i}(t)\right)^2 dt, \tag{3.17}$$

where q_{emp} is the TCF calculated via (3.16), while q_{C_i} is the TCF associated with C_i. In other words, we consider a kind of L^2-type distance between the empirical TCF and the TCF of a copula C_i fitted to the observations. Analogously, the dissimilarity between the i–th and the j–th copula is computed as

$$\delta(C_i, C_j) = \int_0^1 \left(q_{C_i}(t) - q_{C_j}(t)\right)^2 dt \tag{3.18}$$

for $1 \leq i \neq j \leq k$.

Obviously, dissimilarities (3.17) and (3.18) are both computed from a finite approximation of the TCF's at some points $t_1 < t_2 < \cdots < t_N$ in $\mathbb{I}$. Nevertheless, even for small sample size, they seem to provide hints to distinguish among different copulas. To this end, a simulation study is conducted in order to check whether the dissimilarity defined above are able to capture the tail behaviour of different copulas.

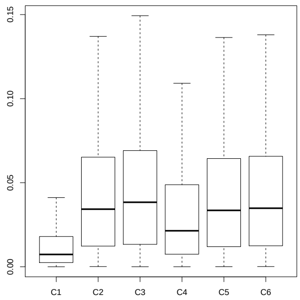
Specifically, we simulate bivariate observations of sample size $n \in \{250, 500\}$, respectively, from Clayton, Gumbel, and Gaussian copulas (with different values of Kendall's τ). Then we fit a parametric copula family to the data (via inversion of Kendall's τ). Finally we calculate the dissimilarity between the empirical copula and the fitted copula via (3.17). The considered families of copulas are Clayton (denoted by C_1), Gumbel (C_2), Frank (C_3), Gaussian (C_4), Plackett (C_5), Galambos (C_6). For $B = 500$ replications, this process produces a set of six matrices containing the dissimilarities between the empirical copula and the fitted copula, computed at each replication. The results are displayed in the form of box plots, for varying Kendall's tau values $\tau \in \{0.25, 0.5, 0.75\}$ (see Figures 3.4 and 3.5).

The following considerations can be drawn:

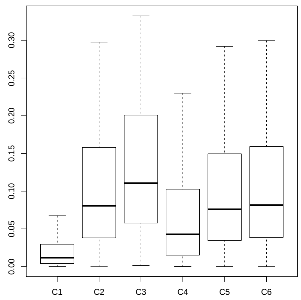
- For the Clayton family (C_1), the tail properties of the simulated copula is identified in all situations, regardless of τ-value and sample size. In fact, both for small and large sample sizes the box plots in panels (a) – (c) suggest that $\delta_{(\text{emp},1)}$ is minimal.

- For the Gumbel family (C_2), the simulation results suggest that the true copula is not always recognized as the best-fit copula. In particular, regardless of sample size, for $\tau \leq 0.5$ it seems that copulas C_3, C_5, C_6 exhibit small distances as well. For $\tau > 0.5$, the identification of the true data generating process seems more reliable.

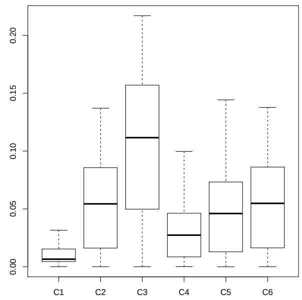
- For the Gaussian family (C_4) with small sample size ($n = 250$), other copulas exhibit a behaviour similar to the Gaussian copula for $\tau = 0.25, 0.5$,



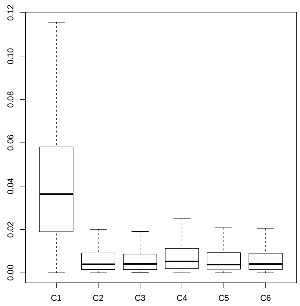
(a) Clayton copula, $\tau = 0.25$



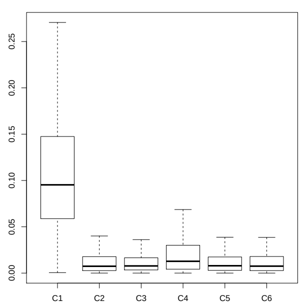
(b) Clayton copula, $\tau = 0.5$



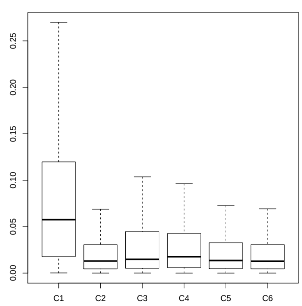
(c) Clayton copula, $\tau = 0.75$



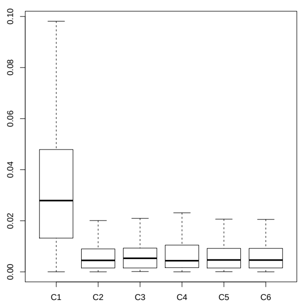
(d) Gumbel copula, $\tau = 0.25$



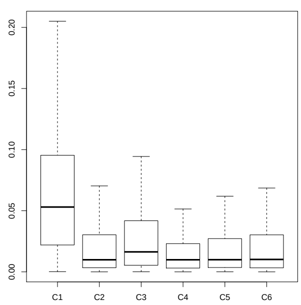
(e) Gumbel copula, $\tau = 0.5$



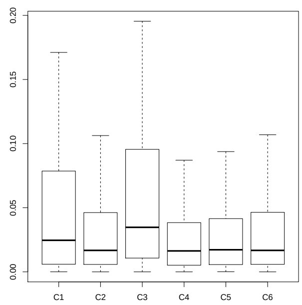
(f) Gumbel copula, $\tau = 0.75$



(g) Normal copula, $\tau = 0.25$

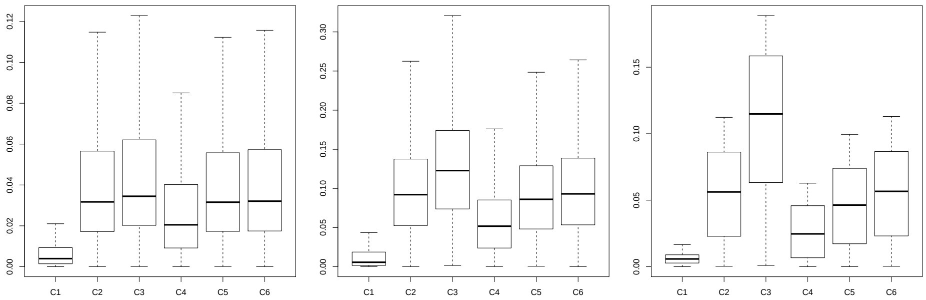


(h) Normal copula, $\tau = 0.5$

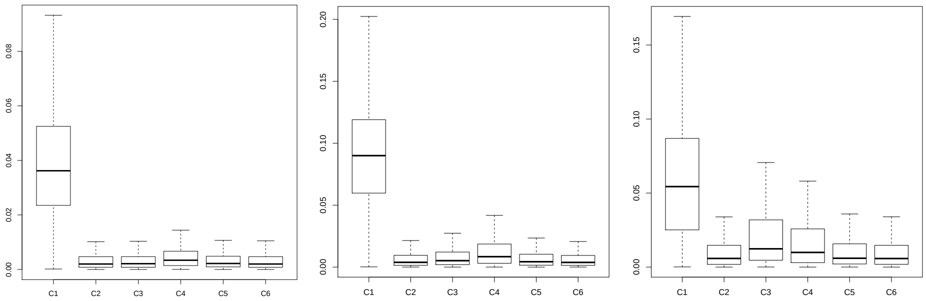


(i) Normal copula, $\tau = 0.75$

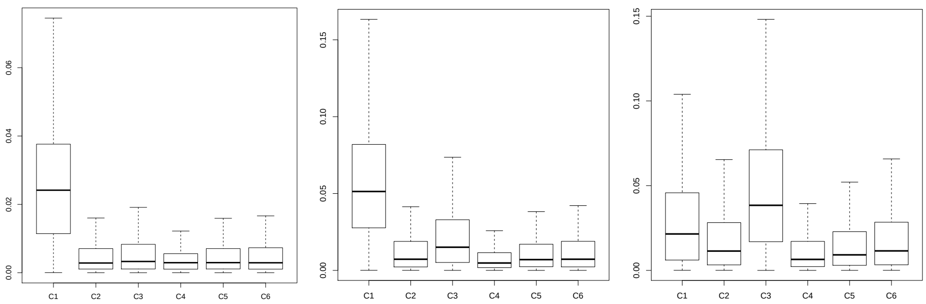
Figure 3.4: Box plots resulting from the simulation study with $n = 250$, for bivariate observations of sample size n from a Clayton (C_1), Gumbel (C_2) and Normal (C_4) copula with $\tau \in \{0.25, 0.5, 0.75\}$.



(a) Clayton copula, $\tau = 0.25$ (b) Clayton copula, $\tau = 0.5$ (c) Clayton copula, $\tau = 0.75$



(d) Gumbel copula, $\tau = 0.25$ (e) Gumbel copula, $\tau = 0.5$ (f) Gumbel copula, $\tau = 0.75$



(g) Normal copula, $\tau = 0.25$ (h) Normal copula, $\tau = 0.5$ (i) Normal copula, $\tau = 0.75$

Figure 3.5: Box plots resulting from the simulation study with $n = 500$, for bivariate observations of sample size n from a Clayton (C_1), Gumbel (C_2) and Normal (C_4) copula with $\tau \in \{0.25, 0.5, 0.75\}$.

while for $\tau = 0.75$ the identification of the true data seems to perform better. For larger sample size and $\tau \geq 0.5$, the Gaussian copula is unambiguously identified as the true copula.

Given such preliminary results, let us illustrate the graphical procedure to find some possible copula candidates to describe the tail behaviour of bivariate observations with unknown dependence structure.

Let $(X_i, Y_i)_{i=1,\ldots,n}$ be a bivariate sample from an unknown copula. Consider the set $\mathscr{C}^1, \ldots, \mathscr{C}^k$ of possible parametric families of copulas that can be suitable to describe the dependence in the given data (*i.e.* the copula test space). The procedure goes as follows.

1. For $i = 1, \ldots, k$ fit a copula C_i from the family $\mathscr{C}^i$ using classical methods (e.g. maximum likelihood estimation, inversion of Kendall's τ, etc.).
2. For $i = 1, \ldots, k$ calculate the dissimilarity between C_i and the empirical copula $\delta_{(\text{emp},i)} := \delta(C_n, C_i)$ by using (3.17).
3. For the k copulas $C_1, \ldots, C_k$, calculate the $K = k(k-1)/2$ mutual dissimilarities $\delta_{(i,j)} := \delta(C_i, C_j)$ by using (3.18).
4. Consider the symmetric matrix $D = (d_{ij})$ (of dimension $k+1$) defined as follows:

$$\begin{aligned} d_{1j} &= \delta_{(\text{emp},j-1)}, && j = 2, \ldots, k+1, \\ d_{ij} &= \delta_{(i-1,j-1)}, && i, j = 2, \ldots, k+1, \quad i < j \\ d_{ii} &= 0, && i = 1, \ldots, k+1. \end{aligned}$$

   Such a D is the dissimilarity matrix that describes the relation among C_n (the empirical copula), $C_1, \ldots, C_k$.

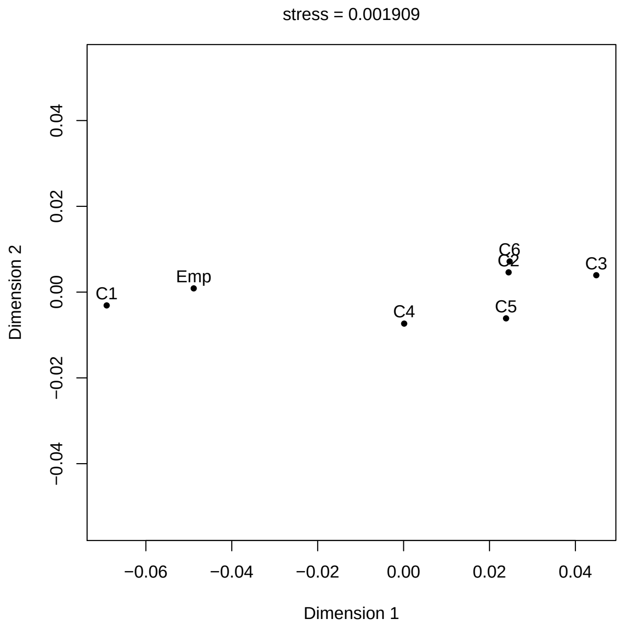
5. Starting with the dissimilarity matrix D, perform a Multidimensional Scaling (MDS) in order to construct a configuration of points in q dimensions, where the Euclidean distances (in the q–dimensional space) between the different copulas has to fit as closely as possible the dissimilarity information (for more details, see Kruskal, 1964a,b). As it is known from classical MDS, the final configuration is such that the distortion caused by a reduction in dimensionality is minimized by means of the so–called *stress function* (for more details, see Härdle & Simar, 2012).

In practice, a suitable visualization is obtained when $q = 2$, although $q > 2$ could be preferred in some cases (depending on the value of the stress function).

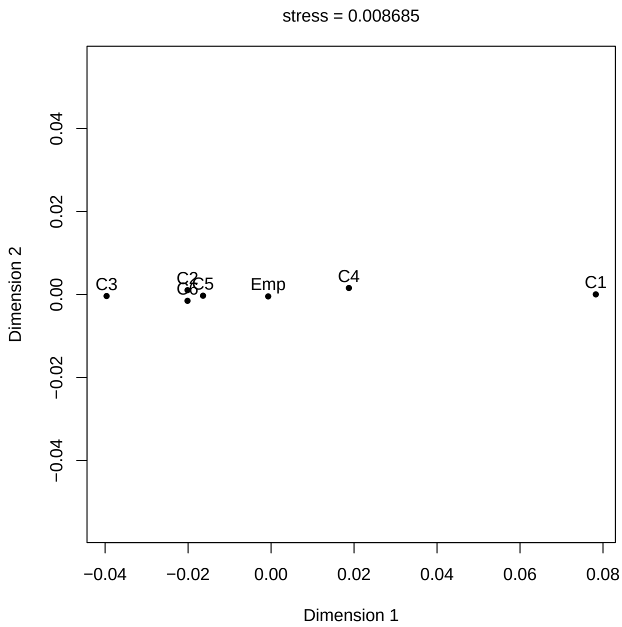
As an illustration of the methodology, consider a set of bivariate observations generated from a known copula and suppose that we would like to check whether the described graphical tool is able to suggest some good candidate copula model for such data. To this end, consider as copula test space the one-parameter copula families $\mathscr{C}^i$, $i = 1, \ldots, k$ mentioned above. Moreover, suppose that the random sample is of size $n = 250$, and is generated by Clayton, Gumbel, Frank, and Gaussian copulas, respectively, with Kendall's tau $\tau = 0.5$. The results for four different situations are displayed in Figures 3.6 and 3.7.

Specifically, for each chart of Figures 3.6 and 3.7, we apply MDS on the matrix of dissimilarities D defined as above, and check the stress function to see whether a 2D representation ($q = 2$) is feasible (in general a good representation should have a stress lower than 2.5%). Then, we plot the k points $p_i = (x_i, y_i)$ corresponding to copula C_i and $p_{\text{emp}} = (x_{\text{emp}}, y_{\text{emp}})$ corresponding to the empirical copula C_n in a 2D graph. Notice that the fitting of parametric copulas has been done via inversion of Kendall's τ. As can be seen, the charts are often useful to identify the true data generating process.

The graphical tool so obtained enables investigation of the goodness-of-fit by means of the relative distances between the empirical TCF and all TCF's of the copulas of the test space at once. By including many other copula families with different characteristics, it is possible to have a 2D visual overview of the whole collection of copulas based on their tail features as expressed by the tail concentration functions. This can be used as a copula selection tool in practical fitting problems, when one wants to choose one or more copulas to model the dependence structure in the data, highlighting the information contained in the tail.

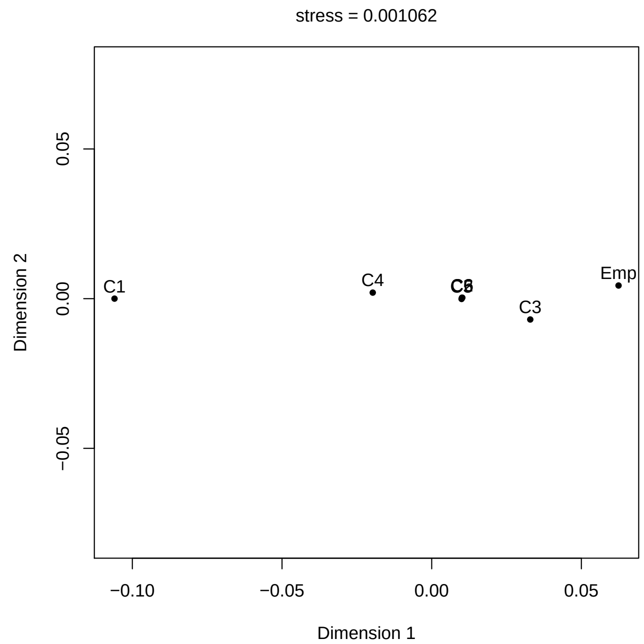


(a) Clayton copula (C_1), $\tau = 0.5$, $n = 100$

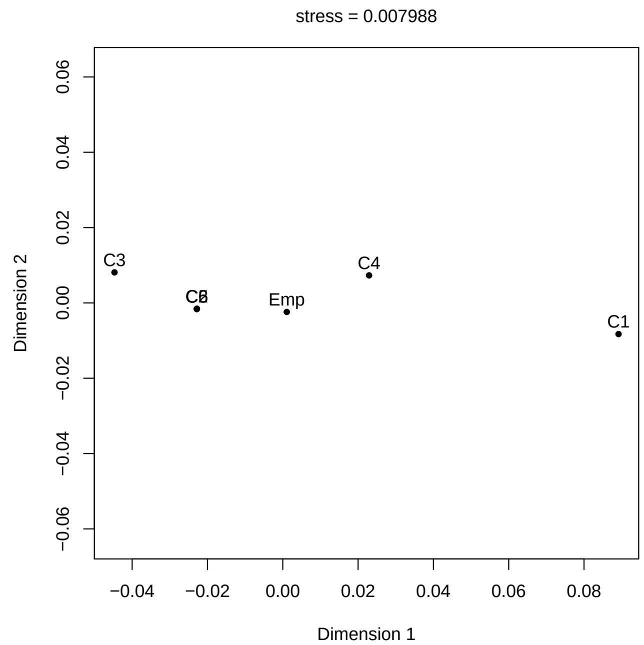


(b) Gumbel copula (C_2), $\tau = 0.5$, $n = 100$

Figure 3.6: Two-dimensional representation of copula test spaces for data generated by a Clayton and a Gumbel copula.



(a) Frank copula (C_3), $\tau = 0.5$, $n = 100$



(b) Normal copula (C_4), $\tau = 0.5$, $n = 100$

Figure 3.7: Two-dimensional representation of copula test spaces for data generated by a Frank and a Normal copula.

Chapter 4

Clustering financial time series by measures of tail dependence

4.1 Introduction

Clustering procedures represent an important tool in finance and economics, since practitioners are often interested in identifying similarities in financial assets for portfolio optimization and/or risk management purposes. In particular, the final goal consists in minimizing the whole risk of a portfolio of assets by adopting some diversification techniques which are based on the selection of different assets from markets and/or regions that one believes to be weakly correlated. In general, the clustering of a group of time series aims at finding sub-groups such that elements within a group have a similar stochastic dependence structure, while elements from distinct groups show a different behaviour. As summarized, for instance, by Liao (2005), these methods can be basically distinguished into three classes, depending upon whether they work (i) directly with the raw data (either in the time or frequency domain), (ii) indirectly with features extracted from the raw data, or (iii) indirectly with models built from the raw data. Clustering procedures typically involve the choice of a convenient dissimilarity measure. To this end, a number of approaches are available in the literature, which are based on different techniques like autoregressive distances (Piccolo, 1990; Corduas & Piccolo, 2008; Otranto, 2008), Mahalanobis-like distances (Caiado & Crato, 2010), variance ratio statistics (Bastos & Caiado, 2013), symbolic data analysis (Brida & Risso, 2010), latent class models (De Angelis, 2013), etc. For a collection of other approaches and applications we refer the reader to Pattarin et al. (2004); Tola et al. (2008) and the references therein. Several clustering methods have focused on the use of Pearson correlation in order to infer the hierarchical structure of a portfolio of financial assets: see, for instance, the book by Kaufman & Rousseeuw (1990) and the works by Mantegna (1999) and Bonanno et al. (2004). Their main idea is to consider a distance between time series that depends on the Pearson cross-correlation coefficient (or rank-based variants like Spearman's correlation), since high positive correlation may be interpreted in terms of some degree of similarity between the time series under consideration. From another perspective, models have been recently introduced in order to estimate the dynamics of correlation coefficients within groups of financial assets, with application to asset allocations (Engle, 2002; Billio et al., 2006; Billio & Caporin, 2009).

Recently, the need for alternatives to classical correlation-based clustering derives from the strong evidence that "classical correlation measures do not give an accurate indication and understanding of the real dependence between risk exposures" (Basel Committee on Banking Supervision, Joint Forum Developments in Modelling Risk Aggregation, October 2010). In particular, dependencies between extreme events such as extreme negative stock returns or large portfolio losses should be adequately modelled to support beneficial asset-allocation strategies especially when there is some contagion effect among the markets under consideration, namely when the positive association among the markets increases in crisis period with respect to tranquil periods. In such a situation, in fact, diversification may fail to work exactly when it is needed most (see, for instance, De Luca et al., 2010; Durante & Jaworski, 2010; Durante & Foscolo, 2013; Durante et al., 2013b). According to the Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), "diversification effects means the reduction in the risk exposure of insurance and reinsurance undertakings and groups related to the diversification of their business, resulting from the fact that the adverse outcome from one risk can be offset by a more favourable outcome from another risk, where those risks are not fully correlated". Thus, clustering techniques tailored to risk management should adopt alternative procedures, by taking into account the information about the tail behaviour of the involved quantities. In fact, if two financial markets have lower tail dependence coefficient different from 0, then they exhibit some positive dependence when both are experiencing very large losses. Such a different approach consists of finding groups that are similar in the sense that time series belonging to the same group tend to comove when they are experiencing large losses.

This chapter begins with some basic definitions and properties of financial returns, focusing on some stylized facts which impact on many financial applications, such as portfolio management, asset allocation, risk management. Then, the potential of copula functions in modelling the joint behaviour of asset returns is discussed. The interest is addressed on the problem of clustering financial time series. In Section 4.3 and 4.4 a general methodology for clustering financial time series via copula-approach is proposed. Preliminarily, a suitable stochastic model may be built in two steps: first, the marginals are fitted (means, variances, and distribution of the standardized residuals), then the standardized residuals of the univariate models are coupled via a suitable copula model. This is the starting point of clustering procedures described in the next two chapters.

4.2 Financial returns

Financial risk management is significantly based on the analysis of time series of returns. As stressed by Campbell et al. (1997), financial studies often involve returns for two main reasons: return series are a complete and scale-free summary of the investment opportunity, and show more attractive statistical properties than price series.

4.2.1 Prices and asset returns

Let P_t denote the price of an asset at time index t, where the t denotes the frequency (*e.g.*, yearly, monthly, daily). In general, the *return* expresses the relative change in the price of a financial asset over a given time interval. In the literature there are two types of returns: simple and compound.

Definition 4.1. For one period from date $t-1$ to date t, the *simple return* is

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}. \tag{4.1}$$

For n periods from date $t-n$ to n, the multiperiod simple return is given by

$$\begin{aligned} R_t(n) &= (1+R_t)(1+R_{t-1})(1+R_{t-2})\cdots(1+R_{t-n+1}) - 1 \\ &= \frac{P_t}{P_{t-1}}\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-n+1}}{P_{t-n}} - 1 = \frac{P_t}{P_{t-n}} - 1 \end{aligned}$$

where $1 + R_t(n) = P_t/P_{t-n}$ is called a compound return. A different representation of returns frequently used by practitioners is based on continuously compounded returns.

Definition 4.2. The continuously compounded return or *log return* is defined as the natural logarithm of the simple gross return of an asset and denoted by r_t:

$$r_t = \log(1+R_t) = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}). \tag{4.2}$$

It is easy to show that the continuously compounded multiperiod return is simply the sum of continuously compounded single-period returns:

$$\begin{aligned} r_t(n) &= \log(1+R_t(n)) = \log((1+R_t)(1+R_{t-1})(1+R_{t-2})\cdots(1+R_{t-n+1})) \\ &= \log(1+R_t) + (1+R_{t-1}) + (1+R_{t-2}) + \cdots + \log(1+R_{t-n+1}) \\ &= r_t + r_{t-1} + \cdots + r_{t-n+1}. \end{aligned}$$

This property is one of the reasons that make continuously compounded returns preferable rather than simple returns. Now, let p be a portfolio that places weight w_k on asset k. Then the simple return of p at time t, $R_{p,t}$, is computed as the weighted sum of the returns of the individual assets:

$$R_{p,t} = \sum_{k=1}^{K} w_k R_{kt},$$

where K is the number of assets and R_{kt} is the simple return of asset k. The above property does not hold for continuously compounded returns of a portfolio, since

$$r_{p,t} = \log\left(\frac{P_{p,t}}{P_{p,t-1}}\right) \neq \sum_{k=1}^{K} w_k \log\left(\frac{P_{k,t}}{P_{k,t-1}}\right),$$

where $r_{p,t}$ is the continuously compounded return of the portfolio at time t. However, for small returns (*e.g.*, daily) one has the approximation

$$r_{p,t} \approx \sum_{k=1}^{K} w_k r_{kt}.$$

4.2.2 Well-known properties of asset returns

In the last decades, the analysis of large data sets of high-frequency price series and the intensive empirical studies on financial time series have revealed a set of properties, common across many instruments, markets and time periods, observed by independent studies and classified as "stylized facts". Due to their generality, they are often qualitative. For a complete review on the subject we refer the reader to Cont (2001) and Danielsson (2011).

Many studies on financial assets highlight some empirical results that only hold some of the time. Return distributions are usually skewed either to the left or to the right, reflecting the asymmetric nature of asset returns. High frequency market returns exhibit negative autocorrelation, but the autocorrelation for the absolute and squared returns is always positive and significant, and decays slowly. A strong positive autocorrelation is usually observable over long periods of time during bull markets, while negative autocorrelation may characterize prolonged bear markets. Almost all financial returns exhibit three statistical properties regardless of asset type, sampling frequency, observation period or market: *volatility clusters*, *heavy tails*, *nonlinear dependence*.

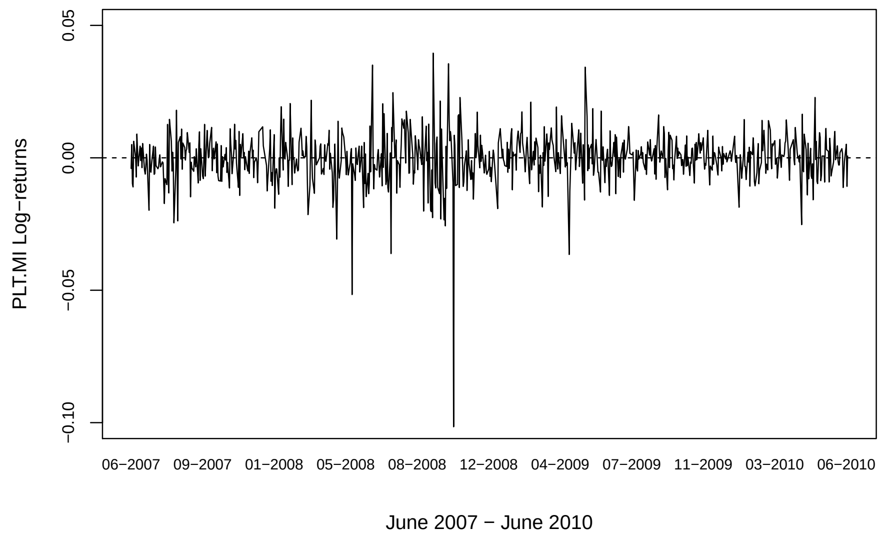


Figure 4.1: Daily log-returns of the adjusted stock price of Parmalat (PLT.MI) for the period June 2007 - June 2010.

Volatility

In financial time series analysis, volatility means the conditional standard deviation of the underlying asset return. Volatility evolves over time and, although it is not directly observable, it has some characteristics that are commonly seen in asset returns. Since the publication of Engle (1982), the phenomenon of *volatility clusters* has been accepted as a stylized fact about asset returns. It can be explained as the tendency of high-volatility events to cluster in time, that is, large price variations are more likely to be followed by large price variations. Thus, one usually observes that a the market goes through periods with high volatility and other periods when volatility is low. The sample autocorrelation function of the squared returns is commonly used to detect volatility clustering, since squared returns are proxies for volatilities. More recent results (see, *e.g.*, Bouchaud & Potters, 2001) have analysed one more important feature of volatility which seems to be rather universal: the volatility of stocks tends to increase when the price drops, referred to as the *leverage* effect. This effect corresponds to a negative correlation between past returns and future volatility, and is asymmetric.
All these properties play an important role in modelling univariate financial time series. The first model designed to capture volatility clusters was the autoregressive conditional heteroscedastic (ARCH) model of Engle (1982). Bollerslev (1986) proposed a useful extension known as the generalized ARCH (GARCH) model which has the potential to incorporate the impact of historical returns. The GARCH-type models belong to the category of conditional volatility models and are based on optimal exponential weighting of historical returns to obtain a volatility forecast. However, as for the ARCH model, GARCH model encounters some weaknesses. For instance, it responds equally to positive and negative shocks. Moreover, only lower order GARCH models are used in most applications, say, GARCH(1,1), GARCH(2,1) and GARCH(1,2) models. Subsequently, a large number of extensions to the GARCH model have been proposed to overcome some weaknesses of the earlier models in handling financial time series. For example, Nelson (1991) proposed the exponential GARCH (EGARCH) model to account for asymmetric effects between positive and negative asset returns. Another widely used GARCH model allowing for leverage effects is the model of Glosten et al. (1993) (GJR-GARCH, also known as threshold-GARCH). Many other volatility models not

mentioned here are available in the literature. We refer the reader to Tsay (2005) for a comprehensive review on financial econometric models and their applications.

Heavy tails

The normality assumption has long dominated conventional asset allocation frameworks. In reality, we can empirically observe that in many cases returns are not independent, and in all cases they are not normally distributed. In statistics, *kurtosis* measures the degree of peakedness of a distribution relative to the tails. Normally distributed variables have kurtosis equal to 3. Usually, a positive excess kurtosis (over 3) indicates the presence of heavy tails (or fat tails). This means that the distribution puts more mass on the tails of its support than a normal distribution does. Thus, a random sample from such a distribution tends to contain more extreme values. The heavy-tailed property of returns has been known since Mandelbrot (1963) and Fama (1963, 1965). According to the survey by Cont (2001), "the (unconditional) distribution of returns seems to display a power-law or Pareto-like tail, with a tail index which is finite, higher than two and less than five for most data sets studied. In particular this excludes stable laws with infinite variance and the normal distribution. However the precise form of the tails is difficult to determine". Non-normality of asset returns has many important consequences in finance, especially in the field of risk management, where the main concern is on the far left tail of the return distribution. Extreme negative past events such as the stock market crash of 1987, the bursting of U.S. technology bubble in 2000-2001 or the recent financial crisis of 2007-2009, have confirmed that assuming normality in risk calculations may cause large underestimations of risk, leading to dramatic consequences. Concluding, financial time series exhibit high variability, as revealed by the heavy-tailed distributions of their increments and the non-negligible occurrence of extreme negative events. These considerations motivate numerous theoretical efforts to understand the intermittent nature of financial time series and to model adequately the tails of the distribution of returns.

Nonlinear dependence

Most statistical models assume that the joint distribution of returns is Gaussian, implying that we can measure dependence by using correlations, such as Pearson's correlation coefficient. It is important to stress that two returns can be uncorrelated (in the linear or Pearson sense) but dependent, since only linear dependencies are detected. Considerable recent research has shown that the dependence between different return series changes according to market conditions and, in particular, correlations under extreme conditions are quite different than under normal conditions. Relying on linear correlation matrices leads to underestimate the probability of joint negative returns during periods of high market volatility. Thus, for many financial applications it is essential to address nonlinear dependence, allowing the dependence structure to vary according to markets behaviour.

4.3 Clustering financial assets: a copula-based approach

Most practical problems in risk management deal with portfolios containing a certain number of assets. The statistical analysis of the risk of such positions requires information on the joint behaviour of the returns of different assets. By considering joint distributions, we turn our focus to how assets behave together during periods of market stress. In particular, copulas allow us to model more accurately an increased incidence of joint negative returns, without any influence of marginal behaviour on the dependence structure. In the literature, copula-based time series models have been used as a tool to handle in a flexible way the link among different univariate time series (see Patton, 2012, and the references therein).
Following these ideas, a suitable stochastic model may be built in two steps: first, the marginals are fitted (means, variances, and distribution of the standardized residuals), then the standardized residuals of the univariate models are coupled via a suitable copula model.
We consider a matrix of d financial time series $(x_{it})_{t=1,\dots,T}$ $(i = 1, 2, \dots, d)$ representing the returns of different financial assets. We assume that each time series $(x_{it})_{t=1,\dots,T}$ is generated by the stochastic process $(\mathbf{X}_t, \mathscr{F}_t)$ such that, for $i = 1, \dots, d$,

$$X_{it} = \mu_i(\mathbf{Z}_{t-1}) + \sigma_i(\mathbf{Z}_{t-1})\varepsilon_{it}, \tag{4.3}$$

where $\mathbf{Z}_{t-1}$ depends on $\mathscr{F}_{t-1}$, the available information up to time $t-1$, and the innovations ε_{it} are distributed according to a distribution function F_i for each t. Moreover, the innovations ε_{it} are assumed to have a constant conditional distribution F_i (with mean zero and variance one, for identification) such that for every t the joint distribution function of $(\varepsilon_{1t}, \ldots, \varepsilon_{dt})$ can be expressed in the form $C(F_1, \ldots, F_d)$ for some copula C.

As known (see, *e.g.*, Jaworski et al., 2010, 2013), the copula C is exactly the function that captures the dependence properties of the time series. As discussed in the previous chapters, the rank-invariant measures of association (Kendall's tau, Spearman's rho and their conditional versions) and the tail dependence coefficients are based on the calculation (in a parametric or non-parametric way) of the respective copula. Here we are assuming that some of the information contained in $\mathscr{F}_{t-1}$ is not relevant for all variables. In particular, each variable depends on its own first lag, but not on the lags of any other variable (see the discussion in Patton, 2009, pp. 772–773). Thus we can construct each marginal distribution model using only the information relevant for each variable, which will likely differ across margins, and then use $\mathscr{F}_{t-1}$ for the copula, to obtain a valid conditional joint distribution. As stressed by Fermanian & Wegkamp (2012), even if the assumption that each variable depends on its own lags seems strong, it has been proved to be reasonable for many empirical studies.

The following steps are implemented in order to group the time series into sub-groups such that elements in each sub-group have strong tail dependence.

1. Choose a suitable copula-based time series model (*e.g.* ARMA-GARCH copula model) in order to model separately the marginal behaviour of each time series and the link between them.
2. Estimate a suitable (pairwise) tail dependence measure among the different time series.
3. Define a dissimilarity matrix by using the information contained in the tail dependence measures and apply a suitable cluster algorithm, according to the general characteristics of the above introduced dissimilarity matrix.

As a relevant feature of our approach, we assume that the multivariate time series process follows a copula-based semi-parametric model that allows to separate the univariate behaviour of each time series from the dependence among them. Moreover, we avoid the specification of a fully parametric model for describing the pairwise dependence between the markets under consideration.

4.4 Fit a copula-based time series model

The choice of the univariate model is made by classical model selection procedures (*e.g.*, Bayesian Information Criteria) and the goodness-of-fit verified by classical tests of homoscedasticity and uncorrelatedness of the residuals (Patton, 2012, 2013). Different models (with different parameters) can be estimated for each univariate time series. In particular, the GARCH-type models have proved to be adequate in modeling return series. Among different extensions, the GJR-GARCH model can be considered in order to capture both the excess of kurtosis and the asymmetric effects (see Section 4.2). For a time series of returns x_t, let ε_t denote the error terms (return residuals, with respect to a mean process). Then, a GJR-GARCH(p,q) (or TGARCH) model assumes the form

$$\varepsilon_t = \sigma_t \eta_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \gamma_i N_{t-i}) \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_j \sigma_{t-j}^2$$

where $\{\eta_t\}$ is a sequence of independent and identically distributed innovations with zero mean and unit variance, N_{t-1} is the indicator function such that

$$N_{t-1} = \begin{cases} 1 & \text{if } \varepsilon_{t-i} < 0, \\ 0 & \text{if } \varepsilon_{t-i} > 0, \end{cases}$$

and α_i, γ_i, β_j are non-negative parameters satisfying $\sum_{i=1}^{\max p,q} (\alpha_i + \beta_i) < 1$. The model uses zero as the threshold to separate the impacts of past shocks, and negative returns have a larger impact to σ_t^2.
In the second step, using the parametric models estimated in previous step, we compute the estimated standardized residuals given, for each $i = 1, \dots, d$ by

$$\widehat{\varepsilon}_{it} = \frac{x_{it} - \widehat{\mu}_i(\mathbf{Z}_{t-1})}{\widehat{\sigma}_i(\mathbf{Z}_{t-1})}.$$

Finally, the estimated standardized residuals are converted to the estimated probability integral transform variables $z_{it} = F_i(\widehat{\varepsilon}_{it})$, where F_i may be estimated from a parametric model (Gaussian, Student t, etc.) or by using the empirical distribution function.

Now, for all $t = 1, \dots, T$ the points $(z_t^1, \dots, z_t^d) := (z_{1t}, \dots, z_{dt})$ contain the information about the link (*i.e.* the copula) among the time series under consideration and can be used in order to compute simple dependence measures or make

inference about the copula of the time series.
As stressed, for instance, by Jondeau & Rockinger (2006), if the marginal model is correctly specified, the estimated probability integral transforms $(z_t^1, \ldots, z_t^d)_{t=1,\ldots,T}$ behave asymptotically like a random sample generated by the copula C. As such, dependence measures calculated from this sample are not biased by serially dependence and/or heteroscedasticity. In particular, this latter aspect is very important since the bias caused by volatility effects may induce serious inaccuracies in the calculation of conditional correlations, as stressed for instance by Forbes & Rigobon (2002) and Bradley & Taqqu (2004). As a further remark, notice that, according to Patton (2012), "the estimated parameters from the conditional mean and variance do not affect the asymptotic distribution of estimated dependence measures such as Spearman's rank correlation and Kendall's tau", that is "we can ignore the error resulting from the estimation of the marginal distribution parameters". In other words, following also the results by Rémillard (2010), we may say that the rank-based dependence measures (like Kendall's tau, Spearman's rho) behave asymptotically like the ones computed from innovations, extending the results of Chen & Fan (2006).

Once obtained the pseudo–observations from the original time series, we adopt two different measures in order to quantify the degree of dependence in the tail of the joint distribution function, namely the lower TDC λ_L and the conditional Spearman's correlation ρ_α. Then, suitable tail dependence-based dissimilarity measures can be defined and used as input in classical cluster analysis tools. The two procedures are described in details in the next two chapters.

Chapter 5

Clustering financial time series via conditional Spearman's rank correlation

5.1 Introduction

The analysis of the association between two random variables has been extensively studied in the literature: for an overview, see the books by Joe (1997), Nelsen (2006), Jaworski et al. (2010), Cherubini et al. (2011), Jaworski et al. (2013). In particular, different kinds of measures have been introduced in order to quantify the association between the variables of interest when they are taking on very large (respectively, small) values: see, for instance, Schmid et al. (2010), Bernard et al. (2013) and the references therein.
One possible way to consider such a kind of dependence is to restrict to a conditional version of the classical Pearson correlation coefficient, as done for instance by Longin & Solnik (2001); Malevergne & Sornette (2006) (the so–called extreme and exceedance correlations). However, as said, Pearson's correlation coefficient is often an inappropriate dependence measure. First, it measures only linear dependence. Secondly, it is not invariant to a change of the univariate margins, and thirdly, it is very sensitive to outliers (Schmid & Schmidt, 2007).

In order to overcome these pitfalls, we suggest a suitable conditional version of Spearman's correlation coefficient ρ_S, where the conditioning set is defined according to some given threshold (see Chapter 3, Section 3.3). Intuitively, we focus the attention on the behaviour of the markets exposed to losses that are judged to be extreme according to a predefined "risky" level. Conditional versions of Spearman's correlation have been also adopted in the detection of contagion among financial markets (see, for instance, Durante & Jaworski (2010) and the references therein). Compared with conditional Pearson's correlation, Spearman's correlation is rank invariant (and, hence, more suitable for checking non-linear comovements in the data) and only depends on the copula of the involved random variables. Moreover, it is adapt to deal with non-Gaussian data (see, for instance, Embrechts et al., 2002; McNeil et al., 2005).

Starting from the approach described in Chapter 4 (Sections 4.3, 4.4), this chapter presents a methodology for clustering financial time series according to their dependence in risky scenarios. The procedure is based on the calculation of suitable pairwise conditional Spearman's correlation coefficients extracted from the series. The performance of the proposed methodology is checked via a simulation study in Section 5.5. An application to the analysis of the components of the Italian FTSE–MIB stock index is given in Section 5.6. Section 5.7 is devoted to final remarks.

5.2 Define a measure of tail dependence

Given two random variables X_i, X_j and a threshold $\alpha \in (0,1)$ representing the risky level, we are interested in the Spearman's correlation of the conditional distribution of $(X_i, X_j \mid (X_i, X_j) \in T_\alpha^{ij})$, where $T_\alpha^{ij} = (-\infty, q_\alpha(X_i)] \times (-\infty, q_\alpha(X_j)]$ is a set of non-zero probability and $q_\alpha(X_i)$ is the α–quantile of X_i for every i. In the following, such a coefficient will be denoted by $\rho_S^{ij}(\alpha)$. From equation (3.13) we know that

$$\rho_S^{ij}(\alpha) = 12 \int\int_{\mathbb{I}^2} C_{T_\alpha^{ij}}(u,v)dudv - 3, \tag{5.1}$$

where $C_{T_\alpha^{ij}}$ is the (threshold) copula associated with the conditional d.f. of $(X_i, X_j \mid (X_i, X_j) \in T_\alpha^{ij})$. The estimation of (5.1) consists in the following steps.

1. We consider the estimated probability integral transform $z_{it} = \widehat{F}_i(\widehat{\varepsilon}_{it})$, where

$\widehat{F}_i$ is the empirical distribution function

$$\widehat{F}_i(t) = \frac{1}{T}\sum_{i=1}^{T} \mathbf{1}_{\widehat{\varepsilon}_{it} \leq t}$$

2. Then we restrict to all points $(z_t^i, z_t^j)_{t=1,\dots,T}$ such that $z_t^i \leq \alpha$ and $z_t^j \leq \alpha$. Let us denote this set by $(z_t^i, z_t^j)_{t=1,\dots,T_\alpha}$. The value $\rho_S^{ij}(\alpha)$ can be estimated by using the standard population version of Spearman's correlation (see, *e.g.*, Genest & Favre, 2007) applied to the (restricted sample) $(z_t^i, z_t^j)_{t=1,\dots,T_\alpha}$.

At the end of the procedure, we obtain a value $\rho_S^{ij}(\alpha)$ representing the association between the financial markets time series i and j when both markets are experiencing severe losses.

Remark 1. Under suitable specification of univariate margins, $(z_t^1, \dots, z_t^d)_t$ behaves asymptotically like a random sample generated by the copula C (and hence the marginals are uniform). In particular, for each i, the α–quantile $(z_t^i)_{t=1,\dots,T}$ is asymptotically equal to α. Due to general results about order statistics (see, *e.g.*, Wilks, 1948), it follows that $(z_t^i, z_t^j)_{t=1,\dots,T_\alpha}$ is also asymptotically a random sample.

5.3 Define a dissimilarity measure

Now, we have to define a suitable dissimilarity matrix that collects the information about the pairwise tail dependence among the time series. We usually assume that a dissimilarity measure δ between time series (x_{1t}) and (x_{2t}) satisfies the following properties:

1. non negativity, *i.e.* $\delta((x_{1t}), (x_{2t})) \geq 0$;
2. identity, *i.e.* $\delta((x_{1t}), (x_{1t})) = 0$;
3. symmetry, *i.e.* $\delta((x_{1t}), (x_{2t})) = \delta((x_{2t}), (x_{1t}))$.

In addition to properties 1-3, we require that the dissimilarity measure δ decreases in a monotone way as (x_{1t}) and (x_{2t}) are more and more similar, according to the idea of similarity one has adopted.

As explained in the previous section, we are considering conditional Spearman's correlation with respect to a given tail set, as specified by a threshold $\alpha \in$

$(0, 1)$. Such an α denotes the "degree" of the risk of the tail scenario we are considering. In practice, we restrict to consider $\alpha \in \{0.05, 0.10, 0.25\}$. Thus, for each pair (i, j) of time series, we calculate the Spearman's correlation $\rho_S^{ij}(\alpha)$ as indicated in the previous step. For instance, $\rho_S^{ij}(0.05)$ refers to the tail dependence between market i and market j when both are experiencing severe losses that happens, on average, with a probability of 5% for each individual market.

Then, for $i, j = 1, \dots, d$, we may define the dissimilarity matrix $\Delta = (\Delta_{ij})$ whose elements are given by

$$\Delta_{ij} = \sqrt{2(1 - \rho_S^{ij}(\alpha))}. \tag{5.2}$$

It is easy to see that properties 1-3 hold for measure (5.2). Moreover it represents a distance matrix, as showed for instance in Bonanno et al. (2004).

Notice that other ways of findings a correlation-type dissimilarity matrix have been provided in the literature, as explained for instance by Kaufman & Rousseeuw (1990). For instance, one might consider $\Delta_{ij} = 1 - \rho_S^{ij}(\alpha)$ or $\Delta_{ij} = 1 - |\rho_S^{ij}(\alpha)|$. In our study, we have found no significant difference among the use of these possible choices.

5.4 Apply a suitable cluster algorithm

The dissimilarity matrix Δ defined above could be used to determine clusters among the d time series of financial returns by means of a suitable procedure. For our purposes, we consider the hierarchical agglomerative clustering techniques frequently used in practice (see, for instance Friedman et al., 2009). As is known, hierarchical agglomerative algorithms start from the finest possible partition (*i.e.* each observation forms a cluster) and, hence, each level merges a selected pair of clusters into a new cluster according to the definition of the distance between two groups. This sequence of nested partitions is best visualized as a top-down tree called a *dendrogram*, such that the dissimilarity between merged clusters is monotone increasing with the level of the merger. Among all the agglomerative strategies we may apply, the three most common clustering procedures (which differ in the computation of the distance between two groups) are *single linkage*, *complete linkage*, *average linkage*. In particular, the single linkage defines the distance between two groups as the smallest value of the individual distances; the complete linkage algorithm defines the distance between two clusters as the maximum distance between their individual components; the average linkage (weighted or not weighted) merges the two latter algorithms, since it computes an average distance. Notice that the dissimilarity matrix need to be further transformed in order to obtain a (Euclidean) distance matrix if one wants to adopt some specific cluster procedures (*e.g.*, Ward method).

5.5 Simulation study

A set of simulation studies has been designed in order to explore the performance of the proposed clustering procedure, with respect to the following parameters:

- the number d of the financial assets;
- the number J of different clusters;
- the strength of the tail correlation inside each subgroup of assets;
- the sample size T.

Such a study aims at clarifying whether, even in a finite sample, the behaviour of conditional Spearman's correlation guarantees that the whole clustering procedure performs in a good way.

Specifically, according to the explanation in Section 5.2, we simulate a sample of all possible realizations $(z_t^1, \ldots, z_t^d)_{t=1,\ldots,T}$ that can be extracted from a multivariate time series model when some appropriate univariate time series models are fitted to the marginals. That is, we simulate directly the estimated probability integral transforms, since the estimation of conditional Spearman's correlation matrix is not asymptotically biased by the estimation of the parameters of the univariate model, as explained above. The realizations $(z_t^1, \ldots, z_t^d)_{t=1,\ldots,T}$ are supposed to be generated by a d–dimensional copula model of the form

$$C(\boldsymbol{u}) = \prod_{j=1}^{J} C_j(u_{j1}, \ldots, u_{jk_j}), \tag{5.3}$$

for all $\boldsymbol{u} = (u_{11}, \ldots, u_{1k_1}, u_{21}, \ldots, u_{2k_2}, \ldots, u_{J1}, \ldots, u_{Jk_J})$, with $k_1 + \cdots + k_J = d$. Hence, C is the product of J copulas C_j, where C_j comes from a Clayton (respectively, survival Gumbel) family with a fixed lower TDC λ. We recall that

Clayton and survival Gumbel copulas are typical examples of copulas exhibiting non-zero TDC. By the very definition, model (5.3) describes the behaviour of $(k_1 + k_2 + \cdots + k_J)$ random variables that can be grouped into J clusters in such a way that variables in the same cluster j are tail dependent with a non-zero tail dependence coefficient driven by the copula C_j. By contrast, random variables in different clusters are supposed to be independent and, hence, they are not tail dependent. In other words, this model is able to interpret the presence of clustered behaviour in the tails. Recently, other copula models have proved to be useful to model tail dependence behaviour among groups and could be used as well (Czado, 2010; Brechmann, 2013). In order to examine the performances of our methods in a variety of situations, we use the following value parameters for model (5.3):

- the threshold used to determine the tail region of interest set to $\alpha = 0.25$;
- the sample size $T = 400, 800$;
- the dimension of the model $d = 32, 64, 128$;
- the number of copulas $J = 4, 8, 16$, with $J < d/2$, considered in (5.3);
- the lower TDC $\lambda = 0.10, 0.30, 0.50$.

Notice that, in the simulation study, there is a trade-off between sample size and choice of the threshold α. In fact, when α becomes smaller, one needs to increase the sample size in such a way that the tail region T_α contains a sufficient number of points that allow the calculation of the conditional Spearman's correlation. In order to overcome such a problem, we prefer here to fix $\alpha = 0.25$ and to assume that the dimension of the sample size is not too large (according to financial practice). It should be mentioned that various similar simulations have been also carried out by using a small α (with large samples) and similar results have been obtained. After simulating a sample, a dissimilarity matrix is created as described in Section 5.3 and the hierarchical clustering via complete linkage is applied. Moreover, for each sample, we determine the optimal number of clusters g by the silhouette index (Kaufman & Rousseeuw, 1990), which reflects the within-cluster compactness and between-cluster separation of a clustering. In detail, for $g = 1, 2, \cdots$, the number of clusters is chosen such that the average silhouette width is maximized over all g. It follows that, if the correct number of clusters is identified, then it coincides with J. Otherwise, the system has misspecified the cluster structure, a fact that will decrease the performance of our methodology.

Finally, in order to examine whether the structure of the obtained clusters matches to the true classification of the instances, two indices are considered: the Rand Index (RI) (Rand, 1971) and the Adjusted Rand Index (ARI) (Hubert & Arabie, 1985). The RI lies between 0 and 1, where the maximum value is taken when two partitions agree perfectly. The ARI is the corrected-for-chance version of the RI, so as to ensure that its maximum value is 1 and its expected value is zero when the partitions are selected at random. It can can yield a value between -1 and 1. Specifically, we calculate the RI and ARI between the cluster structure obtained from the described method and the theoretical one (as induced from model (5.3)). The calculations are repeated 250 times, and the average RI (respectively, ARI) is considered. We also considered more replications without obtaining relevant changes in the results. The results are presented in Tables 5.1, 5.2 and 5.3.

As can be seen, overall the simulations give reasonable results, even for a small sample size. However, as the sample size increases, the performance generally improves, as expected. Moreover a stronger tail dependence within groups (according to λ) clearly improves the results, regardless of copula family, J-value and sample size. The performance of the tested procedure seems also to be related to the type of dependence (*i.e.* the copula) that is involved since, as known, different copulas with the same tail dependence coefficient may have different (finite) tail behaviour, *i.e.*, different lower threshold copula. The number of clusters does not seem to have a clear influence on the results, even if in several cases the performances seem to become weaker as J increases. This can be explained by the fact that we do not fix *a priori* the number of clusters to be considered but we allow the algorithm to select this number according to the silhouette value.

5.6 Application to FTSE-MIB Index

In order to illustrate our approach we analyse daily log-returns of the components of the FTSE–MIB index in the period from June 4, 2007 to June 29, 2012. The data were downloaded from Datastream and are formed by the log-returns of adjusted stock prices $(p_t^i)_{t=1,\dots,T}$ of 34 assets (6 assets have been removed). By considering only the days when all assets were operating, we collect $T = 1292$ observations. The stocks are listed in Table 5.4.

We preliminary fit GJR-GARCH(1,1) models to the univariate time series with

Student-t distributed errors to account for heavy tails. For all time series we then perform Box-Pierce and Ljung-Box tests at lags 1 and 5, to check for residual autocorrelation, ARCH tests at lags 1 and 5, for autoregressive conditional heteroscedasticity and Kolmogorov-Smirnov test to check for the Student hypothesis for the standardized residuals. The estimation results show a reasonable fit for all time series (see Table 5.5). By using the procedure of Section 4.4, we derive hence the sample $(z_t^1, \dots, z_t^{34})$ on $[0,1]^{34}$. In order to restrict our analysis to extreme observations we fix a threshold α denoting the "degree" of risk of the scenario we are considering (for our application we set $\alpha = 0.05, 0.10, 0.25$). In particular, for each choice of the level α, we compute $d(d-1)/2 = 561$ coefficients associated with the pairs (z_t^i, z_t^j), $i \neq j$, $i, j = 1, \dots, 34$ conditional to the fact that (z_t^i, z_t^j) takes values on $[0, \alpha]^2$, and denote them by $\rho_S^{ij}(\alpha)$.

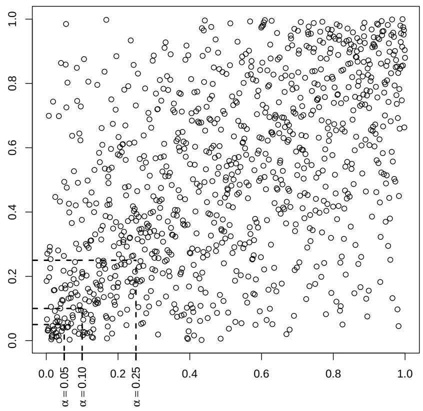


Figure 5.1: Illustration of tail regions defined by the threshold α in a bi-dimensional plot of two series of residuals.

As pointed out before, the starting point for our clustering procedure is a distance defined through the correlation matrix. To this end, hierarchical agglomerative clustering algorithms are applied directly to the matrix $\Delta = (\Delta_{ij})_{i,j=1,\dots,34}$ with $\Delta_{ij} = \sqrt{2(1-\rho_S^{ij}(\alpha))}$. As a clustering procedure, we adopt the complete linkage method, which has proved to be useful in a variety of situations. As a matter of fact, other methods like single linkage and average linkage may be applied as well. The dendrograms produced by complete linkage scheme in each of the three extreme scenarios are displayed in Figure 5.3 together with the associated

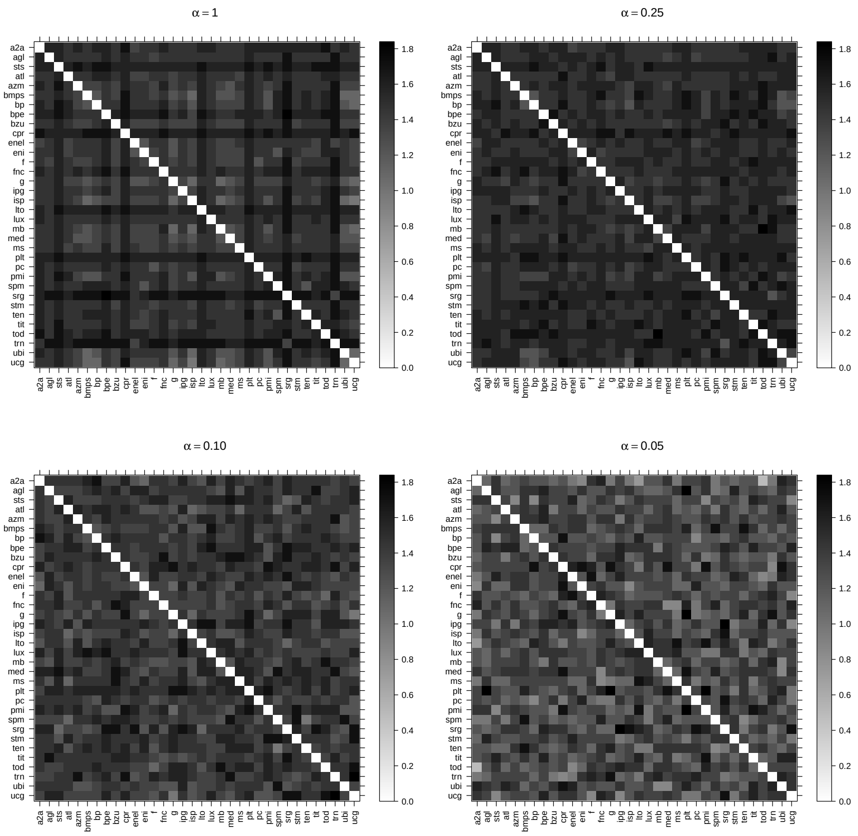


Figure 5.2: Heat Maps for 34 FTSE–MIB components displaying the pairwise dissimilarities for different threshold levels $\alpha \in \{1, 0.25, 0.10, 0.05\}$.

threshold level $\alpha = 0.05$, $\alpha = 0.10$, $\alpha = 0.25$, respectively. For completeness we also report the complete linkage dendrogram when we use unconditional Spearman's correlation (*i.e.* we assume $\alpha = 1$). From here, the number of clusters can be estimated by using classical results in the literature (see, *e.g.*, Everitt, 1979). Moreover, notice that in most real life situations, the number of clusters can be either specified by the user based on his prior knowledge (for instance, geographic or economic considerations) or estimated via a suitable procedure (see, *e.g.*, Milligan & Cooper, 1985; Gordon, 1999).

As can be seen from the dendrograms, the clustering hierarchy varies according to the threshold level α. In fact, it may happen that the dependence between two assets changes according to the different "crisis" periods both the assets are ex-

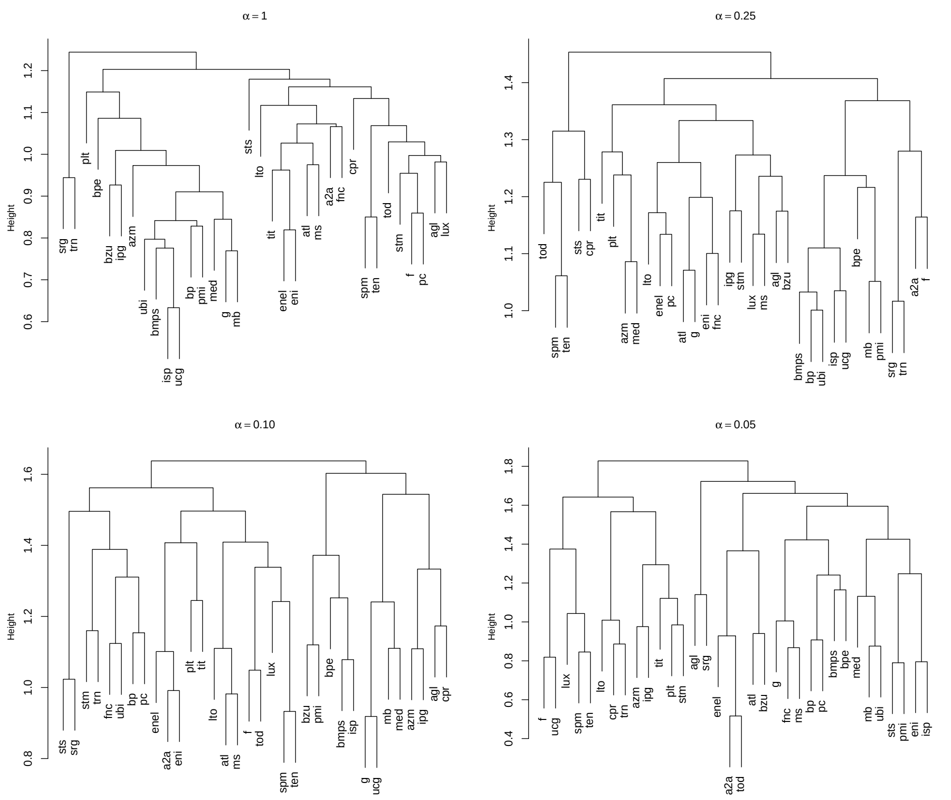


Figure 5.3: Dendrogram plots of 34 FTSE–MIB components obtained by using hierarchical clustering (`hclust` function) with complete linkage method in scenario $\alpha = 1$ (unconditional Spearman's correlation) and extreme scenarios $\alpha = 0.25, 0.10, 0.05$.

periencing. For example, looking at the dendrograms in Figure 5.3 one can try to understand how the dependences within a specific sector in the data set (*e.g.*, bank institutions) evolve for different conditioning. The fact that such dependencies change is well known in the literature, often under the name of financial contagion (Durante & Foscolo, 2013). In particular, the clusters obtained in the unconditional case are quite different from the clusters obtained in a risky scenario. This is particularly important for possible consequences in portfolio management. In fact, in crisis periods, diversification effects can be mitigated when one does not take into account the "extreme" dependence between the different assets.

Finally, a remark is needed here. From a practical point of view, first the risk manager should fix the threshold level α, which reflects her/his investment strategy, then the cluster procedure should be performed. If one wants to be conservative against extreme risks, a threshold $\alpha = 0.05$ could be a reasonable choice, since it reflects the behaviour of the markets in very risky scenarios. It will be a matter of future investigations the determination of algorithmic procedures in order to select the optimal α for a given set of time series. To this end some proposals have been formulated in Jaworski & Pitera (2013).

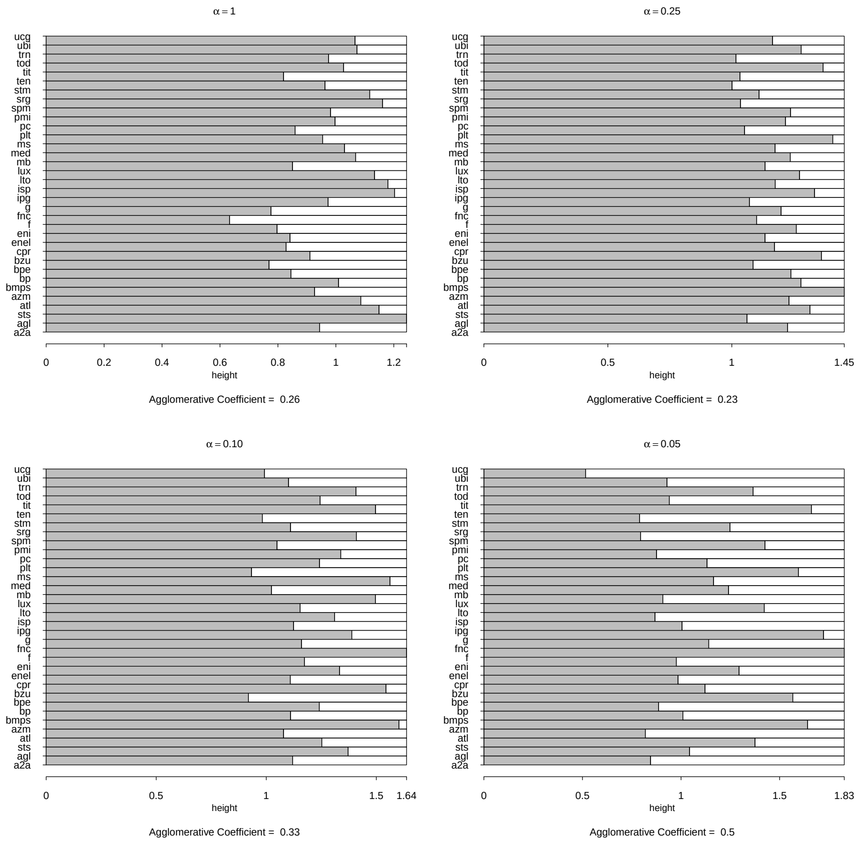


Figure 5.4: Barplots visualizing the agglomerative hierarchical clustering (complete linkage) of 34 FTSE–MIB components for $\alpha \in \{1, 0.25, 0.1, 0.05\}$, from `agnes` algorithm (Kaufman & Rousseeuw, 1990). Agglomerative coefficients measure the amount of clustering structure found.

5.7 Discussion

In this Chapter, a new methodology to cluster financial time series according to a suitable measure of tail dependence has been proposed. The motivation underlying our approach can be found in the well-recognized fact that "classical correlation measures do not give an accurate indication of the real dependence between risk exposures".

We focus on the behaviour of the markets when they are exposed to losses that are judged to be extremely critical according to a predefined threshold level. To do this, we measure the strength of the (positive) association between the time series in a given tail region of their domain by considering suitable pairwise conditional Spearman's correlation coefficients extracted from the original series. The procedure does not work directly with the time series data but requires a preliminary filtering of the univariate time series, by means of a GARCH-type model. As discussed in Section 4.3 of Chapter 4, a copula-based approach allows us to separate the univariate behaviour of each time series from the dependence structure at a multivariate level.

The second step of the procedure is based on the construction of a dissimilarity matrix, collecting all the estimated conditional correlations, which only depend on the associated threshold copula (Section 5.2). These dissimilarities are then used to perform traditional cluster analysis and, in particular, hierarchical clustering. It is worth mentioning that the methodology we suggest does not require the specification of a full (parametric) dependency model between the markets under consideration, but it is essentially non-parametric, thus allowing for a much greater degree of flexibility in modelling the data. Finally, a simulation study (giving overall good results) and an application to the assets from the Italian FTSE-MIB index are presented (Sections 5.5, 5.6).

The results could provide possible evidence that the correlations between markets during stress periods are quite different than under normal conditions, with important consequences in portfolio management problems. In particular, the proposed clustering strategy can be useful to investigate extreme comovements between financial time series in such a way that may be helpful to portfolio selection.

Table 5.1: Simulation study results for $d = 32$, $\alpha = 0.25$.

		$T = 400$		
J	Copula family	λ	RI	ARI
4	Clayton	0.10	0.7624	0.2825
		0.30	0.8506	0.6761
		0.50	0.9597	0.9108
	Survival Gumbel	0.10	0.9155	0.7877
		0.30	0.9984	0.9954
		0.50	0.9998	0.9993
8	Clayton	0.10	0.8235	0.1782
		0.30	0.8841	0.4365
		0.50	0.9581	0.7960
	Survival Gumbel	0.10	0.9080	0.5387
		0.30	0.9949	0.9729
		0.50	0.9995	0.9974
		$T = 800$		
J	Copula family	λ	RI	ARI
4	Clayton	0.10	0.8287	0.5412
		0.30	0.9696	0.9187
		0.50	0.9988	0.9966
	Survival Gumbel	0.10	0.9917	0.9807
		0.30	1.0000	1.0000
		0.50	1.0000	1.0000
8	Clayton	0.10	0.8573	0.3205
		0.30	0.9571	0.7786
		0.50	0.9982	0.9900
	Survival Gumbel	0.10	0.9726	0.8545
		0.30	0.9999	0.9997
		0.50	1.0000	1.0000

Table 5.2: Simulation study results for $d = 64$, $\alpha = 0.25$.

J	Copula family	λ	RI	ARI
		$T = 400$		
4	Clayton	0.10	0.7678	0.1689
		0.30	0.8193	0.6194
		0.50	0.9598	0.9030
	Survival Gumbel	0.10	0.9441	0.8534
		0.30	0.9990	0.9973
		0.50	0.9996	0.9990
8	Clayton	0.10	0.8671	0.1385
		0.30	0.9068	0.4395
		0.50	0.9669	0.8306
	Survival Gumbel	0.10	0.9435	0.6770
		0.30	0.9985	0.9921
		0.50	0.9998	0.9989
16	Clayton	0.10	0.9107	0.0842
		0.30	0.9277	0.2660
		0.50	0.9659	0.6563
	Survival Gumbel	0.10	0.9423	0.4071
		0.30	0.9956	0.9531
		0.50	0.9991	0.9908
		$T = 800$		
J	Copula family	λ	RI	ARI
4	Clayton	0.10	0.8151	0.4917
		0.30	0.9716	0.9270
		0.50	0.9986	0.9961
	Survival Gumbel	0.10	0.9941	0.9838
		0.30	1.0000	1.0000
		0.50	1.0000	1.0000
8	Clayton	0.10	0.8951	0.3328
		0.30	0.9691	0.8429
		0.50	0.9986	0.9931
	Survival Gumbel	0.10	0.9880	0.9396
		0.30	1.0000	1.0000
		0.50	1.0000	1.0000
16	Clayton	0.10	0.9215	0.1979
		0.30	0.9659	0.6504
		0.50	0.9981	0.9799
	Survival Gumbel	0.10	0.9776	0.7653
		0.30	0.9999	0.9994
		0.50	1.0000	1.0000

Table 5.3: Simulation study results for $d = 128$, $\alpha = 0.25$.

$T = 400$				
J	Copula family	λ	RI	ARI
4	Clayton	0.10	0.6821	0.3204
		0.30	0.7826	0.6007
		0.50	0.9458	0.8687
	Survival Gumbel	0.10	0.9656	0.9085
		0.30	0.9997	0.9990
		0.50	0.9997	0.9992
8	Clayton	0.10	0.8301	0.1446
		0.30	0.8762	0.4278
		0.50	0.9583	0.8025
	Survival Gumbel	0.10	0.9533	0.7663
		0.30	0.9993	0.9966
		0.50	0.9997	0.9987
16	Clayton	0.10	0.8883	0.0796
		0.30	0.9077	0.2500
		0.50	0.9593	0.6428
	Survival Gumbel	0.10	0.9396	0.4586
		0.30	0.9977	0.9779
		0.50	0.9996	0.9959
$T = 800$				
J	Copula family	λ	RI	ARI
4	Clayton	0.10	0.7661	0.5644
		0.30	0.9677	0.9139
		0.50	0.9988	0.9968
	Survival Gumbel	0.10	0.9978	0.9939
		0.30	0.9996	0.9937
		0.50	1.0000	1.0000
8	Clayton	0.10	0.8756	0.3441
		0.30	0.9681	0.8456
		0.50	0.9980	0.9906
	Survival Gumbel	0.10	0.9946	0.9739
		0.30	0.9870	0.8750
		0.50	1.0000	1.0000
16	Clayton	0.10	0.8978	0.1789
		0.30	0.9630	0.6676
		0.50	0.9978	0.9791
	Survival Gumbel	0.10	0.9857	0.8675
		0.30	1.0000	0.9999
		0.50	1.0000	1.0000

Table 5.4: List of the analysed components of FTSE–MIB.

Symbol	Name
a2a	A2A
agl	AUTOGRILL
sts	ANSALDO STS
atl	ATLANTIA
azm	AZIMUT HOLDING
bmps	BANCA MPS
bp	BANCO POPOLARE
bpe	BCA POP. EMILIA R.
bzu	BUZZI UNICEM
cpr	CAMPARI
enel	ENEL
eni	ENI
f	FIAT
fnc	FINMECCANICA
g	GENERALI
ipg	IMPREGILO
isp	INTESA SANPAOLO
lto	LOTTOMATICA
lux	LUXOTTICA GROUP
mb	MEDIOBANCA
med	MEDIOLANUM
ms	MEDIASET
plt	PARMALAT
pc	PIRELLI&C
pmi	BCA POP. MILANO
spm	SAIPEM
srg	SNAM
stm	STMICROELECTRONICS
ten	TENARIS
tit	TELECOM ITALIA
tod	TOD'S
trn	TERNA
ubi	UBI BANCA
ucg	UNICREDIT

Table 5.5: Diagnostic tests for the components of FTSE–MIB.

	BP (1)	BP (5)	LB (1)	LB (5)	ARCH (1)	ARCH (5)	KS
a2a	0.90	0.37	0.90	0.37	0.38	0.89	0.50
agl	0.82	0.39	0.82	0.38	0.88	0.94	0.41
sts	0.94	0.20	0.94	0.20	0.73	0.92	0.72
atl	0.76	0.86	0.76	0.85	0.21	0.36	0.78
azm	0.58	0.66	0.58	0.66	0.84	0.64	0.18
bmps	0.32	0.71	0.32	0.70	0.91	0.94	0.46
bp	0.79	0.88	0.79	0.88	0.61	0.03	0.57
bpe	0.21	0.65	0.21	0.64	0.02	0.01	0.25
bzu	0.80	0.49	0.80	0.49	0.55	0.10	0.05
cpr	0.74	0.76	0.74	0.76	0.78	0.99	0.94
enel	0.93	0.94	0.93	0.94	0.30	0.81	0.11
eni	0.89	0.69	0.89	0.69	0.12	0.73	0.03
f	0.80	0.25	0.80	0.25	0.95	0.90	0.04
fnc	0.36	0.10	0.36	0.09	0.39	0.80	0.51
g	0.59	0.78	0.59	0.78	0.44	0.68	0.14
ipg	0.32	0.17	0.32	0.16	0.93	0.95	0.49
isp	0.55	0.24	0.55	0.24	0.02	0.10	0.13
lto	0.68	0.63	0.68	0.63	0.54	0.73	0.79
lux	0.84	0.00	0.84	0.00	0.40	0.75	0.43
mb	0.77	0.95	0.77	0.95	0.95	0.92	0.84
med	0.23	0.62	0.23	0.62	0.79	0.60	0.58
ms	0.65	0.90	0.65	0.90	0.76	0.95	0.52
plt	0.36	0.80	0.36	0.80	0.64	0.87	0.25
pc	0.95	0.97	0.95	0.97	0.70	0.91	0.23
pmi	0.44	0.89	0.44	0.89	0.85	1.00	0.47
spm	0.59	0.84	0.59	0.83	0.21	0.14	0.08
srg	0.45	0.36	0.45	0.36	0.45	0.92	0.37
stm	0.64	0.58	0.64	0.57	0.81	0.51	0.50
ten	0.73	0.93	0.73	0.93	0.45	0.39	0.34
tit	0.89	0.09	0.89	0.08	0.54	0.99	0.13
tod	0.60	0.44	0.60	0.44	0.96	0.72	0.77
trn	0.72	0.94	0.72	0.94	0.67	0.95	0.20
ubi	0.47	0.76	0.47	0.75	0.34	0.81	0.54
ucg	0.34	0.27	0.34	0.27	0.57	0.62	0.08

Chapter 6

Clustering financial time series via tail dependence coefficient

6.1 Introduction

In this chapter a procedure for clustering financial time series according to a suitable copula-based tail coefficient is proposed.

A recent work by De Luca & Zuccolotto (2011) proposes a dissimilarity measure for financial time series clustering based on parametric estimation of pairwise lower tail dependence coefficients. Developing the ideas presented there, we follow an alternative approach which avoids to specify any model assumption on the pairwise dependence structure of the involved time series since it is only based on the rank statistics derived from the observations. Moreover, it is also shown that, while a multidimensional scaling is suggested by De Luca & Zuccolotto (2011) as a further transformation of dissimilarities, this step could be avoided without deteriorating the overall results.

The chapter is organized as follows. Sections 6.2 and 6.3 describe the proposed cluster algorithm, whose performance is checked via a simulation study in Section 6.4. An application to the analysis of MSCI Developed Market indices is given in Section 6.5, allowing a direct comparison with the results of De Luca & Zuccolotto (2011). In Section 6.6 a two-stage portfolio selection procedure is developed and empirical calculations on the EURO STOXX 50 are provided. In particular, by exploiting the proposed tail dependence-based risky measures, a first-step cluster analysis is carried out for discerning between assets that behave similarly during risky scenarios; while the second step concerns the selection of a weighted portfolio from a group of assets in such a way that the assets are diversified in their tail behaviour. Section 6.7 concludes.

6.2 Non-parametric estimation of tail dependence

Once we have obtained the pseudo–observations $(z_t^1, \ldots, z_t^d)_{t=1,\ldots,T}$ from the original time series according to procedure in Chapter 4, Section 4.4, we adopt the concept of tail dependence coefficients. We recall that, as discussed in Chapter 3, Section 3.2, if (X, Y) is a continuous bivariate random vector with copula C, then the lower and upper tail dependence coefficients (shortly, TDC's) only depend on C and are defined, respectively, by

$$\lambda_L(C) = \lim_{t \to 0^+} \frac{C(t,t)}{t} \quad \text{and} \quad \lambda_U(C) = \lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1-t}. \tag{6.1}$$

For what follows, it is important to notice that

$$\lambda_L(C) = \lambda_U(\widehat{C}), \tag{6.2}$$

where $\widehat{C}$ is the survival copula associated with C and given by

$$\widehat{C}(u, v) = u + v - 1 + C(1-u, 1-v).$$

Estimators of tail dependence coefficients have been considered several times in the literature (see, *e.g.*, Frahm et al., 2005). In particular, they are popular in the class of extreme value copulas (see, for instance, Beirlant et al., 2006; Gudendorf & Segers, 2010; Salvadori et al., 2007).

We recall that a copula C is called an *extreme value copula* (EVC) if $C(u^t, v^t) = C^t(u, v)$ for all $t > 0$, $u, v \in [0, 1]$. A result of Pickands (1981) states that C is an EVC if and only if

$$C(u, v) = (uv)^{A\left(\frac{\log v}{\log(uv)}\right)}, \ (u, v) \in [0, 1]^2, \tag{6.3}$$

where $A : [0, 1] \to [1/2, 1]$ is continuous, convex and satisfies the constraint $\max\{t, 1-t\} \le A(t) \le 1$ for all $t \in [0, 1]$. The function A is referred to as the *dependence function* associated with C.

Non-parametric estimation procedures of the dependence function A have been extensively considered in the literature (see, for instance, Gudendorf & Segers, 2010). In particular, if C is an EVC, then

$$\lambda_U(C) = 2 - 2A\left(\frac{1}{2}\right). \tag{6.4}$$

In other words, the estimation of the dependence function A provides an estimation for the upper TDC. Among various possible choices, a good choice is given by the estimator $\widehat{A}^{CFG}$ proposed by Capéraà et al. (1997) (and further studied by Genest & Segers, 2009), due to the fact that it seems preferable to other similar estimators (see the discussion by Genest & Segers, 2009). In the sequel, we will denote by $\widehat{\lambda}_U^{CFG}$ the estimator of the upper TDC obtained from the estimation of the dependence function via formula (6.4).

Now, in order to use an *Extreme Value Theory* (EVT) approach for the estimation of lower TDC of our time series, we adopt the procedure suggested by Frahm et al. (2005, section 3.5).

Let C be the copula associated with the pseudo–observations $(z_t^1, \ldots, z_t^d)_{t=1,\ldots,T}$ of the considered financial returns, obtained by applying the empirical distribution function to the estimated standardized residuals from the univariate models. Let $\widehat{C}$ be the copula associated with the pseudo–observations of the corresponding losses (*i.e.* the opposite of the returns) given by $\widetilde{z}_t^i = 1 - z_t^i$ for every $i = 1, \ldots, d$ and $t = 1, \ldots, T$. As a matter of fact, $\widehat{C}$ may not be an EVC. However, under suitable conditions, it belongs to the so–called *domain of attraction* of an EVC C^* (Gudendorf & Segers, 2010). Moreover, it has been proved that $\widehat{C}$ and C^* have the same upper TDC (Abdous et al., 1999, Lemma 1). Thus, instead of estimating directly the lower TDC from C (or, equivalently, the upper TDC from $\widehat{C}$), we may estimate it by using the estimator $\widehat{\lambda}_U^{CFG}$ applied to the EVC C^*. Obviously, C^* is unknown, but its empirical version can be obtained by extracting block maxima from the loss observations.

Specifically, the pairwise lower TDC is calculated via the following procedure.

1. Given a pair of pseudo–observations $(z_t^i, z_t^j)_{t=1,\ldots,T}$, $i \neq j$, set $\widetilde{z}_t^i = 1 - z_t^i$ and $\widetilde{z}_t^j = 1 - z_t^j$. Namely, we pass from the copula C of the pseudo-observations to the survival copula $\widehat{C}$ of the pseudo–observations
2. We extract from $(\widetilde{z}_t^i, \widetilde{z}_t^j)_{t=1,\ldots,T}$ the maxima of observations over m blocks

of $l = T/m$ elements, by obtaining the time series $(\widetilde{M}_t^i, \widetilde{M}_t^j)_{t=1,\dots,m}$.

3. We estimate the bivariate dependence function A_m^{ij} for the m block maxima via the non-parametric rank-based CFG estimator by Capéraà et al. (1997) and get the estimation $\widehat{\lambda}_U^{CFG}$ of the upper TDC in (6.4).
4. By using (6.2), $\widehat{\lambda}_U^{CFG}$ coincides with the lower TDC of $(z_t^i, z_t^j)_{t=1,\dots,T}$.

Notice that, as usual in the block-maxima approach (see, for instance, Embrechts et al., 1997), a trade-off necessarily takes place in determining the number and size of blocks: a larger size leads to a more accurate determination of the EVC C^* in the domain of attraction; while a large number of blocks gives more data for the estimation of the dependence function A.

6.3 Apply a cluster algorithm on dissimilarities

The definition of a suitable dissimilarity function between each pair of time series is a fundamental step before performing a cluster analysis. As in De Luca & Zuccolotto (2011), here we have to transform the estimated lower TDC's through a monotonic function in such a way that the obtained dissimilarity between two time series is small when their tail dependence is high, and monotonically increases when their tail dependence decreases. Thus, for $i, j = 1, \dots, d$, a matrix $\Delta = (\Delta_{ij})$ is defined whose elements are given by

$$\Delta_{ij} = -\log\left(\hat{\lambda}_{ij}^L\right), \tag{6.5}$$

where $\hat{\lambda}_{ij}^L$ is the lower tail dependence coefficient between time series i and j estimated non-parametrically through the procedure described in the previous section. Equation (6.5) defines a dissimilarity matrix, which can be used as input for hierarchical clustering algorithms. When, instead, partitioning methods are used, such a matrix has to be further treated in order to obtain a corresponding distance matrix. Notice that here we choose a different function to derive the matrix Δ compared to (5.2). Here, $\hat{\lambda}_{ij}^L \in [0, 1]$, implying that (6.5) ranges from 0 to infinity.

Starting from the dissimilarity matrix defined in (6.5) we can perform cluster analysis of the time series by following two different approaches:

1. apply an agglomerative hierarchical algorithm (*e.g.*, single, average, complete linkage) directly to the matrix $\Delta = (\Delta_{ij})$;

2. perform a non-metric Multidimensional Scaling (MDS) in order to obtain the representation $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$ of normalized residuals in $\mathbb{R}^q$. Then, the q-dimensional point configuration obtained can be used as an input for K-means algorithm.

The first clustering process is based on the hierarchical classification of the objects, producing the dendrogram that shows how clusters are formed at each stage of the procedure. We recall that, in hierarchical clustering, partitions are obtained by cutting off the dendrogram at an arbitrary point. One advantage of hierarchical clustering is that the number of clusters is not required as a parameter. In the second approach we consider the points in $\mathbb{R}^q$ obtained from MDS and cluster them with the classical K-means algorithm. As already explained in Chapter 3, Section 3.5, the algorithm works for a given value of the dimension q, which has to be given in input. Starting from an initial point configuration for $q = 2$, the Shepard-Kruskal algorithm (Kruskal, 1964a,b) iteratively improves the points configuration so as to minimize the stress function and have a good approximation of the original entries. The outcome is then used to perform K-means algorithm, a partitioning method in which an iterative algorithm minimizes the sum of distances from each object to its cluster centroid, over all clusters. Cluster centroids are computed differently for each distance measure, to minimize the sum with respect to the chosen measure. It is a faster method than hierarchical clustering, but the number of clusters has to be fixed in advance. The final output is a set of clusters that are as compact and well separated as possible.

6.4 Simulation study

A simulation study is conducted to check the clustering performances of the proposed methodology. In analogy with the model defined in (5.3), for all $\boldsymbol{u} = (u_{11}, \ldots, u_{1k_1}, u_{21}, \ldots, u_{2k_2}, \ldots, u_{J1}, \ldots, u_{Jk_J})$, with $k_1 + \cdots + k_J = d$, we consider the following d–dimensional copula model

$$C(\boldsymbol{u}) = \prod_{j=1}^{J} C_j(u_{j1}, \ldots, u_{jk_j}). \tag{6.6}$$

Again, C is the product of J copulas from Clayton and survival Gumbel family with a fixed lower TDC λ. Model (6.6) is quite convenient in this context since it gives a direct overview of the degree of tail dependence within and across subgroups. We generate T data from (6.6) and consider different scenarios according to:

- the sample size $N = 500, 1000$;
- the dimension of the model $d = 32, 64, 128$;
- the number of different clusters $J = 4, 8, 16$, with $J < d/2$;
- the lower TDC $\lambda = 0.25, 0.50, 0.75$.

For each model, the dissimilarity matrix (6.5) is computed and two clustering procedures (hierarchical and non-hierarchical) are applied:

1. complete linkage algorithm on dissimilarity matrix $\Delta_{ij} = -\log\left(\hat{\lambda}^L_{ij}\right)$;
2. K-means partitioning algorithm on the points in the final configuration returned by Shepard-Kruskal's non-metric MDS.

When performing MDS, we set $q = 2$ in order to avoid computational burden provided by considering $q > 2$. Notice, however, that under this restriction we may in any case obtain that the stress function is lower than 2.5%. Now, supposed that the number of clusters k to be selected is fixed and equal to J for each simulation, we calculate the RI and ARI between the obtained cluster structure (from the sampled data) and the expected cluster structure (as derived from the chosen model). The calculations are repeated 250 times, and the average index is considered. The results are reported in Tables 6.1 – 6.6. As the sample size increases, the performance generally improves, as expected, regardless of copula model and number of clusters. Moreover, as the dimension d increases, the performance seems to decrease, even if the changes are not so evident. In general, an increasing number of clusters seems to give better performances. The different dependence structure (in terms of TDC) matters; in fact, a stronger cluster separation (as obtained by a larger TDC) increases the performances. The output is also influenced by the choice of the copula family. This is due to a different tail behaviour that cannot only be captured by the TDC, as also explained in Chapter 3 about tail concentration function (for more considerations about the tail of a copula see Jaworski, 2010). Finally, complete linkage clustering procedure outperforms K-means, which requires an additional step (multidimensional scaling) to convert the dissimilarity matrix into a distance matrix. We also consider the approach via complete linkage when the

Table 6.1: Simulation study results for $N = 500$, $d = 32$, $k = J$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.8072	0.4845	0.7734	0.3715
		0.50	0.9733	0.9264	0.8662	0.6401
		0.75	1.0000	1.0000	0.8766	0.6802
	Survival Gumbel	0.25	0.7815	0.4216	0.7637	0.3411
		0.50	0.9729	0.9253	0.8604	0.6213
		0.75	0.9998	0.9993	0.8718	0.6736
8	Clayton	0.25	0.8792	0.3596	0.8448	0.1610
		0.50	0.9690	0.8335	0.8888	0.4117
		0.75	0.9997	0.9983	0.9306	0.6491
	Survival Gumbel	0.25	0.8710	0.3210	0.8392	0.1331
		0.50	0.9714	0.8463	0.8949	0.4429
		0.75	0.9999	0.9997	0.9244	0.6204

number of cluster is not fixed, but determined by the silhouette index criterion, as previously done in Section 5.5. These results are reported in Tables 6.7 and 6.8 (similar results are obtained with the remaining copula models and are, for this reason, not reported). As can be seen, the performance remains overall good and comparable with the results obtained in the case the number of clusters is fixed.

6.5 Illustration from MSCI Developed Markets Index

In order to illustrate our approach we analyse daily returns of time series of Morgan Stanley Capital International (MSCI) Developed Markets indices designed to measure the equity market performance of developed markets. The Dataset includes the following markets: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the United Kingdom and the United States. We restrict to the time series of daily log–returns $(x_t^1, \ldots, x_t^d)$, $d = 23, t = 1, \ldots, T$, in the period from June 4, 2002 to June 10, 2010 ($T = 2093$ observations; Source: Datastream) in order to provide a direct comparison with the results by De Luca & Zuccolotto (2011). We preliminary apply a univariate Student-t AR-GARCH model to each time series of returns and

Table 6.2: Simulation study results for $N = 1000$, $d = 32$, $k = J$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.9206	0.7857	0.8670	0.6298
		0.50	0.9998	0.9993	0.9176	0.7829
		0.75	1.0000	1.0000	0.9028	0.7471
	Survival Gumbel	0.25	0.8894	0.7054	0.8328	0.5318
		0.50	0.9987	0.9965	0.9117	0.7695
		0.75	1.0000	1.0000	0.9155	0.7822
8	Clayton	0.25	0.9325	0.6384	0.8531	0.2056
		0.50	0.9985	0.9919	0.9079	0.5079
		0.75	1.0000	1.0000	0.9422	0.7052
	Survival Gumbel	0.25	0.9146	0.5445	0.8488	0.1810
		0.50	0.9991	0.9951	0.9065	0.5022
		0.75	1.0000	1.0000	0.9421	0.7087

Table 6.3: Simulation study results for $N = 500$, $d = 64$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.7576	0.4263	0.7873	0.4220
		0.50	0.9710	0.9236	0.8838	0.6925
		0.75	0.9997	0.9992	0.8804	0.7005
	Survival Gumbel	0.25	0.7293	0.3752	0.7636	0.3574
		0.50	0.9750	0.9348	0.8736	0.6660
		0.75	1.0000	1.0000	0.8599	0.6430
8	Clayton	0.25	0.8660	0.3662	0.8271	0.1555
		0.50	0.9712	0.8617	0.8749	0.3972
		0.75	0.9993	0.9966	0.9225	0.6429
	Survival Gumbel	0.25	0.8580	0.3313	0.8238	0.1381
		0.50	0.9721	0.8669	0.8711	0.3791
		0.75	0.9997	0.9984	0.9233	0.6455
16	Clayton	0.25	0.9278	0.2653	0.9087	0.0759
		0.50	0.9777	0.7695	0.9191	0.1828
		0.75	0.9996	0.9963	0.9406	0.4134
	Survival Gumbel	0.25	0.9236	0.2229	0.9076	0.0619
		0.50	0.9761	0.7544	0.9180	0.1774
		0.75	0.9994	0.9940	0.9412	0.4199

Table 6.4: Simulation study results for $N = 1000$, $d = 64$, $k = J$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.8786	0.7084	0.8709	0.6508
		0.50	0.9991	0.9975	0.9187	0.7932
		0.75	1.0000	1.0000	0.9055	0.7621
	Survival Gumbel	0.25	0.8424	0.6297	0.8459	0.5823
		0.50	0.9995	0.9985	0.9297	0.8207
		0.75	1.0000	1.0000	0.9064	0.7640
8	Clayton	0.25	0.9301	0.6708	0.8372	0.2047
		0.50	0.9990	0.9950	0.8839	0.4394
		0.75	1.0000	1.0000	0.9337	0.6934
	Survival Gumbel	0.25	0.9096	0.5840	0.8320	0.1780
		0.50	0.9992	0.9962	0.8831	0.4359
		0.75	1.0000	1.0000	0.9358	0.7040
16	Clayton	0.25	0.9533	0.5199	0.9100	0.0884
		0.50	0.9982	0.9811	0.9211	0.2014
		0.75	1.0000	1.0000	0.9427	0.4360
	Survival Gumbel	0.25	0.9468	0.4541	0.9092	0.0775
		0.50	0.9984	0.9830	0.9213	0.2032
		0.75	1.0000	1.0000	0.9414	0.4201

Table 6.5: Simulation study results for $N = 500$, $d = 128$, $k = J$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.6472	0.2760	0.7769	0.4006
		0.50	0.9694	0.9219	0.8714	0.6620
		0.75	1.0000	1.0000	0.8788	0.6937
	Survival Gumbel	0.25	0.5927	0.2095	0.7622	0.3603
		0.50	0.9691	0.9233	0.8738	0.6670
		0.75	1.0000	0.9999	0.8779	0.6918
8	Clayton	0.25	0.8026	0.2880	0.8252	0.1802
		0.50	0.9688	0.8583	0.8612	0.3554
		0.75	0.9994	0.9971	0.9116	0.6048
	Survival Gumbel	0.25	0.8343	0.2807	0.8195	0.1688
		0.50	0.9721	0.8731	0.8586	0.3422
		0.75	1.0000	0.9999	0.9110	0.6023
16	Clayton	0.25	0.9183	0.2669	0.8996	0.0859
		0.50	0.9754	0.7770	0.9108	0.1935
		0.75	0.9993	0.9938	0.9329	0.4093
	Survival Gumbel	0.25	0.8930	0.0835	0.8900	0.0496
		0.50	0.9753	0.7780	0.9109	0.1973
		0.75	0.9994	0.9941	0.9336	0.4143

Table 6.6: Simulation study results for $N = 1000$, $d = 128$, $k = J$.

			Complete linkage		K-means	
J	Copula family	λ	RI	ARI	RI	ARI
4	Clayton	0.25	0.7998	0.5585	0.8461	0.5865
		0.50	0.9991	0.9977	0.9210	0.7961
		0.75	1.0000	1.0000	0.8877	0.7204
	Survival Gumbel	0.25	0.7667	0.5109	0.8313	0.5474
		0.50	0.9991	0.9977	0.9257	0.8088
		0.75	1.0000	1.0000	0.9005	0.7516
8	Clayton	0.25	0.8943	0.5954	0.8375	0.2367
		0.50	0.9993	0.9966	0.8733	0.4124
		0.75	1.0000	1.0000	0.9207	0.6442
	Survival Gumbel	0.25	0.8621	0.4180	0.8413	0.2813
		0.50	0.9991	0.9960	0.8727	0.4095
		0.75	1.0000	1.0000	0.9224	0.6522
16	Clayton	0.25	0.9514	0.5621	0.9020	0.1086
		0.50	0.9985	0.9857	0.9149	0.2313
		0.75	1.0000	1.0000	0.9377	0.4501
	Survival Gumbel	0.25	0.8961	0.1224	0.8936	0.0881
		0.50	0.9988	0.9886	0.9146	0.2280
		0.75	1.0000	1.0000	0.9378	0.4505

Table 6.7: Simulation study results for $N = 500$, $d = 32$, k not fixed

			Complete linkage	
J	Copula family	λ	RI	ARI
4	Clayton	0.25	0.7648	0.5637
		0.50	0.8699	0.8439
		0.75	0.9406	0.9715
	Survival Gumbel	0.25	0.7484	0.5341
		0.50	0.8757	0.8521
		0.75	0.9272	0.9512
8	Clayton	0.25	0.7728	0.4541
		0.50	0.8959	0.7908
		0.75	0.9831	0.9814
	Survival Gumbel	0.25	0.7582	0.4187
		0.50	0.8942	0.7898
		0.75	0.9648	0.9696

the standardized residuals are computed. For all time series we then perform Box-Pierce and Ljung-Box tests at lags 1 and 5, to check for residual autocorrelation, ARCH tests at lags 1 and 5, for autoregressive conditional heteroscedasticity and Kolmogorov-Smirnov test to check for the Student hypothesis for the standardized residuals. The estimation results show a reasonable fit for all time series (see Table 6.9). The standardized residuals from each time series are rescaled to the interval $[0, 1]$ thus obtaining the sample $(z_t^1, \ldots, z_t^{23})$ on $[0, 1]^{23}$ which represents the empirical copula among the time series of returns. For the estimation of TDC, we consider $m = 91$ block maxima where each block contains $2093/91 = 23$ elements from each time series of residuals (*i.e.* we focus approximately to monthly maxima). Then, the pairwise lower TDC's λ_{ij}^L are estimated non-parametrically by the procedure described in Section 6.2. The total number of estimated coefficients is $d(d-1)/2 = 253$, resulting in a 23×23 symmetric matrix. The starting point for our clustering procedure is the dissimilarity matrix (6.5) based on estimates $\hat{\lambda}_{ij}^L$ so that two strongly tail dependent assets are grouped together and weakly tail dependent assets are far away. Among hierarchical clustering techniques the complete linkage method is used to achieve more useful hierarchies than single or average linkage from a pragmatic point of view. Moreover, it can be used on data that are not restricted to Euclidean distances. Looking at the dendrogram produced by

Table 6.8: Simulation study results for $N = 1000$, $d = 32$, k not fixed.

			Complete linkage	
J	Copula family	λ	RI	ARI
4	Clayton	0.25	0.8953	0.7752
		0.50	0.9390	0.9237
		0.75	0.9597	0.9769
	Survival Gumbel	0.25	0.8484	0.7221
		0.50	0.9473	0.9345
		0.75	0.9541	0.9673
8	Clayton	0.25	0.8683	0.6518
		0.50	0.9828	0.9640
		0.75	0.9947	0.9922
	Survival Gumbel	0.25	0.8412	0.5694
		0.50	0.9838	0.9672
		0.75	0.9988	0.9973

complete linkage scheme (Figure 6.1), we find out that the hierarchical structure can be interpreted in terms of geographic proximity: the lower tail dependence tends to be higher within European markets, where the Scandinavian countries are grouped together as well as USA and Canada; Pacific countries tend to be divided in two separate clusters where New Zealand and Australia are joined together as well as Hong Kong, Japan and Singapore. At a first level of merging, $k = 3$ can be considered a good solution, where k denotes the number of clusters selected. Table 6.10 reports the corresponding cluster composition.

The second procedure we apply in our case study is the one summarized by point (2) in Section 6.3. Here we repeat the analysis by increasing q until the minimum stress of the corresponding optimal configuration is lower than 2.5%. The final configuration results in a set of $d = 23$ points of dimension $q = 10$, corresponding to a stress value of 0.0242. Figure 6.2 displays the two-dimensional MDS configuration, characterized by a stress value $\min(s) = 0.2283$. The obtained 10-dimensional points configuration can be used as input for K-means algorithm. As said, unlike hierarchical clustering, K-means clustering requires that the number of clusters to extract be specified in advance. The `NbClust` package of R can be used as a guide (Charrad et al., 2013). Additionally, a plot of the total within-groups sums of squares against the number of clusters can be helpful. A bend in the graph can suggest the appropriate number of clusters. From the left plot in Figure 6.3 we observe that the decreasing profile in the within groups sum of squares when k increases from 4 to 5 seems to be higher than the decreasing profile when k increases from 5 to 6, suggesting that a 5-clusters solution may be a good fit to the data. In the right part of Figure 6.3, 7 of 24 criteria provided by the `NbClust` package suggest a 5-clusters solution. The cluster memberships are listed in Table 6.11. The cluster solutions we carry out from the two procedures can be directly compared with the results obtained by De Luca & Zuccolotto (2011), although in our analysis we perform hierarchical clustering in addition to partitioning clustering and adopt different criteria in the choice of the number of clusters. Moreover, we would like to stress again that our procedure does not require any parametric assumption on the copula linking the pairwise financial assets, which can be considered the main advantage of the proposed method.

Table 6.9: Diagnostic tests for the components of MSCI Data.

	BP (1)	BP (5)	LB (1)	LB (5)	ARCH (1)	ARCH (5)	KS
JAPAN	0.88	0.58	0.88	0.58	0.55	0.17	0.03
USA	0.83	0.60	0.83	0.60	0.01	0.15	0.00
AUSTRALIA	0.28	0.77	0.28	0.77	0.32	0.24	0.00
CANADA	0.50	0.87	0.50	0.86	0.94	0.47	0.00
GERMANY	0.32	0.25	0.32	0.25	0.43	0.10	0.01
UK	0.64	0.10	0.64	0.10	0.42	0.45	0.01
AUSTRIA	0.35	0.81	0.35	0.81	0.64	0.42	0.01
DENMARK	0.54	0.95	0.54	0.95	0.16	0.10	0.02
FRANCE	0.54	0.20	0.54	0.20	0.53	0.35	0.00
GREECE	0.24	0.60	0.24	0.60	0.25	0.35	0.05
ITALY	0.49	0.26	0.49	0.26	0.18	0.03	0.01
NEW.ZEALAND	0.42	0.60	0.42	0.60	0.87	0.03	0.08
NORWAY	0.34	0.66	0.34	0.66	0.25	0.10	0.00
SINGAPORE	0.33	0.45	0.33	0.45	0.05	0.10	0.07
SWEDEN	0.49	0.27	0.49	0.27	0.60	0.34	0.04
SWITZERLAND	0.51	0.12	0.51	0.12	0.90	0.18	0.00
BELGIUM	0.17	0.73	0.17	0.73	0.01	0.18	0.00
FINLAND	0.25	0.33	0.25	0.33	0.88	0.99	0.67
HONG.KONG	0.41	0.23	0.41	0.23	0.47	0.01	0.00
IRELAND	0.79	0.93	0.79	0.93	0.34	0.77	0.13
NETHERLANDS	0.35	0.23	0.35	0.23	0.62	0.18	0.01
PORTUGAL	0.14	0.14	0.14	0.14	0.21	0.07	0.42
SPAIN	0.44	0.50	0.44	0.50	0.96	0.00	0.03

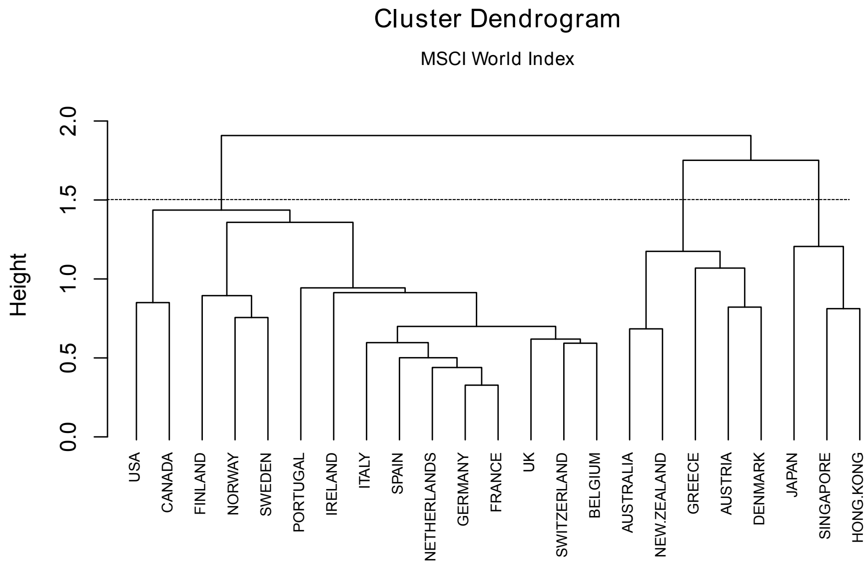


Figure 6.1: Dendrogram of the MSCI World Index Data constituents according to complete linkage clustering. Cutting at height 1.5 a 3-clusters solution is obtained.

Table 6.10: Hierarchical Clustering of MSCI World Index Data.

Cluster 1	Cluster 2	Cluster 3
BELGIUM	AUSTRALIA	HONG KONG
CANADA	AUSTRIA	JAPAN
FINLAND	GREECE	SINGAPORE
FRANCE	DENMARK	
GERMANY	NEW ZEALAND	
IRELAND		
ITALY		
NETHERLANDS		
NORWAY		
PORTUGAL		
SPAIN		
SWEDEN		
SWITZERLAND		
UK		
USA		

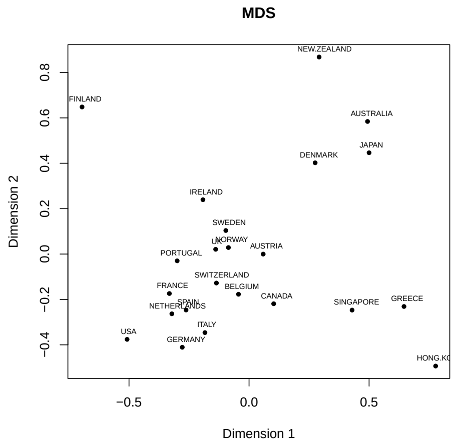


Figure 6.2: Two-dimensional MDS configuration for MSCI World Index Data.

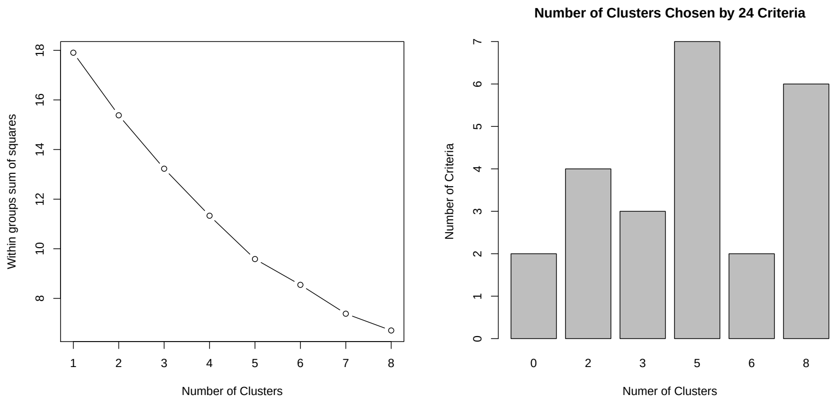


Figure 6.3: *Left* Within groups sum of squares versus the number of clusters in a K-means solution. *Right* Recommended number of clusters using 24 criteria provided by the `NbClust` package.

6.6 Application to portfolio selection framework

By exploiting tail dependence-based risky measures described so far, a two-stage portfolio selection procedure is proposed, with the aim to increase the diversification benefits in a bear market. A first-stage cluster analysis is carried out for discerning between assets with the same performance during risky scenarios. In the second stage, a mean-variance efficient frontier is computed by fixing a number of assets per portfolio and by selecting only one item from each cluster.

Table 6.11: K-means clustering of MSCI World Index Data.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
BELGIUM	AUSTRIA	DENMARK	AUSTRALIA	GREECE
FRANCE	CANADA	FINLAND	JAPAN	HONG.KONG
GERMANY	IRELAND	NORWAY	NEW.ZEALAND	SINGAPORE
ITALY	USA	SWEDEN		
NETHERLANDS				
PORTUGAL				
SPAIN				
SWITZERLAND				
UK				

The idea of diversification by grouping assets is not new (see, *e.g.*, Panton et al., 1976). The further step of selecting assets taking into account group constraints and determining weights via Markowitz's approach has been used, for instance, in Hui (2005), where the groups are determined by factor analysis, and in Cesarone et al. (2013). Finally, De Luca & Zuccolotto (2011) consider all possible portfolios with group constraints; however, again, their clustering procedure is different since it assumes a parametric form of the dependence structure.

As an illustration of our approach, we consider time series related to EURO STOXX 50 stock index and its components in the period from January 2, 2003 to July 31, 2011. Moreover, as out-of-sample period, we will also show the performance of our procedure in the period from August 1, 2011 to September 9, 2011. The period has been selected in consideration of the fact that EURO STOXX 50 was experiencing severe losses. We preliminary apply a univariate Student-t ARMA(1,1)-GARCH(1,1) model to each time series of log–returns of 50 constituents of the index to remove autocorrelation and heteroscedasticity from the data. Then, we compute the standardized residuals in order to check the adequacy of the fit. Having obtained the standardized residuals, we adopt two measures of tail dependence described in Chapter 3 (Sections 3.2 and 3.3):

- the lower tail dependence coefficient λ_L;
- the conditional Spearman's correlation ρ_α, for $\alpha = 0.10$.

As regards the estimation of these quantities we rely on two specific techniques as described in Section and 6.2 and 5.2, respectively:

- the estimate of the lower TDC is derived from the estimate of the extreme-value copula in the domain of attraction of C;
- the estimation of conditional Spearman's ρ_α is related to the calculation of the Spearman's correlation in a sub-sample extracted from the pseudo–observations and dependent on the threshold α.

Both these estimations are obtained via non-parametric procedures and do not require any parametric assumption on the unknown copula linking the time series of interest. Then, the dissimilarity between two time series is defined by $\Delta = (\Delta_{ij})$, for $i, j = 1, \ldots, d$, whose elements are given by

$$\Delta_{ij} = \sqrt{2(1 - \widehat{m}_{ij})}, \tag{6.7}$$

where $\widehat{m}_{ij}$ is the tail dependence measure between time series i and j, that is estimated via one of the two procedures mentioned above. Starting from the dissimilarity matrix defined in (6.7) we can perform a cluster analysis of the time series by different techniques. Here, for a comparative analysis, we focus on two methods: the hierarchical agglomerative algorithms as applied in the previous applications, and in particular, complete linkage, and the *fuzzy clustering algorithm*. The latter is a partitioning method that takes into account some ambiguity in the data, which often occurs, and allows each object to belong to one or more than one cluster according to a membership coefficient. The main advantage of fuzzy clustering is that it yields much more detailed information on the data structure compared to other partitioning techniques. In order to perform fuzzy cluster analysis we can consider FANNY algorithm Kaufman & Rousseeuw (1990), which handles either interval-scaled measurements or dissimilarities. The algorithm aims at minimizing an objective function which is a kind of total dispersion, depending on dissimilarities and membership coefficients. Once the number of clusters is chosen, the algorithm returns some general information on the type of data and the actual memberships for each object in each cluster are listed. Moreover, as in our case, a crisp partition of the financial assets can be determined from the membership value of each time series. In both methods, the optimal number of clusters is chosen by the silhouette index criterion. The results are contained in Tables 6.12–6.15.

As it can be seen, the compositions of the sub–groups seem to differ with respect to both the tail dependence measure and the clustering algorithms. Again, we consider RI (ranging between 0 and 1), and ARI (ranging between -1 and 1), as

Table 6.12: Hierarchical clustering of EURO STOXX 50 based on conditional Spearman's correlation ρ_α with $\alpha = 0.1$

Cluster	Asset							
1	E.IND	F.SGE	D.BASX	D.BAYX	D.RWEX	D.SIEX	D.DTEX	D.SAPX
	H.UNIL	F.LVMH	F.CRFR	I.ISP	F.EI	E.IBE		
2	H.ASML	B.ABI	M.NOK1	D.EONX	F.FTEL	H.MT	F.BSN	F.AIR
	F.OR.F	F.SQ.F						
3	I.ENEL	F.DG.F	D.BMWX	I.ENI	F.TAL	F.UBL		
4	F.BNP	E.REP	H.ING	D.DAIX	E.SCH	F.GOB	E.BBVA	D.ALVX
	D.DBKX	D.MU2X	CRGI	I.G	I.UCG	E.TEF	H.PHIL	F.MIDI
	F.QT.F	F.GSZ						
5	D.VO3X	F.EX.F						

Table 6.13: Hierarchical clustering of EURO STOXX 50 based on lower TDC measure.

Cluster	Asset							
1	E.IND	D.DAIX	D.BMWX					
2	H.ASML	B.ABI	D.RWEX	D.EONX	D.DTEX	D.SAPX	H.UNIL	F.CRFR
	F.EX.F	F.SQ.F	F.QT.F	F.GSZ				
3	I.ENEL	M.NOK1	F.GOB	D.BAYX	F.FTEL	CRGI		
4	F.BNP	E.REP	H.ING	E.SCH	F.SGE	F.DG.F	E.BBVA	D.ALVX
	D.BASX	D.SIEX	H.MT	F.BSN	F.LVMH	F.AIR	I.G	F.OR.F
	I.UCG	I.ISP	E.TEF	F.UBL	F.MIDI	E.IBE		
5	D.DBKX	I.ENI	D.MU2X	D.VO3X	F.TAL	H.PHIL	F.EI	

a measure of agreement among the obtained cluster solutions. If we fix the clustering method and compare the results obtained by changing the tail dependence measure, the obtained grouping compositions seem to be similar, as can be seen by the values reported in Table 6.16, although they do not coincide. In fact, the two tail dependence measures underline different aspects of tail dependence (finite and asymptotic tail behaviour). Analogously, if we fix the tail dependence measure but allow us to use different clustering procedures, the obtained grouping compositions seem to be similar (even more than in the previous case). In other words, the effects of a changing clustering procedure seem to be less evident than those

Table 6.14: Fuzzy clustering of EURO STOXX 50 based on conditional Speaman's correlation ρ_α with $\alpha = 0.1$.

Cluster	Asset							
1	E.IND	E.REP	M.NOK1	H.MT	F.UBL	F.EI		
2	H.ASML	F.BNP	H.ING	E.SCH	E.BBVA	D.BAYX	D.BMWX	D.RWEX
	I.ENI	D.MU2X	D.DTEX					
3	I.ENEL	D.EONX	CRGI	F.BSN	F.CRFR	E.TEF	E.IBE	F.GSZ
4	B.ABI	F.SGE	F.DG.F	D.BASX	F.FTEL	D.VO3X	D.SAPX	H.UNIL
	F.EX.F	F.AIR	F.OR.F	F.SQ.F				
5	D.DAIX	F.GOB	D.ALVX	D.DBKX	D.SIEX	F.TAL	F.LVMH	I.G
	I.UCG	I.ISP	H.PHIL	F.MIDI	F.QT.F			

Table 6.15: Fuzzy clustering of EURO STOXX 50 based on lower TDC measure.

Cluster	Asset							
1	E.IND	F.DG.F	D.BMWX	H.MT	F.LVMH	I.ISP	E.TEF	E.IBE
	F.GSZ							
2	H.ASML	B.ABI	F.SGE	D.EONX	D.DTEX	H.UNIL	F.BSN	F.EX.F
	F.AIR	F.UBL	F.QT.F					
3	I.ENEL	M.NOK1	D.BAYX	I.ENI	F.FTEL	D.SAPX	CRGI	F.TAL
	F.CRFR	H.PHIL	F.EI	F.SQ.F				
4	F.BNP	E.REP	H.ING	E.SCH	F.GOB	E.BBVA	D.ALVX	D.BASX
	D.DBKX	D.SIEX	I.G	F.OR.F	I.UCG	F.MIDI		
5	D.DAIX	D.RWEX	D.MU2X	D.VO3X				

obtained by changing the tail dependence measure.

Once the clustering procedure is completed, the assets have been grouped into a predefined number K of clusters. Then our possible portfolio will be selected on the basis of the following steps.

1. Determine all possible portfolios composed by K assets such that each asset belongs to a different cluster.
2. For these portfolios, calculate the optimal weight assigned to each of its assets with classical Markowitz portfolio selection procedure (Markowitz, 1952). We recall that this procedure provides a general way to maximize

Table 6.16: Rand Index and Adjusted Rand Index between cluster compositions obtained by using ρ_α and the lower TDC from hierarchical clustering (respectively, fuzzy clustering), and between cluster compositions obtained from the two clustering techniques performed on the same tail dependence-based dissimilarity measure.

	Hierarchical Clustering	Fuzzy Clustering	$\rho_{0.10}$	TDC
RI	0.62	0.68	0.68	0.73
ARI	0.03	0.02	0.08	0.26

investor's expected utility under certain conditions, namely to produce portfolios that are able to minimize the total portfolio variance.

3. Given all possible portfolios composed in such a way, plot the graph of their standard deviation against their expected return.
4. Determine the portfolios that are the vertices in the convex efficient frontier of the standard deviation/expected return graph.

According to his/her preference the investor could hence choose one of the portfolios that are on the convex frontier. The proposed approach has the following features:

- It suggests to select the assets of the portfolio by taking into account the grouping structure given by the clustering algorithms. Thus, two assets from the same group (cluster) cannot be included in the same portfolio.
- Once the assets have been selected, their weights are determined by classical methods, like Markowitz approach.
- All the portfolios composed in the previous two steps provide a graphical representation of the possible choices of the investor (see, for instance, Figure 6.4). Based on his/her information, one investment strategy could be selected.
- If no preference is required by the investor, the point with the smallest risk on the convex frontier, namely the global minimum variance portfolio, can be chosen.

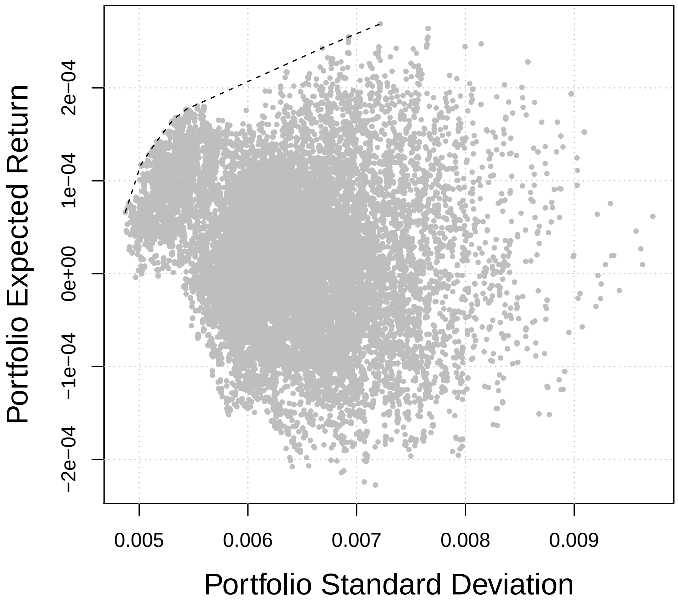


Figure 6.4: Standard deviation-Expected return plot of 5-asset portfolios generated from TDC and hierarchical clustering.

Now, if we restrict to the TDC-based cluster analysis (the other results are quite similar), we may notice in Figure 6.4 the graph of standard deviation against expected returns of all 33264 portfolios composed with our procedure by using hierarchical clustering, while the same picture is obtained in Figure 6.5 by fuzzy clustering. In both cases, we highlight the portfolios in the convex efficient frontier.

The returns of these portfolios in the frontier are compared with the returns of naive minimum variance portfolio built from the whole set of assets and to the benchmark index EURO STOXX 50. As can be seen, the performance of the portfolios in the efficient frontier is generally better than the benchmark and, in several cases, outperforms the global minimum variance portfolio. This seems to confirm the idea that, when markets are experiencing a period of losses, a diversification strategy could be beneficial. The composition of the minimum variance portfolio in the convex frontiers of Figure 6.6 is reported in Table 6.17. Notice that both gave large weight to one single asset, while the other are different. Anyway, as can be read from the basic statistics of the selected portfolios (Table 6.18), the returns of

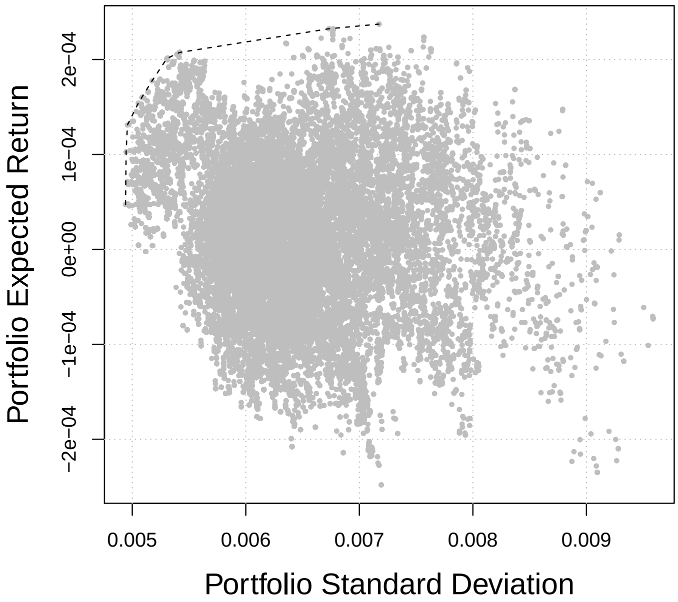


Figure 6.5: Standard deviation-Expected return plot of 5-asset portfolios generated from TDC and fuzzy clustering.

the portfolio obtained by hierarchical clustering (HC Portfolio) and the returns obtained by fuzzy clustering (Fanny Portfolio) are quite similar. Thus, the clustering method does not have a strong influence on the overall results.

6.7 Discussion

In this Chapter, an approach to cluster financial time series according to their asymptotic tail behaviour has been described. The procedure is based on the non-parametric estimation of copula-based lower tail dependence coefficients, used to quantify the extent to which extremely negative events tend to occur simultaneously for pairs of financial assets.

First, a suitable copula-based time series model is chosen (see Chapter 4, Section 4.3) in order to model separately the marginal behaviour of each time series and the link between them; in the second step, the pairwise tail dependence coefficients are estimated via the *block maxima method*. It consists of fitting an extreme

Table 6.17: Composition of the minimum variance portfolios in the convex frontiers of Figure 6.6, selected by hierarchical and fuzzy clustering applied on the lower TDC-based measure.

Hierarchical Clustering					
Assets	F.EI	F.FTEL	F.BSN	D.SAPX	E.IND
Weights	0.41	0.25	0.15	0.14	0.05
Fuzzy Clustering					
Assets	F.EI	H.UNIL	E.TEF	D.RWEX	I.G
Weights	0.43	0.20	0.15	0.13	0.09

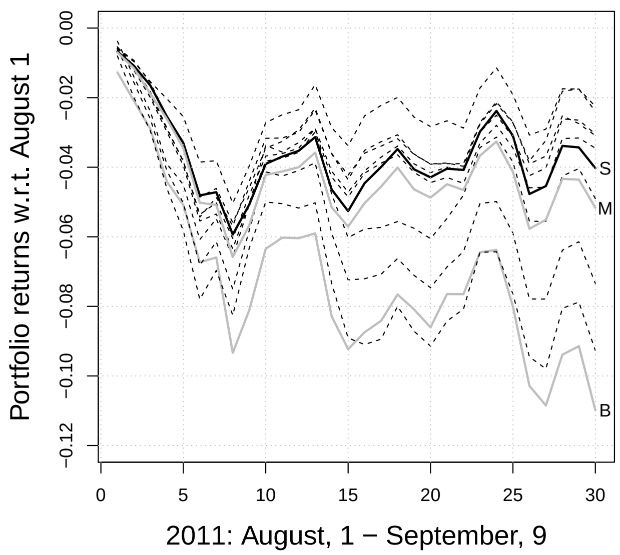
Table 6.18: Basic statistics related to the log–returns of selected minimum variance portfolios in the convex frontier by hierarchical and fuzzy clustering. Period: August 1, 2011 – September 9, 2011.

	Mean	S.D.	Skewness	5% VaR	5% E.S.	Sharpe Ratio
HC Portfolio	0.0001	0.0049	-0.0426	-0.0077	-0.0116	0.0115
FANNY Portfolio	0.0001	0.0051	0.0296	-0.0077	-0.0118	0.0108
Naive MVP	0.0001	0.0048	-0.1527	-0.0076	-0.0114	0.0149
EURO STOXX 50	-0.0001	0.0071	0.0636	-0.0111	-0.0170	-0.0125

value distribution to a sample of maxima over blocks extracted from an underlying series. In particular, the estimation of the dependence function A associated with an extreme value copula provides an estimation for the upper tail dependence coefficients of the transformed returns and, in turn, for the lower tail dependence coefficients of the original time series (Section 6.2). Then, the matrix of pairwise dissimilarities is defined by transforming the tail coefficients matrix, in such a way it can be used for clustering purposes. Two standard procedures are adopted, namely, the agglomerative hierarchical clustering and K-means algorithm. The latter requires a non-metric Multidimensional Scaling procedure to convert the information coming from the dissimilarity matrix into a points configuration in a high-dimensional Euclidean space.

A small simulation study is conducted to check the clustering performances of the proposed methodology, by using different copula models. Results show that the performance of the method is quite promising, even for large dimension of the model. In order to illustrate the proposed approach, we analyse daily returns of time series of Morgan Stanley Capital International Developed Markets indices, designed to measure the equity market performance of developed markets. The results can be compared with the case study in the work by De Luca & Zuccolotto (2011), where a similar procedure is investigated by using a parametric estimation.

The usefulness of the presented methodology is further validated by the application to portfolio selection framework (Section 6.6). The main idea is to select the assets of the portfolio by taking into account the grouping structure given by the clustering solution. A two-stage portfolio diversification strategy is proposed and tested on time series related to EURO STOXX 50 stock index and its components. The results seem to confirm the need for alternative diversification strategies when markets are experiencing a period of simultaneous large losses.



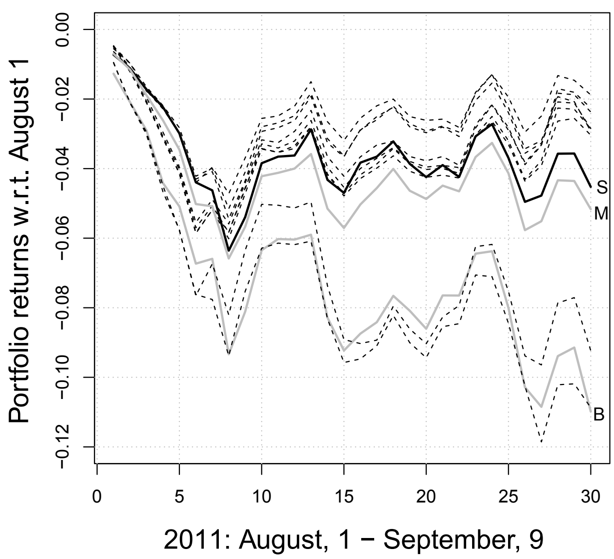


Figure 6.6: *Top* 9 minimum variance portfolios on the efficient frontier of all the possible 5-asset portfolios obtained via hierarchical clustering. *Bottom* 10 minimum variance portfolios on the efficient frontier of all the possible 5-asset portfolios obtained via fuzzy clustering. The returns of the minimum variance portfolio in the frontier (denoted by *S*) and the other portfolios in the frontier (black dotted lines) are compared with the returns of EURO STOXX 50 (denoted by *B*) and with the returns of global minimum variance portfolio(denoted by *M*) composed of all 50 assets.

Chapter 7

Conclusions

Copulas have proved to be useful in several application fields, especially when the major issue is to understand/quantify a risk coming from different sources. In many situations, the global risk strongly depends on the behaviour of the copula in the tails of the distribution. The flexibility in the specification of the dependence structure in the tail represents one of the main advantages of copula models. The main goal of this thesis is to provide new insights into the study of tail dependence of copulas in order to develop tools that may enhance practical applications.

Chapter 3 discusses the theory behind different tail dependence measures that can be expressed in terms of copula functions (whose mathematical background is given in Charter 2). Then, the problem of detecting different tail behaviours is investigated. We analyse several aspects of tail dependence of copulas both at finite and at infinite scale. In particular, the development of graphical copula tools may help in the choice of the relevant copula for the problem at hand, especially when classical goodness-of-it techniques may not be efficient. To this end, a modification of the graphical tool in Michiels & De Schepper (2013) has been proposed, in order to identify which families of copulas are closer to the empirical copula in the tail dependence behaviour. The main ingredient of the suggested graphical tool is the introduction of a suitable dissimilarity measure based on the notion of tail concentration function.

As discussed in the thesis, financial problems often deal with the minimization of the whole risk of a portfolio of assets by adopting some diversification techniques, *i.e.* by investing in assets that do not behave similarly especially in crisis periods. The diversification issue naturally poses the question of investigating the relationship between financial time series and of checking whether they can be grouped together according to some similarity criterion. However, it should be taken into account that in many situations classical correlation-based clustering procedures do not give an accurate indication and understanding of the real dependence between risk exposures, especially when there is some contagion effect among the markets under consideration. Hence, clustering techniques tailored to risk management should adopt alternative procedures, by taking into account the information about the tail behaviour of the involved quantities. In the thesis, it is argued and it is shown that copula models and clustering strategies based on measures of tail dependence could allow to successfully address asset allocation problems.

In Chapter 5 and Chapter 6 two strategies for clustering financial time series are presented, according to a suitable dissimilarity measure that accounts for a kind of extreme (tail) dependence among the markets under consideration. Specifically, the aim consists in creating groups of time series such that elements of each group tend to comove when they are experiencing very large losses. To this end, we adopt two different measures of tail dependence, namely the (lower) tail dependence coefficient and the conditional Spearman's correlation. These measures express two different ways of looking at tail dependence since they focus, respectively, on asymptotic tail dependence and finite tail dependence. From these coefficients, estimated for each pair of time series, a dissimilarity matrix is defined and used as input for classical clustering techniques. As a relevant feature of the proposed procedures, we consider copula-based time series models that allow to separate the univariate behaviour of each time series from the dependence among them (see Section 4.3). Moreover, we adopt non-parametric estimation procedures in modelling the pairwise dependence between the time series, thus avoiding any model assumption. Simulation studies are conducted to check the clustering performances of the two different procedures, with quite promising results. Illustrations from financial datasets are provided, showing the practical implementation of the described techniques. Section 5.7 and 6.7 provide short discussions on some considerations that could be drawn looking at the results and possible advantages in using the suggested procedures.

The last section in Chapter 6 is devoted to an application in portfolio selection framework. Specifically, a two-stage portfolio diversification strategy is proposed, in order to increase the diversification benefits in a bear market. By exploiting tail dependence-based risky measures previously introduced, a first-step cluster analysis is performed for discriminating between assets with the same performance during risky scenarios. Hierarchical and fuzzy techniques are used in the illustration, showing that they perform in a similar way. Then, a mean-variance efficient frontier is computed by fixing a number of assets per portfolio.

Empirical calculations could provide possible evidence that investing on selected index components in trouble periods may improve the risk-averse investor portfolio performance. The proposed portfolio selection procedure is intended to be used by an investor to have more insights into the relationships among different assets in crisis periods. In particular, it may serve to warn against the automatic use of standard portfolio selection procedures that may not work when the markets are expected to experience bearish periods.

References

ABDOUS, B., GHOUDI, K. & KHOUDRAJI, A. (1999). Non-parametric estimation of the limit dependence function of multivariate extremes. *Extremes* 2 245–268.

BASTOS, J. A. & CAIADO, J. (2013). Clustering financial time series with variance ratio statistics. *Quantitative Finance* (in press).

BEIRLANT, J., GOEGEBEUR, Y., SEGERS, J., TEUGELS, J., DE WAAL, D. & FERRO, C. (2006). *Statistics of Extremes: Theory and Applications*. Wiley Series in Probability and Statistics. Wiley.

BERNARD, C., BRECHMANN, E. & CZADO, C. (2013). Statistical assessments of systemic risk measures. In J. P. Fouque & J. A. Langsam, eds., *Handbook on Systemic Risk*. Cambridge University Press, 165–179.

BILLIO, M. & CAPORIN, M. (2009). A generalized dynamic conditional correlation model for portfolio risk evaluation. *Mathematics and Computers in Simulation* 79 2566 – 2578.

BILLIO, M., CAPORIN, M. & GOBBO, M. (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters* 2 123–130.

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31 307 – 327.

BONANNO, G., CALDARELLI, G., LILLO, F., MICCICHÉ, S., VANDEWALLE, N. & MANTEGNA, R. N. (2004). Networks of equities in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems* 38 363–371.

BOUCHAUD, J.-P. & POTTERS, M. (2001). More stylized facts of financial markets: leverage effect and downside correlations. *Physica A: Statistical Mechanics and its Applications* 299 60 – 70.

BRADLEY, B. & TAQQU, M. (2004). An extreme value theory approach to the allocation of multiple assets. *International Journal of Theoretical and Applied Finance* 7 1031–1068.

BRECHMANN, E. C. (2013). Hierarchical Kendall copulas: Properties and inference. *Canadian Journal of Statistics* (to appear).

BRIDA, J. G. & RISSO, W. A. (2010). Hierarchical structure of the German stock market. *Expert Systems with Applications* 37 3846–3852.

CAIADO, J. & CRATO, N. (2010). Identifying common dynamic features in stock returns. *Quantitative Finance* 10 797–807.

CAMPBELL, J. Y., LO, A. W. C. & MACKINLAY, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

CAPÉRAÀ, P., FOUGERES, A.-L. & GENEST, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* 84 567–577.

CESARONE, F., SCOZZARI, A. & TARDELLA, F. (2013). A new method for mean-variance portfolio optimization with cardinality constraints. *Annals of Operations Research* 205 213–234.

CHARRAD, M., GHAZZALI, N., BOITEAU, V. & NIKNAFS, A. (2013). *An examination of indices for determining the number of clusters: NbClust Package*. R package version 1.4, URL `http://CRAN.R-project.org/package=NbClust`.

CHEN, X. & FAN, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135 125–154.

CHERUBINI, U., GOBBI, F., MULINACCI, S. & ROMAGNOLI, S. (2011). *Dynamic Copula Methods in Finance*. John Wiley & Sons, Ltd., Chichester.

CHERUBINI, U., LUCIANO, E. & VECCHIATO, W. (2004). *Copula Methods in Finance*. Wiley Finance Series. Wiley, Chichester.

CLAYTON, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65 141–151.

CONT, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. In *Quantitative Finance*, vol. 1. Institute of Physics Publishing, 223–236.

CORDUAS, M. & PICCOLO, D. (2008). Time series clustering and classification by the autoregressive metric. *Computational Statistics & Data Analysis* 52 1860 – 1872.

CZADO, C. (2010). Pair-copula constructions of multivariate copulas. In P. Jaworski, F. Durante, W. Härdle & T. Rychlik, eds., *Copula Theory and its Applications*, vol. 198 of *Lecture Notes in Statistics - Proceedings*. Berlin Heidelberg: Springer, 93–109.

DANIELSSON, J. (2011). *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. The Wiley Finance Series. Wiley.

DE ANGELIS, L. (2013). Latent class models for financial data analysis: Some statistical developments. *Statistical Methods and Applications* 22 227–242.

DE LUCA, G., RIVIECCIO, G. & ZUCCOLOTTO, P. (2010). Combining random forest and copula functions: A heuristic approach for selecting assets from a financial crisis perspective. *Intelligent Systems in Accounting, Finance & Management* 17 91–109.

DE LUCA, G. & ZUCCOLOTTO, P. (2011). A tail dependence-based dissimilarity measure for financial time series clustering. *Advances in Data Analysis and Classification* 5 323–340.

DEHEUVELS, P. (1978). Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. *Publ. Inst. Statist. Univ. Paris* 23 37.

DOBRIĆ, J., FRAHM, G. & SCHMID, F. (2013). Dependence of stock returns in bull and bear markets. *Dependence Modeling* (in press).

DURANTE, F., FERNÁNDEZ-SÁNCHEZ, J. & SEMPI, C. (2013a). A note on the notion of singular copula. *Fuzzy Sets and Systems* 211 120–122.

DURANTE, F. & FOSCOLO, E. (2013). An analysis of the dependence among financial markets by spatial contagion. *Int. J. Intell. Syst.* 28 319–331.

DURANTE, F., FOSCOLO, E. & SABO, M. (2013b). A spatial contagion test for financial markets. In R. Kruse, M. R. Berthold, C. Moewes, M. n. Gil, P. Grzegorzewski & O. Hryniewicz, eds., *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, vol. 190 of *Advances in Intelligent Systems and Computing*. Springer Berlin Heidelberg, 313–320.

DURANTE, F. & JAWORSKI, P. (2010). Spatial contagion between financial markets: a copula-based approach. *Appl. Stoch. Models Bus. Ind.* 26 551–564.

DURANTE, F. & SEMPI, C. (2010). *Copula Theory: An Introduction*. Copula

Theory and Its Applications. Lecture Notes in Statistics 198. Springer-Verlag Berlin Heidelberg.

EMBRECHTS, P., KLÜPPELBERG, C. & MIKOSCH, T. (1997). *Modelling Extremal Events: For Insurance and Finance*, vol. 33 of *Applications of Mathematics*. Springer.

EMBRECHTS, P., MCNEIL, A. J. & STRAUMANN, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In M. Dempster, ed., *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, 176–223.

ENGLE, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50 987–1007.

ENGLE, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20 339–350.

EVERITT, B. S. (1979). Unresolved problems in cluster analysis. *Biometrics* 169–181.

FAMA, E. F. (1963). Mandelbrot and the Stable Paretian Hypothesis. *The Journal of Business* 36 420–429.

FAMA, E. F. (1965). The behavior of stock-market prices. *The Journal of Business* 38 34–105.

FANG, K., KOTZ, S. & NG, K. (1987). Symmetric Multivariate and Related Distributions. *Chapman&Hall, London* .

FERMANIAN, J.-D. & WEGKAMP, M. H. (2012). Time-dependent copulas. *Journal of Multivariate Analysis* 110 19–29.

FORBES, K. J. & RIGOBON, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance* 57 2223–2261.

FRAHM, G., JUNKER, M. & SCHMIDT, R. (2005). Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance: Mathematics and Economics* 37 80 – 100.

FRANK, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae* 19 194–226.

FRIEDMAN, J., HASTIE, T. & TIBSHIRANI, R. (2009). The elements of statistical learning: Data mining, inference, and prediction. *Springer Series in Statistics* .

GENEST, C. & FAVRE, A.-C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *J. of Hydrologic Engineering* 347–368.

GENEST, C. & SEGERS, J. (2009). Rank-based inference for bivariate extreme-value copulas. *The Annals of Statistics* 37 2990–3022.

GLOSTEN, L. R., JAGANNATHAN, R. & RUNKLE, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48 1779–1801.

GORDON, A. D. (1999). Classification. *Monographs on Statistics and Applied Probability* 82.

GUDENDORF, G. & SEGERS, J. (2010). Extreme-value copulas. In P. Jaworski, F. Durante, W. Härdle & T. Rychlik, eds., *Copula Theory and its Applications*, vol. 198 of *Lecture Notes in Statistics - Proceedings*. Berlin Heidelberg: Springer, 127–145.

GUMBEL, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* 55 698–707.

HÄRDLE, W. & SIMAR, L. (2012). *Applied multivariate statistical analysis*. Springer, 3rd ed.

HOUGAARD, P. (1986). A class of Multivariate failure time distributions. *Biometrika* 73 671–678.

HUA, L. & JOE, H. (2012). Tail comonotonicity: Properties, constructions, and asymptotic additivity of risk measures. *Insurance: Mathematics and Economics* 51 492–503.

HUBERT, L. & ARABIE, P. (1985). Comparing partitions. *Journal of Classification* 2 193–218.

HUI, T.-K. (2005). Portfolio diversification: a factor analysis approach. *Applied Financial Economics* 15 821–834.

JAWORSKI, P. (2010). Tail behaviour of copulas. In P. Jaworski, F. Durante, W. Härdle & T. Rychlik, eds., *Copula Theory and its Applications*, vol. 198 of *Lecture Notes in Statistics - Proceedings*. Berlin Heidelberg: Springer, 161–186.

JAWORSKI, P. (2013). The limiting properties of copulas under univariate conditioning. In P. Jaworski, F. Durante & W. Härdle, eds., *Copulae in Mathematical and Quantitative Finance*, Lecture Notes in Statistics. Springer Berlin Heidelberg, 129–163.

JAWORSKI, P., DURANTE, F. & HÄRDLE, W. K., eds. (2013). *Copulae in Mathematical and Quantitative Finance*, vol. 213 of *Lecture Notes in Statistics - Proceedings*. Springer Berlin Heidelberg.

JAWORSKI, P., DURANTE, F., HÄRDLE, W. K. & RYCHLIK, T., eds. (2010). *Copula Theory and its Applications*, vol. 198 of *Lecture Notes in Statistics - Proceedings*. Springer Berlin Heidelberg.

JAWORSKI, P. & PITERA, M. (2013). On spatial contagion and multivariate GARCH models. *Applied Stochastic Models in Business and Industry* (in press).

JOE, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis* 46 262–282.

JOE, H. (1997). *Multivariate Models and Dependence Concepts*. Monographs in Statistics and Probability 73. Chapman and Hall, London.

JONDEAU, E. & ROCKINGER, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. *Journal of International Money and Finance* 25 827 – 853.

KAUFMAN, L. & ROUSSEEUW, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. Wiley and Sons Inc., New York.

KENDALL, M. G. (1938). A new measure of rank correlation. *Biometrika* 30 81–93.

KIMBERLING, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae* 10 152–164.

KRUSKAL, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29 1–27.

KRUSKAL, J. B. (1964b). Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 29 115–129.

LIAO, T. (2005). Clustering of time series data - a survey. *Pattern Recognition* 38 1857 – 1874.

LONGIN, F. & SOLNIK, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance* 56 649–676.

MAI, J.-F. & SCHERER, M. (2012). *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, vol. 4 of *Series in quantitative finance*. Imperial College Press London.

MALEVERGNE, Y. & SORNETTE, D. (2006). *Extreme Financial Risks*. Springer-Verlag.

MANDELBROT, B. (1963). The variation of certain speculative prices. *The Journal of Business* 36 392–417.

MANNER, H. & SEGERS, J. (2011). Tails of correlation mixtures of elliptical copulas. *Insurance: Mathematics and Economics* 48 153–160.

MANTEGNA, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems* 11 193–197.

MARKOWITZ, H. (1952). Portfolio selection. *The Journal of Finance* 7 77–91.

MCNEIL, A. J., FREY, R. & EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, New Jersey.

MCNEIL, A. J. & NEŠLEHOVÁ, J. (2009). Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistics* 3059–3097.

MICHIELS, F. & DE SCHEPPER, A. (2008). A copula test space model how to avoid the wrong copula choice. *Kybernetika* 44 864–878.

MICHIELS, F. & DE SCHEPPER, A. (2013). A new graphical tool for copula selection. *Journal of Computational and Graphical Statistics* 22 471–493.

MIKOSCH, T. (2006). Copulas: Tales and facts. *Extremes* 9 3–20.

MILLIGAN, G. W. & COOPER, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50 159–179.

MOORE, D. S. & SPRUILL, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics* 599–616.

NELSEN, R. B. (1991). Copulas and association. In *Advances in Probability Distributions with Given Marginals*. Springer, 51–74.

NELSEN, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. New York: Springer, 2nd ed.

NELSON, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59 pp. 347–370.

OAKES, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B (Methodological)* 44 414–422.

OTRANTO, E. (2008). Clustering heteroskedastic time series by model-based procedures. *Computational Statistics* & *Data Analysis* 52 4685 – 4698.

PANTON, D. B., LESSIG, V. P. & JOY, O. M. (1976). Comovement of international equity markets: a taxonomic approach. *Journal of Financial and Quantitative Analysis* 11 415–432.

PATTARIN, F., PATERLINI, S. & MINERVA, T. (2004). Clustering financial time series: an application to mutual funds style analysis. *Computational Statistics & Data Analysis* 47 353 – 372.

PATTON, A. (2009). Copula–based models for financial time series. In T. Mikosch, J.-P. Kreiß, R. A. Davis & T. G. Andersen, eds., *Handbook of Financial Time Series*. Springer Berlin Heidelberg, 767–785.

PATTON, A. J. (2012). A review of copula models for economic time series. *Journal of Multivariate Analysis* 110 4 – 18.

PATTON, A. J. (2013). Copula methods for forecasting multivariate time series. In G. Elliott & A. Timmermann, eds., *Handbook of Economic Forecasting*, vol. 2. Elsevier, Oxford, 899–960.

PICCOLO, D. (1990). A distance measure for classifying ARIMA models. *Journal of Time Series Analysis* 11 153–164.

PICKANDS, J. (1981). Multivariate extreme value distributions. In *Proceedings 43rd Session International Statistical Institute*, vol. 2. 859–878.

RAND, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association* 66 846–850.

RÉMILLARD, B. (2010). Goodness-of-fit tests for copulas of multivariate time series. *Available at SSRN 1729982* .

SALVADORI, G., DE MICHELE, C., KOTTEGODA, N. & ROSSO, R. (2007). *Extremes in Nature: An Approach to Using Copulas*, vol. 56 of *Water Science and Technology Library*. Springer.

SCHMID, F. & SCHMIDT, R. (2007). Multivariate conditional versions of Spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis* 98 1123 – 1140.

SCHMID, F., SCHMIDT, R., BLUMENTRITT, T., GAISSER, S. & RUPPERT, M. (2010). Copula-based measures of multivariate association. In P. Jaworski, F. Durante, W. Härdle & T. Rychlik, eds., *Copula Theory and its Applications*, vol. 198 of *Lecture Notes in Statistics - Proceedings*. Berlin Heidelberg: Springer, 209–236.

SCHWEIZER, B. & SKLAR, A. (1974). Operations on distribution functions not derivable from operations on random variables. *Studia Mathematica* 52 43–52.

SCHWEIZER, B. & SKLAR, A. (1983). Probabilistic metric spaces. North-Holland Series in Probability and Applied Mathematics. North-Holland Publishing Co., New York, NY.

SCHWEIZER, B. & WOLFF, E. F. (1976). Sur une mesure de dépendance pour les variables aléatoires. *C. R. Acad. Sci. Paris Sér. A-B* 283 Aiv, A659–A661.

SCHWEIZER, B. & WOLFF, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics* 879–885.

SIBUYA, M. (1959). Bivariate extreme statistics, I. *Annals of the Institute of Statistical Mathematics* 11 195–210.

SKLAR, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Universite de Paris* 8 229–231.

SKLAR, A. (1996). Random variables, distribution functions, and copulas: a personal look backward and forward. In *Distributions with Fixed Marginals and Related Topics*, vol. 28. Institute of Mathematical Statistics, 1–14.

TOLA, V., LILLO, F., GALLEGATI, M. & MANTEGNA, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control* 32 235 – 258.

TSAY, R. S. (2005). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics. Wiley.

VENTER, G. G. (2002). Tails of copulas. In *Proceedings of the Casualty Actuarial Society*, vol. 89. 68–113.

WILKS, S. S. (1948). Order statistics. *Bulletin of the American Mathematical Society* 54 6–50.

Roberta Pappadà

CURRICULUM VITAE

Contact Information

University of Padova
Department of Statistics
via Cesare Battisti, 241-243
35121 Padova. Italy.

Tel. +39 049 827 4111
e-mail: pappada@stat.unipd.it
roberta.pappada84@gmail.com

Current Position

Since January 2011 (expected completion: February 2014)

PhD Student in Statistical Sciences, University of Padova.

Thesis title: Copula-based measures of tail dependence with Applications

Supervisor: Prof. Nicola Torelli

Co-supervisor: Dr. Fabrizio Durante.

Research interests

- Copula models for time series
- Multivariate dependence models with copulas
- Financial risk measurement
- Extreme-value Theory (applications in finance and hydrology)

Education

February 2007 – December 2009
MSc in Mathematics.
University of Salento, Department of Mathematics
Title of dissertation: "Bivariate Extreme Value Copulas"
Supervisor: Dr. Gianfausto Salvadori
Final mark: 110/110 (with honours)

September 2003 – February 2007
BSc in Mathematics and Informatics.
University of Salento, Department of Mathematics
Title of dissertation: "Dynamic and Stochastic routing problems"
Supervisor: Prof. Chefi Triki
Final mark: 110/110 (with honours)

Visiting periods

April 2013 – July 2013
Center for Mathematics, Technische Universität München
Munich, Germany .
Supervisor: Prof. Claudia Czado

Computer skills

- C ++, R, Stata, MATLAB, Mathematica, Octave, SQL, Java
- Microsoft Office applications, LYX, TEX (LATEX, BIBTEX)
- Microsoft Windows, Linux

Language skills

Italian: native language; English: fluent; French: moderate; Spanish: basic

Publications

F. Durante, R. Pappadà and N. Torelli (2013) Clustering of financial time series in risky scenarios. Advances in Data Analysis and Classification (DOI: 10.1007/s11634-013-0160-4) Published online: 22 December 2013.

F. Durante, R. Pappadà and N. Torelli (2013) Clustering financial time series by measures of tail dependence. In T. Minerva, I. Morlini, and F. Palumbo, editors, *CLADAG 2013. 9th Meeting of the Classification and Data Analysis Group. Book of Abstracts* (ISBN: 978-88-6787-117-9), 4 pages. CLEUP, 2013.

F. Durante, R. Pappadà and N. Torelli (2012) Clustering of financial time series in risky scenarios. In *Abstracts of 6th CSDA International Conference on Computational and Financial Econometrics (CFE 2012) and 5th International Conference of the ERCIM Working Group on Computing & Statistics (ERCIM 2012)* (ISBN: 978-84-937822-2-1), p. 96. Conference Center "Ciudad de Oviedo" (Spain), 1-3 December 2012.

F. Durante and R. Pappadà (2012) Clustering of financial time series in extreme scenarios. In *Atti della XLVI Riunione Scientifica Società Italiana di Statistica* (ISBN: 978-88-6129-882-8), 4 pages. Rome (Italy), 20-22 June 2012. CLEUP, Padova, 2012.

Submitted Papers

F. Durante, J. Fernández-Sánchez and R. Pappadà (2014) Copulas, diagonals and tail dependence. Submitted.

F. Durante, E. Foscolo, R. Pappadà, H. Wang (2013) A portfolio diversification strategy via tail dependence measures. Submitted.

F. Durante, R. Pappadà and N. Torelli (2013) Clustering of extreme observations via tail dependence estimation. Submitted.

Conference presentations

Pappadà, R. (2013) Clustering financial time series by measures of tail dependence. (oral) *CLADAG 2013. 9th Meeting of the Classification and Data Analysis Group.*, Modena, Italy, September 18–20.

Pappadà, R. (2012) Clustering of financial time series in risky scenarios. (oral) *ERCIM 2012. 5th International Conference of the ERCIM WG on COMPUTING & STATISTICS.*, Oviedo , Spain, December 1–3.

Pappadà, R. (2012) Clustering of financial time series in extreme scenarios. (oral) *46th Scientific Meeting of the Italian Statistical Society*, Rome, Italy, June 20–22.

References

Dr. Fabrizio Durante
School of Economics and Management, Free University of Bozen - Bolzano
Piazza Università, 1, I-39100 Bolzano, Italy
Phone: +39 0471 013493
e-mail: fabrizio.durante@unibz.it

Prof. Nicola Torelli
Department of Economics, Business, Mathematics and Statistics
University of Trieste
Via Valerio, 4/1 34127 Trieste, Italy
Phone: +39 40 558 7032
e-mail: nicola.torelli@econ.units.it

Prof. Claudia Czado
Center for Mathematics, Technische Universität München
Parkring 13, D-85748 Garching bei München, Munich Germany
Phone: +49 89 289 17428
e-mail: cczado@ma.tum.de